Bias-corrected estimation for speculative bubbles in stock prices

Robinson Kruse*
University of Cologne and CREATES, Aarhus University

Hendrik Kaufmann
Quoniam Asset Management, Frankfurt am Main

Christoph Wegener
IPAG Business School, Paris and Center for Risk and Insurance, Hannover

April 5, 2018

Abstract

We provide a comparison of different finite-sample bias-correction methods for possibly explosive autoregressive processes. We compare the empirical performance of the downward-biased standard OLS estimator with an OLS and a Cauchy estimator, both based on recursive demeaning, as well as a second-differencing estimator. In addition, we consider three different approaches for bias-correction for the OLS estimator: (i) bootstrap, (ii) jackknife and (iii) indirect inference. The estimators are evaluated in terms of bias and root mean squared errors (RMSE) in a variety of practically relevant settings. Our findings suggest that the indirect inference method clearly performs best in terms of RMSE for all considered levels of persistence. In terms of bias-correction, the jackknife works best for stationary and unit root processes, but with a typically large variance. For the explosive case, the indirect inference method is recommended. As an empirical illustration, we reconsider the "dot-com bubble" in the NASDAQ index and explore the usefulness of the indirect inference estimator in terms of testing, date stamping and calculations on overvaluation.

Key words: Explosive behavior · Bias-correction · Indirect inference · Bubbles

JEL classification: C13, C22, G12

***Corresponding author:** University of Cologne, Faculty of Management, Economics and Social Science, Albertus-Magnus-Platz, 50923 Cologne, Germany, e-mail address: `kruse-becher@wiso.uni-koeln.de` and CREATES, Aarhus University, Department of Economics and Business, Fuglesangs Allé 4, DK-8210 Aarhus V, Denmark. E-mail address: `rkruse@creates.au.dk`

1 Introduction

Estimating the persistence of financial and economic time series is a long standing issue in applied econometrics. The most common framework for assessing the persistence is the autoregressive model. A major practical problem is the inherent downward bias of the conventional OLS estimator. Its bias increases along two dimensions: a small sample size and a true autoregressive parameter in the vicinity of unity are disadvantageous. Given a relatively small sample size arising from e.g. sample splitting, rolling windows, low frequency or data availability, it is a complicated task to estimate the persistence if the process (i) is either stationary, but highly persistent, (ii) exhibits a unit root or (iii) behaves explosively. The bias function is highly nonlinear and changes its derivative in the local-to-unity region.

In finance and economics, it is a well established fact that most time series are characterized by high persistence and stochastic trends, see e.g. Nelson and Plosser (1982) and Schotman and van Dijk (1991). During periods of bubbles (or crises) some financial and economic time series are likely to exhibit even explosive behavior. Recent examples for time series with temporary explosive roots are stock prices (Phillips et al., 2011), house prices (Caspi, 2016, Engsted et al., 2016 and Shi, 2017), exchange rates (Steenkamp, 2017, Hu and Oxley, 2017) and commodity prices (Gutierrez, 2013, Etienne et al., 2014 and Figuerola-Ferretti et al., 2015). Moreover, there is also evidence in art markets (Kräussl et al., 2016, Assaf, 2017) and target balances (Potrafke and Reischmann, 2014).

This work compares several bias-correction estimators and techniques by means of a Monte Carlo study. Among these are the (i) Cauchy estimator by So and Shin (1999a) which builds upon another (ii) recursive mean adjusted estimator, see So and Shin (1999b). The Cauchy estimator is approximately median-unbiased for unit root and explosive processes. This property makes it attractive in comparison to the classic analytic median bias-corrections proposed in Andrews (1993), Andrews and Chen (1994) and Roy and Fuller (2001) which rule out explosiveness by construction. We also consider a recently proposed (iii) estimator based on second differencing by Chen and Kuo (2013). In addition, we consider the (iv) bootstrap, (v) jackknife and (vi) indirect inference approaches. In more detail, we study the bootstrap-based bias-correction procedure by Kim (2003) which builds on Kilian (1998). The jackknife correction (based on Efron, 1979) is recently studied in Chambers (2013), and the indirect inference estimator (see e.g. MacKinnon and Smith, 1998 and Gouriéroux et al., 2000) is suggested by Phillips et al. (2011).

While the main body of the literature focusses on stationary autoregressive models and on the unit root case, mildly explosive behavior has received much less attention. As a consequence, the finite-sample properties of recently suggested techniques are only partly explored and a comprehensive comparison with a focus on explosive roots has not been conducted yet. Due to the nonlinearity in the bias function, it is unclear whether existing recommendations for the stationary and unit root case carry over to the situation with explosive roots. Even some bias-corrected estimators exclude explosive behavior by construction. For a practitioner it is important to know which estimator (if any) performs best overall, i.e. for stationary, unit root and explosive processes. This paper intends to fill these gaps and to provide recommendations for practical applications.

In our Monte Carlo study, we evaluate the performance of the estimators by means of bias and root mean squared errors (RMSE) in small samples. We find that the indirect inference estimator performs very well in terms of RMSE and routinely outperforms its competitors. The recommendations for bias-correction (without considering the variance of estimators) are more diverse. We distinguish situations where the practitioner (i) aims at using a robust method with balanced performance across the levels of persistence or (ii) can either rule out stationarity or explosiveness a priori. Regarding (i), indirect inference and bootstrap are the best robust choices in terms of bias-reduction. In case (ii) the jackknife is highly recommendable in absence of explosiveness, while the bootstrap and indirect inference perform very well for explosive series. Interestingly, these two methods rank second and third for stationary and unit root series which makes their use advisable in case (i). It is worthwhile to emphasize that the indirect inference estimator is the clear winner in terms of RMSE overall.

We provide an empirical illustration to the real NASDAQ composite price index between 1973:01 and 2005:06. This time span should cover explosive regimes due to the presence of the "dot-com bubble". The data has been intensively studied in the academic literature (see, for example Phillips, Wu, and Yu, 2011; Homm and Breitung, 2012; Harvey, Leybourne, and Sollis, 2017) and facilitates to emphasize the relevance of precise persistence estimation. For ease of comparison, we consider the OLS and indirect inference estimator in a rolling window fashion in order to test for speculative bubbles, date stamp the start and the end of the "dot-com bubble" and to estimate the bubble growth rate.

The paper is organized as follows. Section 2 describes the different estimators and bias-correction techniques. We present our simulation results in Section 3. The empirical illustration is given in Section 4. Conclusions are drawn in Section 5.

2 Finite-sample bias-corrections

2.1 Bias of the OLS estimator

The inherent bias of the OLS estimator in autoregressive models is our point of departure. The complicated estimation of autoregressive processes in finite samples sparked a fruitful area of research, see e.g. Kiviet and Phillips (2012) for a recent survey on the vast literature.[1] We focus on the possibility of mild explosive behavior in a simple and widely applied autoregressive

[1] Kendall (1954), Shaman and Stine (1988), Tjøstheim and Paulsen (1983), Tanaka (1984) and Abadir (1993) provide analytic derivations of asymptotic expansions which can be used for bias-correction, see also Abadir (1995) for the context of unit root testing. Bao and Ullah (2007) provide general results on the second-order bias and the mean squared error.

framework with unknown mean. For ease of presentation, we consider a simple first-order model, but our discussion extends to higher-order models in a straightforward way. It is given by

$$y_t = \mu + \rho y_{t-1} + \varepsilon_t \ , \tag{1}$$

where ε_t is a zero-mean white noise process with variance σ^2. We consider the cases of stationarity and unit roots, i.e. $|\rho| < 1$ and $\rho = 1$, and the case where ρ satisfies $\rho = 1 + c/T$, with $c > 0$ and T being the sample size. In the latter case, the autoregressive parameter is local-to-unity in the sense that $\rho \to 1$ as $T \to \infty$. For finite T, ρ deviates moderately from unity.[2]

It is well known that the OLS bias in the given AR(1) model depends on the true value of the autoregressive parameter ρ and the sample size T. For instance, Bao (2007) uses Nagar-type expansions to show that the expected value of the OLS estimator $\widehat{\rho}$ has the form[3]

$$E\left(\widehat{\rho}\right) = \rho - \frac{1+3\rho}{T} + \frac{1}{T^2}\left[\frac{3\rho - 9\rho^2 - 1}{1-\rho} + \frac{\mu^2}{\sigma^2}\frac{1+3\rho}{(1-\rho)^2}\right] + o\left(T^{-2}\right) \ .$$

The smaller the sample size, the more severe is the downward bias. For fixed T, the downward bias is strongest when $\rho \to 1$. To the best of our knowledge, no analogous approximations have been yet derived for the explosive case with unknown mean. An exception is Phillips (2012), where the author shows that the bias behaves as $O(\rho^{-T})$ for the special case when $\mu = 0$ and fixed $\rho > 1$.[4] As we are mainly interested in the practically more relevant case with an unknown mean, we focus on numerical bias-correction methods in the following. However, the analytic bias result for the explosive case with zero intercept already provides several interest insights: the bias function is highly nonlinear and changes its derivative in the local-to-unity region, see also MacKinnon and Smith (1998) for early experimental evidence which we replicate by some introductory simulations in Figure 1.

— FIGURE 1 ABOUT HERE —

Figure 1 shows the bias of the OLS estimator $\widehat{\rho}$ for ρ in $y_t = \mu + \rho y_{t-1} + \varepsilon_t$ for two different sample sizes of $T = 25$ and $T = 50$.[5] It can be seen that the bias reduces much quicker for explosive processes as ρ increases. It approaches zero for large (in comparison to the sample size) values of ρ. But, the estimation of mildly explosive processes with roots near unity is still heavily biased. The problem persists for higher-order autoregressive models for which analytic bias formulas depend on the autoregressive lag order.

[2] Asymptotic theory for mildly explosive autoregressions is developed in Phillips and Magdalinos (2007). Wang and Yu (2015) provide asymptotic results for explosive autoregressions.

[3] For the following result, normality is assumed in addition to $\rho < 1$ and $y_0 = 0$. More detailed expressions can be found in Bao (2007) for the case of non-normality and non-zero initialization. Stationarity is, however, required for such Nagar-type expansions.

[4] An analogous (and more complicated) formula for the mildly explosive case with $\rho = 1 + c/T$, $c > 0$ is also provided.

[5] The true autoregressive parameter ρ (on the x-axis) ranges from 0.6 to 1.2. For simplicity, we set $\mu = 0$, $y_0 = 0$ and $\varepsilon_t \overset{iid}{\sim} N(0,1)$.

Our simulation study in Section 3 takes several important dimensions of bias reduction into account. Beside a broad comparison of estimators and bias-correction techniques (as discussed in the following subsections), we consider the uncertainty of not knowing whether the series is in fact stationary, unit root non-stationary or explosive. In the latter case, it obviously plays an important role *how* explosive the series is. As this is ultimately unknown, it is desirable to use some *robust* estimator in practice. Only in some situations one may reasonably rule out certain cases by assumption. Moreover, we shall take the price of bias reduction in terms of possibly inflated variances into account by additionally comparing root mean squared errors. We proceed by presenting three alternative estimators. These include the recursively demeaned OLS, the Cauchy and the second difference-based estimator. Thereafter, we turn our attention to three bias-correction techniques for the standard OLS estimator. First, a bootstrap method by Kim (2003) is discussed, followed by a simple jackknife approach advocated by Chambers and Kyriacou (2013) and finally, we describe the indirect inference approach used in Phillips et al. (2011).

2.2 Alternative estimators

2.2.1 Recursive mean adjustment

An alternative way to handle the deterministic component μ is to consider recursive mean adjustment, see So and Shin (1999b). OLS estimation of the autoregressive parameters is then applied to the series $\bar{y}_t = t^{-1}\sum_{s=1}^{t} y_s$. The ordinary mean adjustment scheme leads to a finite-sample correlation of the regressors with the error term which is a source of the bias. By applying a recursive mean adjustment, this problem is mitigated. Simulation results in So and Shin (1999b) suggest that the bias and the mean squared error can be improved substantially. The resulting estimator is labeled as $\widehat{\rho}^R$:

$$\widehat{\rho}^R = \frac{\sum_{t=2}^{T}(y_t - \bar{y}_{t-1})(y_{t-1} - \bar{y}_{t-1})}{\sum_{t=2}^{T}(y_{t-1} - \bar{y}_{t-1})^2} \,. \tag{2}$$

The standard OLS estimator $\widehat{\rho}$ only differs from $\widehat{\rho}^R$ in that the OLS estimator uses $\bar{y}_T$ instead of $\bar{y}_{t-1}$.

2.2.2 Cauchy estimator

The Cauchy estimator, see So and Shin (1999a), builds upon $\widehat{\rho}^R$ as it uses recursive mean adjustment as well. The estimator for an AR(1) process is given by

$$\widehat{\rho}^C = \frac{\sum_{t=2}^{T}(y_t - \bar{y}_{t-1})\mathsf{sign}(y_{t-1} - \bar{y}_{t-1})}{\sum_{t=2}^{T}|y_{t-1} - \bar{y}_{t-1}|} \tag{3}$$

where $\mathsf{sign}(x_t) = 1$ if $x_t \geq 0$ and $\mathsf{sign}(x_t) = -1$ if $x_t < 0$. This estimator can be interpreted as an instrumental variable estimator where $\mathsf{sign}(y_{t-1} - \bar{y}_{t-1})$ serves as the instrument. Moreover, it can also be seen as a weighted LS estimator with weights $w_t = |y_{t-1} - \bar{y}_{t-1}|^{-1}$. This interpretation is based on the fact that $\mathsf{var}(y_t|y_{t-1},...) = \sigma w_t^{-1}$. The Cauchy estimator $\widehat{\rho}^C$ exhibits some nice properties. It is approximately median-unbiased for all values of ρ whereas the OLS estimator is median-biased.[6] Another important feature is that the theoretical foundation of estimator allows for explosive autoregressive processes. This is in contrast to the approximately median-unbiased estimators by Andrews (1993), Andrews and Chen (1994) and Roy and Fuller (2001), for example.

2.2.3 Second difference-based estimator

Chen and Kuo (2013) suggest an estimator for the autoregressive parameter in AR(1) processes which is based on the second difference of y_t. Their OLS estimator is based on the following transformed regression model

$$2\Delta_2 y_t = \rho\Delta_2 y_{t-1} + \zeta_t,$$

where $\Delta_2 = (1 - L^2)$ with L denoting the lag operator. The error term ζ_t is given as $\zeta_t = 2\Delta_2 y_t - \rho\Delta_2 y_{t-1}$. Importantly, the regressor $\Delta_2 y_{t-1}$ is uncorrelated with the error ζ_t. The estimator for ρ is obtained by applying OLS to the transformed regression model. The resulting estimator reads

$$\widehat{\rho}^{SD} = \frac{2\sum_{t=2}^{T}\Delta_2 y_t \Delta_2 y_{t-1}}{\sum_{t=2}^{T}\Delta_2 y_{t-1}^2}.$$

The second difference-based estimator $\widehat{\rho}_{SD}$ exhibits the asymptotic property that $\sqrt{T}(\widehat{\rho}_{SD} - \rho)$ is distributed as a Normal random variable with zero mean and a variance of two for $-1 < \rho \leq 1$. As the authors argue, the estimator is inconsistent for explosive processes as $\Delta_2 y_t$ is non-stationary, but it maintains its properties for mildly explosive processes for which $\rho^{2T}/\sqrt{T} \to 0$ holds. In fact, this only allows for very mild deviations in the direction of explosiveness. Interestingly, the simulation results in Chen and Kuo (2013) exclude the mildly explosive case. Mean adjustment does not play a role for this estimator as the constant deterministic term already vanishes after taking the first difference.[7]

2.3 Bias-correction techniques

2.3.1 Bootstrap

The first bias-correction technique we consider is the bootstrap. Kim (2003) proposes the use of the bootstrap for a correction of the OLS bias which builds upon Kilian (1998). The bias of the OLS estimator can be estimated as follows: Estimate the model $y_t = \mu + \rho y_{t-1} + u_t$ via OLS and obtain the estimates $\widehat{\theta} = (\widehat{\mu}, \widehat{\rho})'$. Generate a pseudo-data set $\{y_t^b\}_{t=1}^T$ based on these estimates according to

$$y_t^b = \widehat{\mu} + \widehat{\rho} y_{t-1}^b + u_t^b \ ,$$

where u_t^b is a random draw with replacement from the OLS residuals $\{\widehat{u}_t\}_{t=1}^T$. B sets of pseudo-data are generated. Each pseudo-data set gives a bootstrap parameter estimate $\widehat{\theta}^b = (\widehat{\mu}^b, \widehat{\rho}^b)'$ by estimating the model $y_t^b = \mu + \rho y_{t-1}^b + v_t$, $b = 1, \ldots, B$. We obtain the sequence $\{\widehat{\theta}^b\}_{b=1}^B$ and the

[6] An estimator $\widetilde{\rho}$ for ρ is said to be median-unbiased if $P(\widetilde{\rho} \geq \rho) \geq 1/2$ and $P(\widetilde{\rho} \leq \rho) \geq 1/2$.
[7] Chen and Kuo (2013) also discuss the case of time trends.

average bias of $\widehat{\theta}^b$ is estimated as $\widetilde{\theta} - \widehat{\theta}$, where $\widetilde{\theta}$ is the sample average of $\{\widehat{\theta}^b\}_{i=1}^B$, i.e.

$$\widetilde{\theta} \equiv \frac{1}{B} \sum_{b=1}^{B} \widehat{\theta}^b \, .$$

Using this bootstrap-based estimator for the bias, a bias-correction for $\widehat{\theta}$ can be directly obtained via

$$\widehat{\theta}^B = \widehat{\theta} - \left(\widetilde{\theta} - \widehat{\theta}\right) = 2\widehat{\theta} - \widetilde{\theta} \, .$$

This estimator computes the OLS estimation bias for a process with parameter values $\widehat{\theta}$ and uses this bias as approximation for the true bias of $\widehat{\theta}$. For further details regarding this estimator, the reader is referred to Kim (2003). The validity of the bootstrap procedure for the explosive case (with finite error variance) is given in Basawa et al. (1989). Datta (1995) shows that the bootstrap also works for heavy-tailed innovations. The unit root case is treated in Basawa et al. (1991), Datta (1996) and extended in Inoue and Kilian (2002) to general deterministic terms and higher-order autoregressive models.[8]

2.3.2 Jackknife

In the following, we consider sub-sample jackknife procedures. Suppose that the full sample Y is divided into m sub-samples Y_j of same length l, $j = 1, \dots, m$, and let $\widehat{\theta}^j$ be the OLS estimate for θ in sub-sample Y_j. Then, the jackknife statistic

$$\widehat{\theta}^J = \left(\frac{T}{T-l}\right)\widehat{\theta} - \left(\frac{l}{T-l}\right)\widetilde{\theta}$$

where $\widetilde{\theta} = \frac{1}{m}\sum_{j=1}^m \widehat{\theta}^j$ satisfies $E(\widehat{\theta}^J) = \theta + O(T^{-2})$ and is thus able to reduce the bias. Chambers (2013) proposes and compares various jackknife techniques to reduce the small sample bias.[9] Here, we focus on one of the methods in the comparison of Chambers (2013): the non-overlapping sub-samples jackknife. This estimator has good bias-correction properties without the considerable increase of the RMSE of higher-order jackknife estimators. The time series is splitted in m non-overlapping sub-samples,

$$Y_j = (y_{[(j-1)T/m+1]}, \dots, y_{[jT/m]})', \qquad j = 1, \dots, m.$$

In the following, we work with $m = 2$ sub-samples, because the procedure with this particular choice of m has the best bias-correction properties according to Table 1 in Chambers (2013). This simplifies the jackknife statistic to

$$\widehat{\theta}^J = 2\widehat{\theta} - \widetilde{\theta} \, .$$

The intuition behind this approach is almost the same as in the bootstrap approach of Kim (2003). The average bias in the sub-samples is higher because of the smaller sample size and therefore a bias-reduction is induced. The difference to the bootstrap procedure is that the

[8]See also Berkowitz and Kilian (2000) for a survey on bootstrap techniques for time series.
[9]See also Chen and Yu (2015) for further developments.

average bias is calculated on sub-samples of the true process and not on pseudo-data.[10] The introduced jackknife procedure is only valid as long as the process is stationary, see Chambers (2013). The unit root case is tackled in Chambers and Kyriacou (2018). To the best of our knowledge, the (mildly) explosive case has not been under consideration so far.

2.3.3 Indirect inference

We now turn to a simulation-based estimator relying on the concept of indirect inference. Indirect inference estimators have a long tradition, see e.g. Gouriéroux et al. (1993) and Smith (1993). The indirect inference estimator suggested in Phillips et al. (2011) allows for explosiveness, see also Phillips (2012) for a recent contribution on its limit theory. The following exposition draws heavily from Phillips et al. (2011) and starts with the OLS estimator $\widehat{\rho}$. Consider a set of simulated series with AR(1) coefficient equal to some ρ, i.e. $\{y_t^h(\rho)\}_{h=1}^H$, $h = 1, 2, \ldots, H$. H denotes the total number of available simulation paths.[11] For each $h \in \{1, 2, ..., H\}$, we obtain an OLS estimate labeled as $\widehat{\rho}^h(\rho)$. The indirect inference estimator is then given by

$$\widehat{\rho}_H^{II} = \arg\min_{\rho \in \Theta} \left\| \widehat{\rho} - \frac{1}{H} \sum_{h=1}^{H} \widehat{\rho}^h(\rho) \right\| ,$$

where Θ is a compact parameter space and $\| \cdot \|$ is the L_2-norm. For $H \to \infty$ one obtains

$$\widehat{\rho}^{II} = \arg\min_{\rho \in \Theta} \| \widehat{\rho} - q(\rho) \| ,$$

where $q(\rho) = E\left(\widehat{\rho}^h(\rho)\right)$ is the binding function. Given invertibility of q, the indirect inference estimator results as

$$\widehat{\rho}^{II} = q^{-1}(\widehat{\rho}) .$$

Hence, this estimator uses a grid of possible true values for ρ and the corresponding average OLS estimates $H^{-1} \sum_{h=1}^{H} \widehat{\rho}^h(\rho)$. The estimate $\widehat{\rho}$ is compared to the average OLS estimates. $\widehat{\rho}^{II}$ corresponds to the value which leads to the average OLS estimate with the minimal distance to $\widehat{\rho}$. The finite-sample bias-correction stems from the simulation of $q(\rho)$. whose precision increases with H. Importantly, the indirect inference estimator is applicable even for mildly explosive processes, see Phillips et al. (2011).

[10] Kaufmann and Kruse (2013) have also experimented with a higher-order non-overlapping sub-samples jackknife estimator, i.e. J(2,3) in the notation of Chambers (2013). Our results show that a further bias-reduction (in comparison to the simple jackknife estimator) can be achieved, but that the increase in variance is substantial. Moreover, the simple jackknife already delivers convincing bias-correction. For this reason, we focus on the simple version in the following.

[11] In order to generate $\{y_t^h(\rho)\}_{h=1}^H$, we assume normal errors in the following. This assumption is not crucial as the performance is nearly unaffected even under skewed and fat-tailed distributions, see Kaufmann and Kruse (2013).

3 Finite-sample properties

3.1 Monte Carlo setup

In this section we investigate the properties of various bias-correction methods via Monte Carlo simulation. The foci of this analysis are the bias-reduction and the RMSE of these estimators. The simulation setup is as follows: We consider autoregressive models of the structure

$$y_t = \mu + \rho y_{t-1} + \beta \Delta y_{t-1} + \varepsilon_t$$

with $\varepsilon_t \sim N(0, \sigma^2)$.[12] The autoregressive parameter ρ measures the persistence of y_t and takes values $\rho = \{0.85, 0.9, 0.95, 0.99, 1, 1.01, 1.02, 1.03, 1.04, 1.05\}$. We consider $T = \{25, 50\}$. The degree of explosiveness is mild in our design and corresponds to typical values in financial applications. The intercept μ is specified along the lines of Phillips et al. (2014) in a local-to-zero fashion as follows: $\mu = \widetilde{\alpha} T^{-\eta}$ such that y_t has a small deterministic drift which depends on the sample size T. The localizing parameter $\eta > 0$ determines the strength of the drift and whether it is dominant ($\eta \leq 0.5$) or dominated by the stochastic trend ($\eta > 0.5$). We set $\widetilde{\alpha} = \{0, 1\}$ and $\eta = \{0.4, 0.6\}$. In case of a truly zero intercept, we allow the initial condition to be non-negligible by setting $y_0 = \sqrt{T}$, see Harvey and Leybourne (2014). Otherwise, the initial condition is set equal to zero for non-stationary processes and equal to $y_0 = \varepsilon_0 / \sqrt{1 - \rho^2}$ under stationarity. In line with Phillips et al. (2014), the fitted autoregression contains always an intercept to account for a potential deterministic drift. Regarding higher-order autoregressive models, we set $\beta = -0.3$ and use first-order models for the estimation. Hence, we are considering the important case of an under-specified model.[13] The number of Monte Carlo repetitions is set to $20,000$ for each single experiment. The number of bootstrap repetitions is 499. The binding function for the indirect inference estimator is simulated with $H = 20,000$.

3.2 Bias

We present results for the bias in Figure 2 for $T = 25$ and in Figure 3 for $T = 50$, respectively. In the first place, we simulate a first-order process with a truly zero intercept. Below, we study the extensions with a drift or second-order autocorrelation.

— FIGURES 2–3 ABOUT HERE —

We start by discussing the results for the smaller sample size and comment on differences to the case of $T = 50$ later on. The downward bias of the OLS estimator is strong for all ρ. The

[12] Additional results for stable distributed errors with fat tails and heteroscedastic errors (via a GARCH(1,1) structure) are included in Kaufmann and Kruse (2013). The main insight from the results for these settings is that the performance of estimators is typically unaffected. Bao and Ullah (2007) show that the second-order bias (for $\rho < 1$) is robust against non-normality. As Bao (2007) points out, this result remains only valid up to order $O(T^{-1})$, i.e. $-(1+3\rho)/T$ under a general error distribution. Skewness and kurtosis influence the bias via $O(T^{-2})$ terms.

[13] Kaufmann and Kruse (2013) also compared the estimators in case of correct higher-order specifications and the exact order of the autoregressive model does alter the main findings. Over-fitting of the autoregressive model is not harmful, while under-fitting turns out to be an important issue.

worst performance is obtained for values of ρ in the vicinity of unity. For explosive processes, the bias is less severe, but still of major relevance. Even for $\rho = 1.05$, the downward bias is of such a magnitude that the OLS estimator does not exceed unity on average. Thus, even when the process is explosive, the OLS estimator has an average bias of approximately -0.1 and will thus yield average point estimates around 0.95.

The performance of the OLS estimator with recursive demeaning is more promising over the whole range of ρ, but it is worthwhile to use the Cauchy estimator instead. This estimator is still downward biased. But, it indicates explosiveness at least for $\rho = 1.05$ where the simulated bias equals -0.039. The bootstrap approach performs substantially better than the Cauchy estimator. For quite strongly explosive series, the bias function switches its sign leading to a small upward bias for $\rho \geq 1.04$. For $\rho < 1$, the bootstrap estimator performs similar to the second differencing estimator. The latter has some additional advantages for very persistent stationary processes and for unit roots. For explosive series, it exhibits a somewhat risky behavior: after a certain threshold (here: 1.02), the estimator turns to be severely upward distorted. This is not entirely surprising given the restrictive condition $\rho^{2T}/\sqrt{T} \to 0$. We conclude that the second difference estimator performs reasonably well for stationary and unit root processes, and also for *very* mildly explosive cases. But, for typical values of ρ in financial and economic applications, it seems that the estimator can not be recommended. Somewhat similarly, the jackknife approach performs extremely well for stationary and unit root series. In fact, it offers superb bias-correction abilities at nearly zero computational cost. For explosive series, however, results are less favorable. Finally, the indirect inference estimator appears to offer a well balanced performance over the whole range of ρ. In many cases, the indirect inference estimator ranks second or third with a small bias.

For $T = 50$, all estimators perform better as more information is available. Again, the jackknife approach has the best performance among all approaches for values $\rho \leq 1$. It has bias problems for $\rho > 1$. The indirect inference estimator and the bootstrap approach also work very well, especially for explosive series. The bias of the standard OLS estimator is very close to zero for $\rho = 1.05$. In this case other estimators are less effective (except for the bootstrap): they typically lead to some minor upward distortion.

Results for the case of a deterministic drift are included in Tables 1 and 2. We distinguish the situations in which the drift is either dominant relative to the stochastic component, i.e. $\eta = 0.4$, or when it is dominated ($\eta = 0.6$). As this issue is relevant for non-stationary processes only, we focus on the corresponding subset for the autoregressive parameter, i.e. $\rho = \{1, 1.01, 1.02, 1.03, 1.04, 1.05\}$. In quite many cases, the indirect inference estimator provides the best bias-correction, especially for $T = 25$ and $\rho > 1$. The bootstrap approach is a close competitor and the jackknife provides most accurate estimation for the unit root case, both under negligible and non-negligible drifts. A comparison of the two distinct cases of $\eta = 0.4$ and $\eta = 0.6$ reveals that the standard OLS estimator is performing worse under a negligible drift.

Note that the bias in the case of the dominated drift component ($\eta = 0.6$) is negative – like in the case of the first-order autocorrelation with a truly zero intercept – while it is positive for the dominant drift ($\eta = 0.4$). The reason is that the presence of the non-negligible drift leads to an overestimation of the persistence. Table 3 contains the results for an AR(2) process. Here, we find that the jackknife estimator performs best, even in the explosive region, followed by the indirect inference and the bootstrap approach. In general, all estimators suffer from ignoring the second-order autocorrelation, but notably the poor performance of the OLS estimator can be improved significantly.

— TABLES 1–3 ABOUT HERE —

3.3 Root mean squared error

Another useful measure for evaluating the small-sample performance of competing estimators is the root mean squared error (RMSE). Both components, bias and variance, contribute to this measure. By considering the RMSE, we are able to judge the price in terms of an increased variance for reducing the bias in relative terms. The results for the root mean squared errors (multiplied with T) are shown in Figure 4 for $T = 25$ and in Figure 5 for $T = 50$, respectively. Similar to the analysis of bias, we begin with a first-order process with a truly zero intercept and continue with extensions thereafter.

— FIGURES 4–5 ABOUT HERE —

Again, we start by interpreting the results for $T = 25$. Some clear-cut results can be obtained. The best performing method (with explicit advantages over its competitors) is indirect inference. Our second recommendation would be the bootstrap, while the second difference estimator is dominated by all others even for stationary and unit root processes, the reason being its relatively large variance. The standard OLS estimator also performs better than the jackknife for all values of ρ. The recursively demeaned OLS estimator outperforms the Cauchy estimator, especially for less persistent series. Both techniques are on equal footing for explosive series. The bootstrap has advantages over the recursively demeaned OLS estimator for a broad range of values around unity. In cases of more extreme persistence (either stationary or explosive), it is as good as the former estimator. The indirect inference estimator is the clear winner in this competition as it offers reasonable bias-correction together with a reasonably low variance. We have seen that bias-correction can be improved, but only at high costs. Taking both issues simultaneously into account, we recommend to use $\widehat{\rho}^{II}$ in practice for small sample sizes of $T = 25$. For $T = 50$, the indirect inference estimator remains our recommended procedure. However, the differences to its competitors are less visible for rather explosive series with $\rho \geq 1.03$. For stationary autoregressive series with moderate persistence the bootstrap performs similarly well. The recursively demeaned OLS estimator still has advantages over the Cauchy estimator. For large values of ρ, we find that the OLS estimator performs much better than for less persistent series. The main reason being that the bias vanishes and that its variance is relatively low as well. The second difference estimator is excluded from Figure 5 due to its relatively bad performance.

Results for the case of a deterministic drift are included in Tables 4 and 5. The indirect inference estimator emerges as the overall best approach with lowest RMSE values across all settings. This excellent performance further strengthens our recommendation to apply this estimator for non-stationary series. In Table 6 the results for a second-order process are reported. The performance of all estimators deteriorate by neglecting the second-order autocorrelation also in terms of the RMSE. However, it is notable that the results for the first-order process in Figure 4 and Figure 5 are directly comparable to those presented in Table 6 as they share the same specification for the deterministic term. Like in the case of the first-order process, the indirect inference estimator outperforms all its competitors for both sample sizes and all considered values for the autoregressive parameter. Thus, the indirect inference estimator is also preferable under a neglected second-order autocorrelation.

— TABLES 4–6 ABOUT HERE —

4 NASDAQ composite index

As an illustration of the indirect inference estimator, we study the logarithm of the real NASDAQ composite price index (without dividends) from 1973:01 to 2005:06 with $T = 389$ monthly observations. This series has received much attention in the related literature on bubble testing, but also by policy makers, because it captures the so-called "dot-com bubble". We consider exactly the same data points as other studies (see e.g. Phillips et al. (2011), Gutierrez (2011), Homm and Breitung (2012), Breitung and Kruse (2013) and Harvey et al. (2017)) and employ established bubble testing methodologies, see e.g. Phillips et al. (2011), for ease of comparison. In our econometric approach we use a rolling window version of the right-tailed ADF (Augmented Dickey-Fuller) procedure as it has been employed by Phillips et al. (2011), Gutierrez (2013) and Chong and Hurn (2016). As a novelty, we make use of the indirect inference estimator not only for persistence estimation, but also for a modified unit root test and date stamping of the explosive subperiods.

With regard to the underlying theoretical framework, we start with the no-arbitrage condition

$$P_t = \frac{1}{1+R} E_t(P_{t+1} + D_{t+1}) \tag{4}$$

to introduce the concept of a financial bubble. P_t is the real stock price, D_t is the real dividend and R is a positive and constant discount rate. Recursive substitution until period k yields to

$$P_t = \underbrace{\sum_{i=1}^{k} \frac{1}{(1+R)^i} E_t\left(D_{t+i}\right)}_{F_t} + \underbrace{\frac{1}{(1+R)^k} E_t\left(P_{t+k}\right)}_{B_t}.$$

If the transversality condition $\lim_{k\to\infty} \frac{1}{(1+R)^k} E_t\left(P_{t+k}\right) = 0$ holds, the price consists out of a fundamental part F_t which is a martingale process. Otherwise, the price contains also a bubble component B_t such that P_t is a submartingale process. This distinction motivates to test for a unit root (martingale process) against explosiveness (submartingale process) in stock prices.

A log-linear approximation of equation (4) leads the solution $p_t = p_t^f + b_t$, where the log of the stock price p_t is composed of a fundamental price p_t^f and a bubble component b_t (expressed in natural logarithms). The fundamental value p_t^f is determined by discounted expected future dividends only. The bubble is a submartingale given by $b_t = (1+g)b_{t-1} + v_t$ with $g > 0$ and v_t being a martingale difference sequence. Obtaining an estimate for ρ larger than one implies that the log stock price grows at a rate of $(\widehat{\rho}-1)\%$ per month. Assuming that the log dividends are non-explosive, b_t is explosive and grows with rate $\widehat{g} \geq (\widehat{\rho}-1)\%$. A precise estimate of ρ is thus informative for the growth rate of the bubble process b_t. Given an initial overvaluation of the stock market of $a_0 > 1$, i.e. $P_0 = a_0 P_0^f$, the bubble starts at $b_0 = \log(P_0/P_0^f) = \log(a_0) > 0$. After t periods, the expected level of the bubble component is $b_t = \widehat{\rho}^{\,t} \cdot \log(a_0)$. This leads to an estimate of overvaluation in the *level* of the stock market index P_t given by $\widehat{a}_t \equiv \exp(b_t) = \exp(\widehat{\rho}^{\,t} \cdot \log(a_0))$. Hence, when estimating the overvaluation in the stock prices, a bias-corrected estimator for ρ is clearly advantageous.

Given the temporary nature of bubbles, we consider switches from unit root behavior to explosiveness until a burst of the bubble. In its simplest form, the data generating process is given as

$$p_t = \begin{cases} p_{t-1} + \varepsilon_t, & t < t_e \ , \quad \text{"non-explosive"} \\ \rho p_{t-1} + \varepsilon_t, & t \geq t_e \ , \quad \text{"explosive"} \\ p_{t_c}^* + \sum_{i=[r_e T]+1}^{t} \varepsilon_i, & t \geq t_p \ , \quad \text{"non-explosive with re-intialization".} \end{cases}$$

The explosive regime starts at time t_e with $\rho > 1$ and collapses by jumping to a new initial value $p_{t_c}^*$ at time $t_p > t_e$. Up to time point t_e, the price has a non-explosive unit root. This specification allows also for transitional mean reverting dynamics to $p_{t_c}^*$ and continues as a unit root process (see Phillips and Yu, 2009; Phillips, Wu, and Yu, 2011). We impose the standard assumption of weak dependence for the innovations ε_t and we suppress the intercept, but our empirical testing strategy (see equation 5) takes a potential drift and serial correlation into account.

The auxiliary Dickey-Fuller test regression

$$p_t = \mu + \rho p_{t-1} + \sum_{j=1}^{J} \lambda_j \Delta p_{t-j} + u_t \qquad (5)$$

is used to test the null hypothesis of a random walk throughout the whole sample $(t = 1, 2, ..., T)$ against the alternative of an explosive process, see Phillips et al. (2011). The test regression contains a drift term μ and J lags of the differenced series $(\Delta p_{t-j}; j \geq 1)$ to account for additional serial correlation beyond the first lag.

In order to account for time-variation, equation (5) is estimated by a rolling window scheme with $w_0 = [r_0 T]$ observations per window. Rolling window is advocated in Gutierrez (2011) and Chong and Hurn (2016) compared the recursive approaches (see Phillips et al., 2011, 2014). Their simulation evidence shows that rolling window tests yield higher detection rates, especially for mild explosive roots, i.e. $\rho = 1 + c/T$, which are more difficult to detect for small values of the localizing parameter $c > 0$. We contribute to this literature by exploring the role of bias-correction via indirect inference techniques. Due to rolling estimation with relatively few observations per window, bias-correction plays an important role as demonstrated in our Monte Carlo study.

For a given window size w, we compute the standard ADF t-statistic (t_w) and its counterpart using the bias-corrected indirect inference estimator (t_w^{II}):

$$t_w = \frac{\widehat{\rho}_w - 1}{\widehat{\sigma}_w} \quad \text{and} \quad t_w^{II} = \frac{\widehat{\rho}_w^{II} - 1}{\widehat{\sigma}_w^{II}} \tag{6}$$

with $\widehat{\sigma}_w$ and $\widehat{\sigma}_w^{II}$ as the estimated standard deviation of the OLS estimator $\widehat{\rho}_w$ and the indirect inference estimator $\widehat{\rho}_w^{II}$, respectively. The upward bias-correction implies larger values of the numerator of the t_w^{II}-statistic which is not offset by an increase of the standard deviation by a similar amount. Hence, a new set of critical values is needed and simulations with 10,000 replications and a sample size set equal to the rolling window size of fifty observations provide those. They are equal to 0.88, 1.14 and 1.71 for the nominal significance level of ten, five and one percent, respectively. Clearly, they exceed those for the conventional t-statistic (-0.44, -0.08, 0.60) by some extent, see Phillips et al. (2011). The break date estimators $\widehat{r}_e$ (for the start of the bubble) and $\widehat{r}_p$ (for its collapse) are given in the OLS case by

$$\widehat{r}_e = \inf_{w \geq w_0} \{w : t_w > cv_\alpha(w_0)\} \text{ and } \widehat{r}_p = \inf_{w \geq \widehat{r}_e + \log(n)/n} \{w : t_w < cv_\alpha(w_0)\} \tag{7}$$

and in the indirect inference case by

$$\widehat{r}_e^{II} = \inf_{w \geq w_0} \left\{w : t_w^{II} > cv_\alpha^{II}(w_0)\right\} \text{ and } \widehat{r}_p^{II} = \inf_{w \geq \widehat{r}_e^{II} + \log(n)/n} \left\{w : t_w^{II} < cv_\alpha^{II}(w_0)\right\} \tag{8}$$

with α as the significance level.

— FIGURE 6 ABOUT HERE —

Figure 6 shows results generated by the rolling window approach using indirect inference estimation. We use $r_0 = 0.13$ such that $w_0 = 50$ which is a sample size included in our Monte Carlo study. The nominal significance level is set to five percent ($cv_{0.05}^{II}(50) = 1.14$; red thick solid line in Figure 6) and we use a first-order model in our rolling window estimations for simplicity (similar to Phillips et al. (2011) when computing the indirect inference estimator).

The rolling window indirect inference t_w^{II}-statistics are significant in the late Nineties in conjunction with the well-established evidence for the "dot-com bubble". The date stamping routine identifies an explosive period from $\widehat{r}_e^{II}$ = October 1995 to $\widehat{r}_p^{II}$ = October 2000 (see the vertical dotted lines in Figure 6). Phillips et al. (2011) find the bubble regime to last from July 1995 to September 2000 when using rolling windows with 77 observations. Figure 6 also shows the results for conventional t_w-statistics based on the rolling OLS estimator with 50 observations per window (similar to the displayed indirect inference statistics t_w^{II}). The t_w-statistics are fluctuating relatively erratic around their critical value during the late Nineties which can be explained by the inherent downward bias. Overall, the trajectory for the t_w^{II}-statistics appears more steady which might be seen as a prevention against misinterpretations in favour of a collapsing bubble.

We estimate the autoregressive root for the subsample from October 1995 to October 2000. It is expected that this estimate exceeds unity such that it reflects the explosive bubble in the NASDAQ price index. Notably, the OLS estimator yields 0.989, while the indirect inference estimator yields 1.046. Hence, the OLS estimator does not even indicate an explosive behavior, but rather suggests a root in the stationary local-to-unity region. As a direct implication, the OLS estimate cannot be used for estimation of overvaluation. Assuming a mild initial overvaluation of ten percent ($P_0 = a_0 P_0^f$ with $a_0 = 1.1$) as in Phillips et al. (2011), the bubble (in logs) starts at $b_0 = \log(a_0) = 0.095$. After 60 months (from the start of the bubble in October 1995 to its burst in October 2000), the overvaluation has accumulated to 312% which is notably larger than the estimate of 173% reported in Phillips et al. (2011). The calculations in Phillips et al. (2011) are based on a sample period from January 1990 to June 2000 rather than their identified explosive bubble period. We argue that the mixture of observations from a non-explosive unit root regime (prior to the bubble start in 1995) and those from the explosive bubble regime (after 1995) lowers the estimated persistence. The included observations from the non-explosive regime might be seen as a contamination when estimating the explosive root. This explains why their indirect inference estimate for the lower bound of the bubble growth rate (1.040) is lower than ours: 1.046. These slight discrepancies are, however, driving the economically significant difference in the estimated overvaluation. As Phillips et al. (2011) discuss by looking at the actual NASDAQ price dynamics, an overvaluation of around 278% would be reasonable. Our estimate of 312% is relatively close.

5 Conclusions

This paper compares different bias-correction techniques for autoregressive processes. Among these are the a bootstrap-based estimator, an indirect inference estimator and a jackknife estimator. In addition, we study the finite-sample performance of an estimator based on second differencing, a recursively demeaned OLS estimator and the Cauchy estimator. We focus on situations where the sample size is relatively small and data is highly persistent, exhibits a unit root or is even mildly explosive such as temporary bubbles in stock prices.

Our simulation study of bias and root mean squared errors of several estimators reveals the following results: The substantial bias of the OLS estimator can be remarkably reduced for any level of persistence. The most promising approaches are the indirect inference estimator, the bootstrap approach and the jackknife estimator. The indirect inference estimator provides excellent bias-correction in various settings together with a reasonably low variance, while the jackknife estimator performs often best in terms of bias-correction, but has a clearly larger variance, rendering this estimator less recommendable in terms of RMSE. The bootstrap approach often is the second best choice.

As an empirical illustration, we consider the widely studied NASDAQ composite price index from 1973:01 to 2005:06 to cover an explosive regime due to the "dot-com bubble". Our results highlight the need for bias-correction in light of precise estimation of the start and end date of the bubble and the implied stock market overvaluation. While the conventional OLS estimator suggests a stationary autoregressive root during the identified bubble period, the indirect inference estimator provides a plausible estimate for the bubble growth rate.

Bias-corrected estimation of persistence is also important in other related fields like vector autoregressions, see Engsted and Pedersen (2014) for a simulation study and Engsted and Nielsen (2012) for a co-explosive model. Clearly, an indirect inference estimator would be worthwhile to study in such an extended (explosive) framework. Moreover, it is well known that bias-correction plays a vital role in improving out-of-sample forecasts for stationary and highly persistent time series, see e.g. Gospodinov (2002), Kim (2003) and Kim and Durmaz (2012). There is, however, limited work on forecasting of explosive time series, see e.g. Grillenzoni (1998) for a notable exception. It would be thus an interesting issue to investigate to what extent forecasts of explosive stock prices during bubble periods could be improved by means of bootstrap, jackknife and indirect inference bias-correction techniques.

References

Abadir, K. M. (1993): "OLS bias in a nonstationary autoregression," *Econometric Theory*, 9, 81–93.

——— (1995): "Unbiased estimation as a solution to testing for random walks," *Economics Letters*, 47, 263 – 268.

Andrews, D. W. K. (1993): "Exactly median-unbiased estimation of first order autoregressive/unit root models," *Econometrica*, 61, 139–165.

Andrews, D. W. K. and H. Y. Chen (1994): "Approximately median-unbiased estimation of autoregressive models," *Journal of Business and Economic Statistics*, 12, 187–204.

Assaf, A. (2017): "Testing for bubbles in the art markets: An empirical investigation," *forthcoming in Economic Modelling*.

Bao, Y. (2007): "The approximate moments of the least squares estimator for the stationary autoregressive model under a general distribution," *Econometric Theory*, 23, 1013–1021.

Bao, Y. and A. Ullah (2007): "The second-order bias and mean squared error of estimators in time-series models," *Journal of Econometrics*, 140, 650–669.

Basawa, I. V., A. K. Mallik, W. P. McCormick, J. Reeves, and R. L. Taylor (1991): "Bootstrapping Unstable First-Order Autoregressive Processes," *The Annals of Statistics*, 19.

Basawa, I. V., A. K. Mallik, W. P. McCormick, and R. L. Taylor (1989): "Bootstrapping Explosive Autoregressive Processes," *The Annals of Statistics*, 17.

Berkowitz, J. and L. Kilian (2000): "Recent developments in bootstrapping time series," *Econometric Reviews*, 19.

Breitung, J. and R. Kruse (2013): "When bubbles burst: econometric tests based on structural breaks," *Statistical Papers*, 54, 911–930.

Caspi, I. (2016): "Testing for a housing bubble at the national and regional level: the case of Israel," *Empirical Economics*, 51, 483–516.

Chambers, M. J. (2013): "Jackknife estimation of stationary autoregressive models," *Journal of Econometrics*, 172, 142–157.

Chambers, M. J. and M. Kyriacou (2013): "Jackknife estimation with a unit root," *Statistics & Probability Letters*, 83, 1677–1682.

——— (2018): "Jackknife bias reduction in the presence of a near-unit root," *Econometrics*, 6, 11.

Chen, J. G. and B. S. Kuo (2013): "Gaussian inference in general AR(1) models based on difference," *Journal of Time Series Analysis*, 34, 447–453.

Chen, Y. and J. Yu (2015): "Optimal jackknife for unit root models," *Statistics & Probability Letters*, 99, 135–142.

Chong, J. and A. Hurn (2016): "Testing for Speculative Bubbles: Revisiting the Rolling Window," Tech. rep., School of Economics and Finance, Queensland University of Technology.

Datta, S. (1995): "Limit Theory and Bootstrap for Explosive and Partially Explosive Autoregression," *Stochastic Processes and their Applications*, 57.

——— (1996): "On asymptotic properties of bootstrap for AR(1) processes," *Journal of Statistical Planning and Inference*, 53.

Efron, B. (1979): "Bootstrap methods: Another look at the jacknife," *Annals of Statistics*, 7, 1–26.

Engsted, T., S. J. Hviid, and T. Q. Pedersen (2016): "Explosive bubbles in house prices? Evidence from the OECD countries," *Journal of International Financial Markets, Institutions and Money*, 40, 14–25.

Engsted, T. and B. Nielsen (2012): "Testing for rational bubbles in a coexplosive vector autoregression," *The Econometrics Journal*, 15, 226–254.

Engsted, T. and T. Q. Pedersen (2014): "Bias-correction in vector autoregressive models: A simulation study," *Econometrics*, 2, 45–71.

Etienne, X. L., S. H. Irwin, and P. Garcia (2014): "Bubbles in food commodity markets: Four decades of evidence," *Journal of International Money and Finance*, 42, 129–155.

Figuerola-Ferretti, I., C. L. Gilbert, and J. R. McCrorie (2015): "Testing for Mild Explosivity and Bubbles in LME Non-Ferrous Metals Prices," *Journal of Time Series Analysis*, forthcoming.

Gospodinov, N. (2002): "Median unbiased forecasts for highly persistent autoregressive processes," *Journal of Econometrics*, 111, 85–101.

Gouriéroux, C., A. Monfort, and E. Renault (1993): "Indirect Inference," *Journal of Applied Econometrics*, 8, 85–118.

Gouriéroux, C., E. Renault, and N. Touzi (2000): "Calibration by simulation for small sample bias correction," in *Simulation-Based Inference in Econometrics: Methods and Applications*, ed. by R. Mariano, T. Schuermann, and M. Weeks, Cambridge University Press, 328–358.

Grillenzoni, C. (1998): "Forecasting unstable and nonstationary time series," *International Journal of Forecasting*, 14, 469–482.

Gutierrez, L. (2011): "Bootstrapping asset price bubbles," *Economic Modelling*, 28, 2488–2493.

——— (2013): "Speculative bubbles in agricultural commodity markets," *European Review of Agricultural Economics*, 40, 217–238.

Harvey, D. I. and S. J. Leybourne (2014): "Asymptotic behaviour of tests for a unit root against an explosive alternative," *Economics Letters*, 122, 64–68.

Harvey, D. I., S. J. Leybourne, and R. Sollis (2017): "Improving the accuracy of asset price bubble start and end date estimators," *Journal of Empirical Finance*, 40, 121–138.

Homm, U. and J. Breitung (2012): "Testing for speculative bubbles in stock markets: a comparison of alternative methods," *Journal of Financial Econometrics*, 10, 198–231.

Hu, Y. and L. Oxley (2017): "Are there bubbles in exchange rates? Some new evidence from G10 and emerging market economies," *Economic Modelling*.

Inoue, A. and L. Kilian (2002): "Bootstrapping autoregressive processes with possible unit roots," *Econometrica*, 70.

Kaufmann, H. and R. Kruse (2013): "Bias corrected estimation in potentially mildly explosive autoregressive models," *CREATES Research Paper 2013-10.*

Kendall, M. G. (1954): "Notes on bias in the estimation of autocorrelation," *Biometrika*, 41, 403–404.

Kilian, L. (1998): "Small sample confidence intervals for impulse response functions," *The Review of Economics and Statistics*, 80.

Kim, H. and N. Durmaz (2012): "Bias correction and out-of-sample forecast accuracy," *International Journal of Forecasting*, 28, 575–586.

Kim, J. H. (2003): "Forecasting autoregressive time series with bias-corrected parameter estimators," *International Journal of Forecasting*, 19, 493–502.

Kiviet, J. F. and G. D. A. Phillips (2012): "Higher-order asymptotic expansions of the least-squares estimation bias in first-order dynamic regression models," *Computational Statistics and Data Analysis*, 56, 3705–3729.

Kräussl, R., T. Lehnert, and N. Martelin (2016): "Is there a bubble in the art market?" *Journal of Empirical Finance*, 35, 99–109.

MacKinnon, J. G. and A. A. Smith (1998): "Approximate bias correction in econometrics," *Journal of Econometrics*, 85, 205–230.

Nelson, C. R. and C. R. Plosser (1982): "Trends and random walks in macroeconmic time series: some evidence and implications," *Journal of Monetary Economics*, 10, 139–162.

Phillips, P. C. and J. Yu (2009): "Limit theory for dating the origination and collapse of mildly explosive periods in time series data," Tech. rep., Sim Kee Book Institute (SKBI) for Financial Economics Discussion Paper, Singapore Management University.

Phillips, P. C. B. (2012): "Folklore theorems, implicit maps, and indirect inference," *Econometrica*, 80, 425–454.

Phillips, P. C. B. and T. Magdalinos (2007): "Limit theory for moderate deviations from a unit root," *Journal of Econometrics*, 136, 115–130.

Phillips, P. C. B., S. Shi, and J. Yu (2014): "Specification Sensitivity in Right-Tailed Unit Root Testing for Explosive Behaviour," *Oxford Bulletin of Economics and Statistics*, 76, 315–333.

Phillips, P. C. B., Y. Wu, and J. Yu (2011): "Explosive behavior in the 1990s NASDAQ: When did exuberance escalate asset values?" *International Economic Review*, 52, 201–226.

Potrafke, N. and M. Reischmann (2014): "Explosive Target balances of the German Bundesbank," *Economic Modelling*, 42, 439–444.

Roy, A. and W. A. Fuller (2001): "Estimation for autoregressive time series with a root near 1," *Journal of Business and Economic Statistics*, 19, 482–493.

Schotman, P. and H. K. van Dijk (1991): "A Bayesian analysis of the unit root in real exchange rates," *Journal of Econometrics*, 49, 195–238.

Shaman, P. and R. A. Stine (1988): "The bias of autoregressive coefficient estimators," *Journal of the American Statistical Association*, 83, 842–848.

Shi, S. (2017): "Speculative bubbles or market fundamentals? An investigation of US regional housing markets," *forthcoming in Economic Modelling.*

Smith, A. A. (1993): "Estimating nonlinear time-series models using simulated vector autoregressions," *Journal of Applied Econometrics*, 8, 63–84.

So, B. S. and D. W. Shin (1999a): "Cauchy Estimators for Autoregressive Processes with Applications to Unit Root Tests and Confidence Intervals," *Econometric Theory*, 15, 165 – 176.

——— (1999b): "Recursive mean adjustment in time-series inferences," *Statistics & Probability Letters*, 43, 65 – 73.

Steenkamp, D. (2017): "Explosiveness in G11 currencies," *forthcoming in Economic Modelling.*

Tanaka, K. (1984): "An asymptotic expansion associated with the maximum likelihood estimators in ARMA models," *Journal of the Royal Statistical Society, Series B (Methodological)*, 46, 58–67.

Tjøstheim, D. and J. Paulsen (1983): "Bias of some commonly-used time series estimates," *Biometrika*, 70, 389–399.

Wang, X. and J. Yu (2015): "Limit theory for an explosive autoregressive process," *Economics Letters*, 126, 176–180.

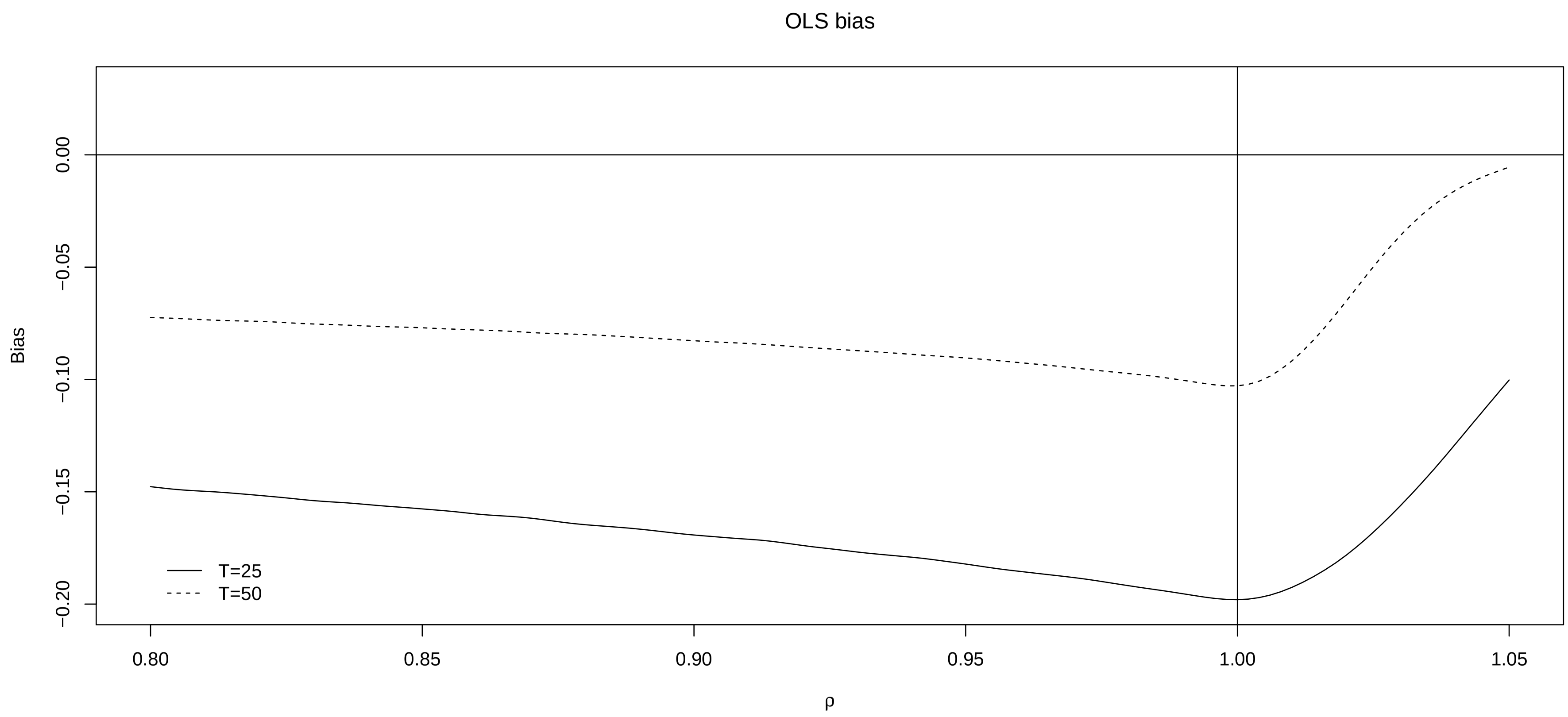

Figure 1: OLS estimation bias for ρ in $y_t = \mu + \rho y_{t-1} + \varepsilon_t$.

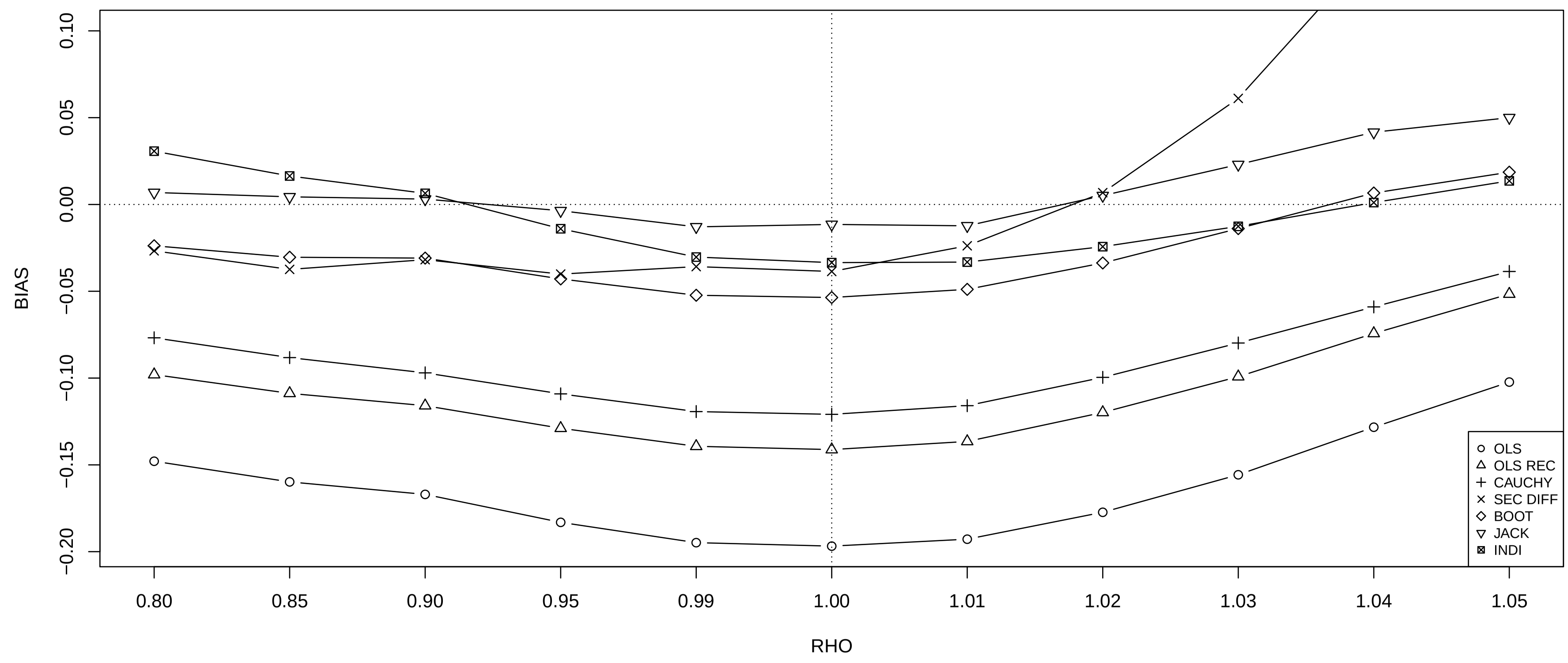

Figure 2: Bias for $T = 25$ for estimators: $\widehat{\rho}$ (OLS), $\widehat{\rho}^R$ (OLS REC), $\widehat{\rho}^C$ (CAUCHY), $\widehat{\rho}^{SD}$ (SEC DIFF), $\widehat{\rho}^B$ (BOOT), $\widehat{\rho}^J$ (JACK) and $\widehat{\rho}^{II}$ (INDI).

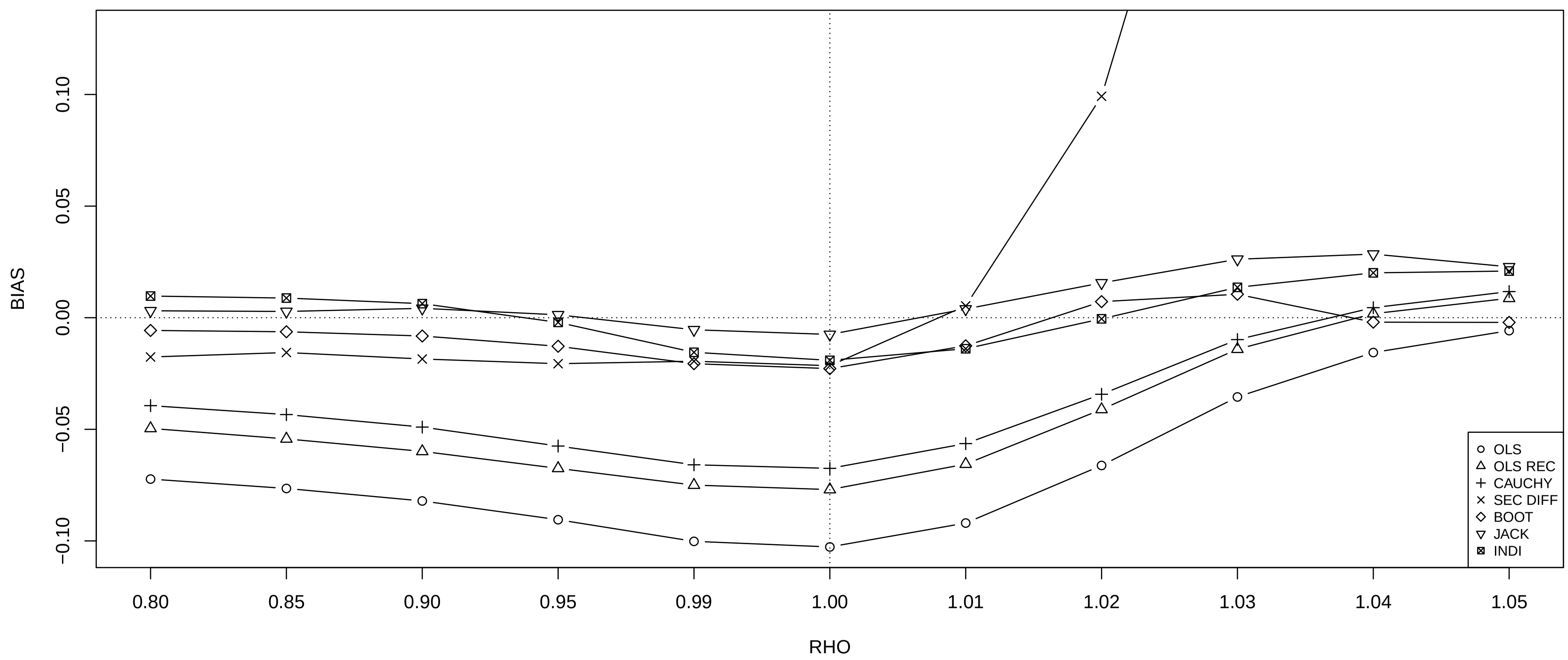

Figure 3: Bias for $T = 50$ for estimators: $\widehat{\rho}$ (OLS), $\widehat{\rho}^R$ (OLS REC), $\widehat{\rho}^C$ (CAUCHY), $\widehat{\rho}^{SD}$ (SEC DIFF), $\widehat{\rho}^B$ (BOOT), $\widehat{\rho}^J$ (JACK) and $\widehat{\rho}^{II}$ (INDI).

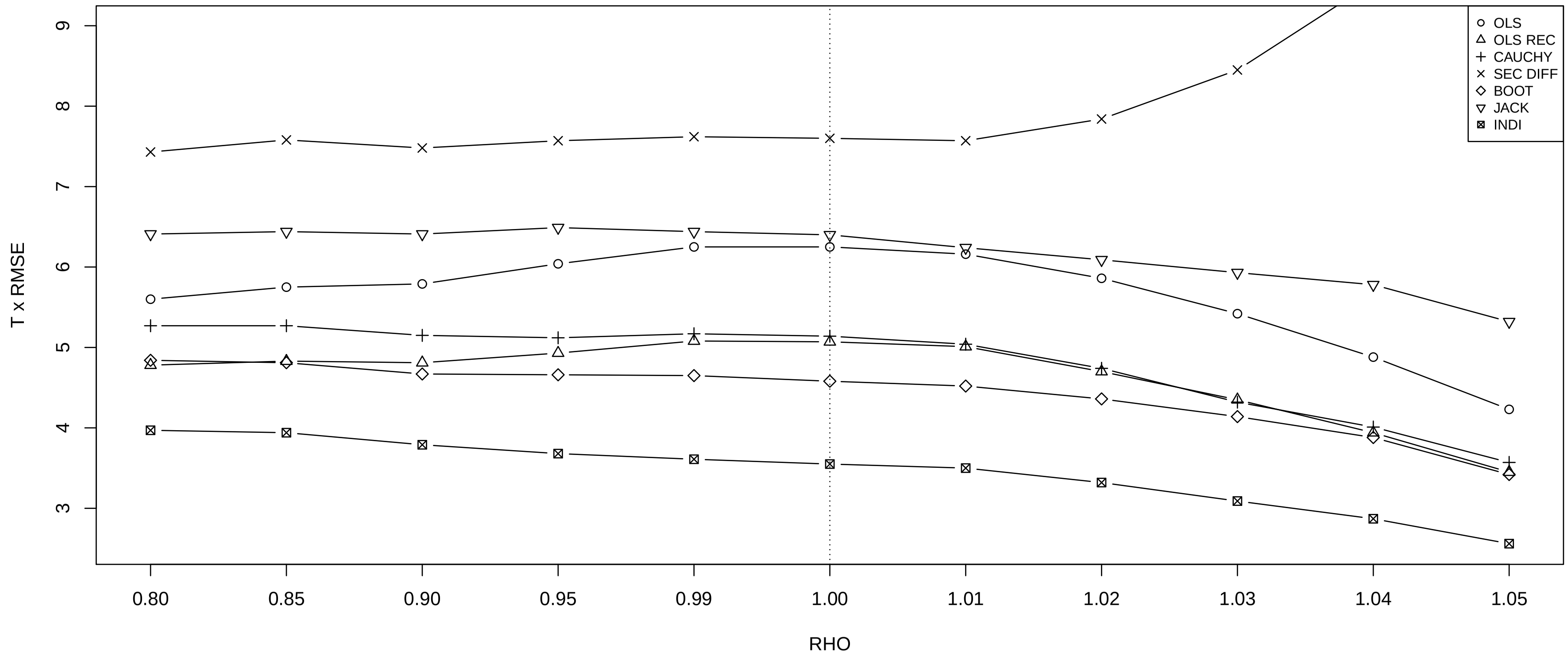

Figure 4: $T \times$RMSE for $T = 25$ for estimators: $\widehat{\rho}$ (OLS), $\widehat{\rho}^R$ (OLS REC), $\widehat{\rho}^C$ (CAUCHY), $\widehat{\rho}^{SD}$ (SEC DIFF), $\widehat{\rho}^B$ (BOOT), $\widehat{\rho}^J$ (JACK) and $\widehat{\rho}^{II}$ (INDI).

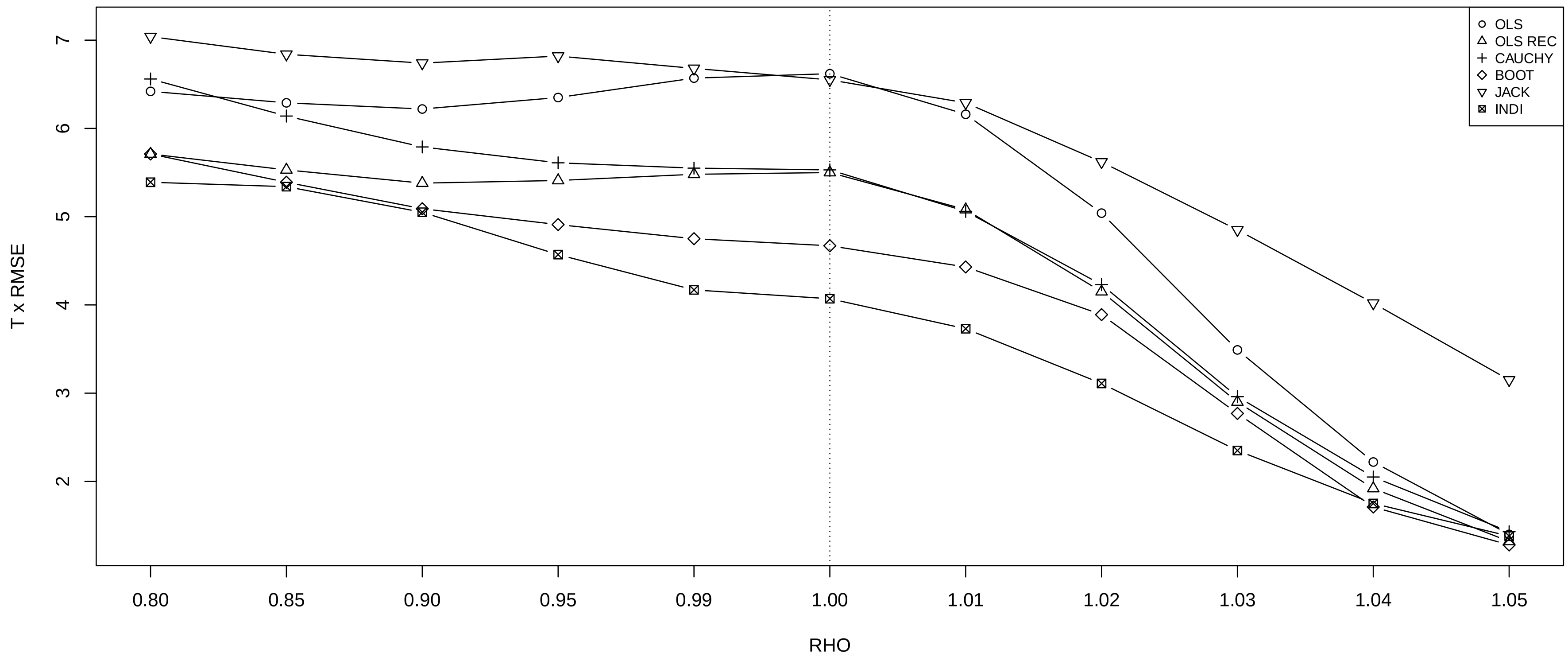

Figure 5: $T \times$RMSE for $T = 50$ for estimators: $\widehat{\rho}$ (OLS), $\widehat{\rho}^R$ (OLS REC), $\widehat{\rho}^C$ (CAUCHY), $\widehat{\rho}^B$ (BOOT), $\widehat{\rho}^J$ (JACK) and $\widehat{\rho}^{II}$ (INDI). The $\widehat{\rho}^{SD}$ estimator (SEC DIFF) is excluded.

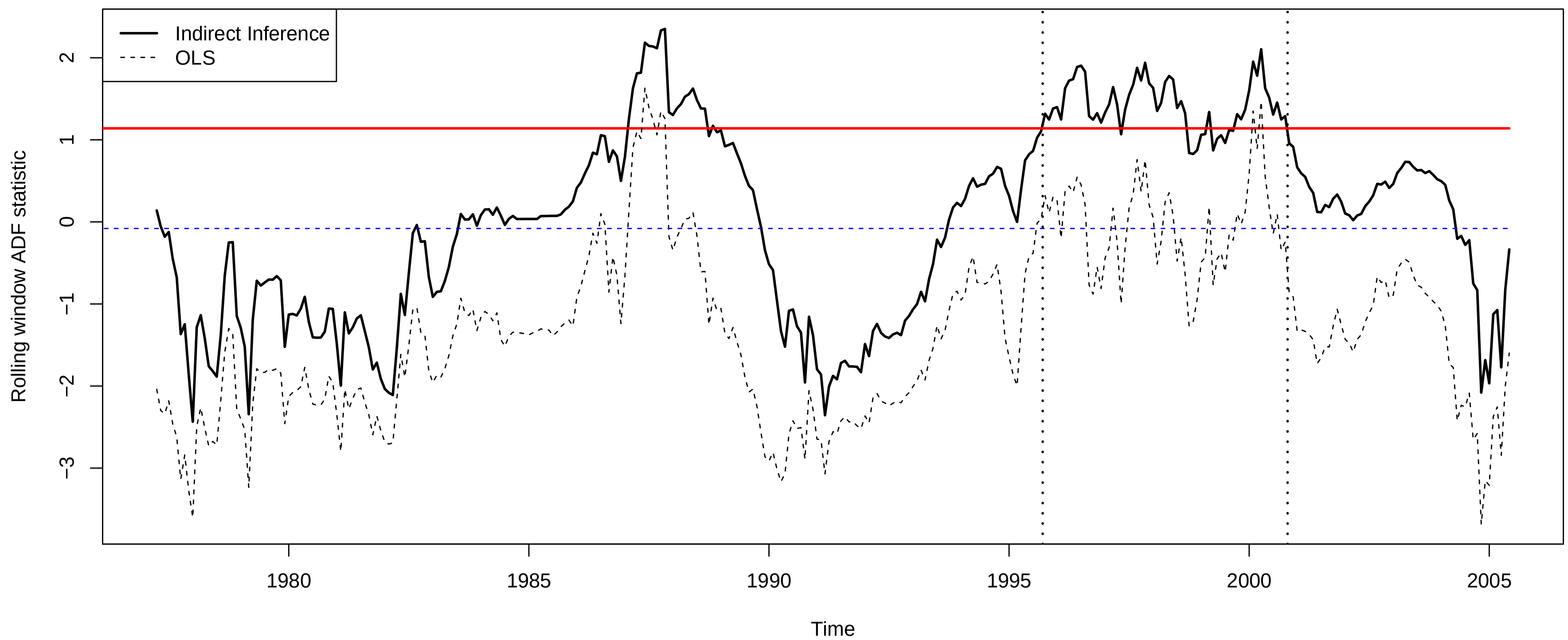

Figure 6: Rolling window ADF statistics for the real NASDAQ index. Indirect inference statistics t_w^{II} (thick black solid line) are compared to OLS t_w (black dashed line). Critical values at the nominal five percent level are displayed as thick red solid line (1.14) and blue dashed line (-0.08) for indirect inference and OLS, respectively. Vertical black dotted lines represent $\widehat{r}_e^{II}$ and $\widehat{r}_p^{II}$, respectively.

Table 1: Bias, dominant drift component ($\eta = 0.4$)

$T = 25$

ρ	$\widehat{\rho}$	$\widehat{\rho}^R$	$\widehat{\rho}^C$	$\widehat{\rho}^{SD}$	$\widehat{\rho}^B$	$\widehat{\rho}^J$	$\widehat{\rho}^{II}$
1.00	-0.132	-0.071	-0.050	0.061	0.013	0.052	0.015
1.01	-0.120	-0.062	-0.039	0.117	0.023	0.052	0.021
1.02	-0.108	-0.048	-0.026	0.164	0.030	0.062	0.026
1.03	-0.117	-0.058	-0.040	0.177	0.018	0.044	0.013
1.04	-0.098	-0.043	-0.028	0.238	0.034	0.060	0.022
1.05	-0.089	-0.041	-0.029	0.290	0.025	0.055	0.023

$T = 50$

ρ	$\widehat{\rho}$	$\widehat{\rho}^R$	$\widehat{\rho}^C$	$\widehat{\rho}^{SD}$	$\widehat{\rho}^B$	$\widehat{\rho}^J$	$\widehat{\rho}^{II}$
1	-0.064	-0.033	-0.023	0.039	0.014	0.033	0.011
1.01	-0.054	-0.026	-0.018	0.099	0.022	0.033	0.013
1.02	-0.046	-0.022	-0.016	0.164	0.022	0.029	0.012
1.03	-0.033	-0.012	-0.009	0.301	0.013	0.029	0.016
1.04	-0.029	-0.010	-0.008	0.432	-0.005	0.029	0.012
1.05	-0.024	-0.008	-0.007	0.554	-0.009	0.028	0.010

Table 2: Bias, dominated drift component ($\eta = 0.6$)

$T = 25$

ρ	$\widehat{\rho}$	$\widehat{\rho}^R$	$\widehat{\rho}^C$	$\widehat{\rho}^{SD}$	$\widehat{\rho}^B$	$\widehat{\rho}^J$	$\widehat{\rho}^{II}$
1.00	-0.182	-0.122	-0.098	-0.015	-0.038	0.001	-0.023
1.01	-0.166	-0.108	-0.091	0.016	-0.022	0.022	-0.012
1.02	-0.166	-0.107	-0.085	0.036	-0.024	0.019	-0.018
1.03	-0.154	-0.098	-0.079	0.061	-0.014	0.018	-0.011
1.04	-0.148	-0.095	-0.079	0.100	-0.008	0.026	-0.012
1.05	-0.143	-0.089	-0.070	0.130	-0.009	0.024	-0.012

$T = 50$

ρ	$\widehat{\rho}$	$\widehat{\rho}^R$	$\widehat{\rho}^C$	$\widehat{\rho}^{SD}$	$\widehat{\rho}^B$	$\widehat{\rho}^J$	$\widehat{\rho}^{II}$
1.00	-0.090	-0.063	-0.054	0.001	-0.010	0.005	-0.008
1.01	-0.085	-0.059	-0.049	0.017	-0.006	0.008	-0.008
1.02	-0.078	-0.053	-0.045	0.062	-0.003	0.004	-0.009
1.03	-0.064	-0.041	-0.035	0.141	0.000	0.017	-0.003
1.04	-0.052	-0.031	-0.027	0.257	-0.006	0.017	-0.001
1.05	-0.045	-0.026	-0.025	0.384	-0.016	0.019	-0.002

Table 3: Bias, AR(2) process

$T = 25$

ρ	$\widehat{\rho}$	$\widehat{\rho}^{R}$	$\widehat{\rho}^{C}$	$\widehat{\rho}^{SD}$	$\widehat{\rho}^{B}$	$\widehat{\rho}^{J}$	$\widehat{\rho}^{II}$
1.00	-0.336	-0.278	-0.251	-0.354	-0.211	-0.100	-0.145
1.01	-0.300	-0.243	-0.224	-0.306	-0.170	-0.052	-0.119
1.02	-0.285	-0.227	-0.201	-0.290	-0.152	-0.035	-0.110
1.03	-0.278	-0.219	-0.196	-0.255	-0.145	-0.031	-0.106
1.04	-0.217	-0.159	-0.140	-0.158	-0.078	0.017	-0.063
1.05	-0.187	-0.131	-0.118	-0.078	-0.049	0.037	-0.044

$T = 50$

ρ	$\widehat{\rho}$	$\widehat{\rho}^{R}$	$\widehat{\rho}^{C}$	$\widehat{\rho}^{SD}$	$\widehat{\rho}^{B}$	$\widehat{\rho}^{J}$	$\widehat{\rho}^{II}$
1.00	-0.177	-0.150	-0.138	-0.316	-0.103	-0.030	-0.088
1.01	-0.162	-0.133	-0.121	-0.283	-0.087	-0.021	-0.078
1.02	-0.121	-0.094	-0.088	-0.211	-0.045	0.012	-0.045
1.03	-0.075	-0.052	-0.047	-0.056	-0.004	0.034	-0.014
1.04	-0.043	-0.023	-0.022	0.201	-0.001	0.039	0.003
1.05	-0.024	-0.007	-0.005	0.504	-0.013	0.028	0.010

Table 4: $T\times$RMSE, dominant drift component ($\eta = 0.4$)

$T = 25$

ρ	$\widehat{\rho}$	$\widehat{\rho}^{R}$	$\widehat{\rho}^{C}$	$\widehat{\rho}^{SD}$	$\widehat{\rho}^{B}$	$\widehat{\rho}^{J}$	$\widehat{\rho}^{II}$
1.00	4.931	3.870	3.973	7.900	4.105	5.936	3.026
1.01	4.555	3.572	3.579	8.630	3.787	5.747	2.767
1.02	4.192	3.195	3.279	8.993	3.525	5.485	2.566
1.03	4.565	3.586	3.561	9.910	3.771	5.567	2.769
1.04	4.084	3.212	3.280	10.657	3.530	5.273	2.545
1.05	3.890	3.198	3.432	11.644	3.220	5.172	2.390

$T = 50$

ρ	$\widehat{\rho}$	$\widehat{\rho}^{R}$	$\widehat{\rho}^{C}$	$\widehat{\rho}^{SD}$	$\widehat{\rho}^{B}$	$\widehat{\rho}^{J}$	$\widehat{\rho}^{II}$
1.00	2.376	1.814	1.872	5.644	1.890	3.047	1.522
1.01	2.114	1.605	1.719	6.150	1.774	2.757	1.320
1.02	2.075	1.728	1.730	7.594	1.798	2.722	1.400
1.03	1.581	1.283	1.404	10.483	1.230	2.314	1.017
1.04	1.620	1.315	1.344	13.782	1.135	2.405	1.089
1.05	1.527	1.322	1.378	16.545	1.054	2.205	1.033

Table 5: $T \times$RMSE, dominated drift component ($\eta = 0.6$)

$T = 25$

ρ	$\widehat{\rho}$	$\widehat{\rho}^R$	$\widehat{\rho}^C$	$\widehat{\rho}^{SD}$	$\widehat{\rho}^B$	$\widehat{\rho}^J$	$\widehat{\rho}^{II}$
1.00	5.976	4.724	4.818	7.579	4.500	6.203	3.541
1.01	5.560	4.442	4.537	7.633	4.166	6.174	3.151
1.02	5.696	4.541	4.414	8.090	4.351	6.300	3.294
1.03	5.391	4.331	4.392	8.473	4.095	5.957	3.086
1.04	5.328	4.381	4.507	9.002	4.133	5.743	3.052
1.05	5.197	4.247	4.092	9.482	3.979	5.702	2.914

$T = 50$

ρ	$\widehat{\rho}$	$\widehat{\rho}^R$	$\widehat{\rho}^C$	$\widehat{\rho}^{SD}$	$\widehat{\rho}^B$	$\widehat{\rho}^J$	$\widehat{\rho}^{II}$
1.00	2.948	2.382	2.343	4.789	2.075	3.086	1.748
1.01	2.907	2.392	2.295	5.225	2.104	3.114	1.760
1.02	2.769	2.282	2.281	6.102	2.013	2.889	1.644
1.03	2.457	2.051	2.034	7.618	1.777	2.764	1.478
1.04	2.179	1.810	1.796	10.182	1.529	2.523	1.327
1.05	2.066	1.773	1.832	13.259	1.412	2.465	1.290

Table 6: $T \times$RMSE, AR(2) process

$T = 25$

ρ	$\widehat{\rho}$	$\widehat{\rho}^R$	$\widehat{\rho}^C$	$\widehat{\rho}^{SD}$	$\widehat{\rho}^B$	$\widehat{\rho}^J$	$\widehat{\rho}^{II}$
1	10.088	8.880	8.738	12.355	8.270	8.733	5.757
1.01	9.151	7.984	8.134	11.687	7.324	8.258	5.239
1.02	8.932	7.752	7.647	11.795	7.198	8.017	5.136
1.03	8.937	7.744	7.760	11.142	7.303	8.298	5.194
1.04	7.430	6.409	6.368	10.340	6.014	7.295	4.299
1.05	6.743	5.806	6.153	10.362	5.516	6.770	3.900

$T = 50$

ρ	$\widehat{\rho}$	$\widehat{\rho}^R$	$\widehat{\rho}^C$	$\widehat{\rho}^{SD}$	$\widehat{\rho}^B$	$\widehat{\rho}^J$	$\widehat{\rho}^{II}$
1	5.462	4.893	4.838	9.822	4.325	4.835	3.793
1.01	5.211	4.614	4.654	9.360	4.165	4.761	3.648
1.02	4.195	3.682	3.837	8.526	3.301	4.048	2.878
1.03	3.060	2.699	2.664	7.805	2.510	3.449	2.112
1.04	2.046	1.780	1.978	9.953	1.640	2.884	1.410
1.05	1.211	0.997	1.050	14.741	0.970	1.990	0.881