Nearly unbiased estimation of autoregressive models for bounded near-integrated stochastic processes*

JOSEP LLUÍS CARRION-I-SILVESTRE[†] and MARÍA DOLORES GADEA[‡] and ANTONIO MONTAÑÉS[§]

[†]*AQR-IREA, Department of Econometrics, Statistics, and Applied Economics, University of Barcelona, Av. Diagonal 690, 08034 Barcelona (e-mail: carrion@ub.edu)*

[‡]*Department of Applied Economics, University of Zaragoza, Gran Vía 4, 50005 Zaragoza, Spain (e-mail: lgadea@unizar.es)*

[§]*Department of Economic Analysis, University of Zaragoza, Gran Vía 4, 50005 Zaragoza, Spain (e-mail: amontane@unizar.es)*

Abstract

The paper investigates the estimation bias of autoregressive models for bounded near-integrated stochastic processes and the performance of the standard procedures in the literature that aim to correct the estimation bias. In some cases, the bounded nature of the stochastic processes worsens the estimation bias effect. The paper extends two popular autoregressive estimation bias correction procedures to cover bounded stochastic processes. Monte Carlo simulations reveal that accounting for the bounded nature of the stochastic processes leads to improvements in the estimation of autoregressive models. Finally, an illustration is given using the unemployment rate of the G7 countries.

JEL classification: C22, C32, E32, Q43

Keywords: Bounded near-integrated stochastic processes, estimation bias, unit root tests, unemployment persistence

Word count: 6131

*We thank the editor and two referees for helpful comments and suggestions. The authors gratefully acknowledge the financial support from the Spanish Ministerio de Ciencia y Tecnología, Agencia Española de Investigación (AEI) and European Regional Development Fund (ERDF, EU) under grants ECO2015-65967-R (A. Montañés), ECO2017-83255-C3-1-P (AEI/ERDF, EU) (J. L. Carrion-i-Silvestre and M. D. Gadea) and ECO2016-81901-REDT.

I. Introduction

Since the seminal paper by Nelson and Plosser (1982), time series data-based analysis has frequently begun with the study of the time properties of the variables. This usually implies the use of some unit root tests, and the statistical inference drawn from their application is important for subsequent analyses. For instance, a quite popular practice is to determine the persistence degree of shocks by means of estimating autoregressive models. This provides very interesting insights about the evolution of the variable being studied, including the analysis of persistence in variables such as real exchange rates, where some practitioners have studied the number of periods that a shock takes to vanish – see Balli *et al.* (2014), among others. Similarly, Watson (2014) studies the effect of the Great Recession on inflation persistence. This type of analysis, however, is not straightforward given that we should take into account that the ordinary least-squares (OLS) estimator is consistent but biased in finite samples, and this bias must be removed in order to appropriately measure the degree of persistence. There are various proposals in the literature which try to correct this finite sample bias. We can cite here the contributions of Andrews (1993), Andrews and Chen (1994), Kilian (1998), Hansen (1999), Rossi (2005) and Perron and Yabu (2009a), among others, which develop different valid techniques to remove the estimation bias.

However, some commonly employed variables may be affected by the presence of bounds. Macroeconomic variables such as nominal interest rates, unemployment rates, exchange rates and the great ratios, among others, are bounded by definition, preventing them from exhibiting a large variance. This feature generates tension in the statistical inference associated with standard unit root tests and, hence, the estimation of the degree of persistence of shocks. The standard order of integration analysis of time series considers that an I(1) non-stationary stochastic process can vary freely within the limit, that is, the constraints that impose the existence of bounds are ignored. The behavior of these types of variables might seem to be stationary when, in fact, they are non-stationary. In this regard, Cavaliere (2005) and Cavaliere and Xu (2014) show that standard unit root tests might reach misleading conclusions if the bounded nature of the time series is not accounted for. Therefore, it is recommendable to analyze the influence of these bounds on the determination of time series properties.

The goal of this paper is to assess whether the use of bias-corrected autoregressive parameter estimates allows us to obtain more accurate empirical economic analyses that build upon the computation of statistics such as shock persistence measures or long-run variance (LRV) estimates. To address this issue, the paper investigates the performance of some of the popular estimation bias correction methods mentioned above when applied to bounded near-integrated stochastic processes. First, we focus on some standard procedures, showing that, in general, the amount of estimation bias that is corrected is small when the bounded nature of the time series is ignored. This suggests the need to extend these standard procedures to incorporate the effect of the bounds on the estimation of autoregressive models for persistent time series.

The paper proceeds as follows. Section II. describes the model for bounded (near-integrated and integrated) stochastic processes. Section III. motivates the analysis showing that standard bias correction methods give poor estimates when applied to bounded stochastic processes. This leads us to propose in Section IV. an extension of bias correction methods that considers this feature. Section V. analyzes the finite sample performance of the suggested approaches. Section VI. provides an empirical illustration, focusing on the unemployment rate persistence of the G7 countries. Finally, Section VII. sets out the conclusions. The proofs and supplementary material are collected in the appendix.

II. The model

Let x_t be a stochastic process with a data generating process (DGP) given by:

$$x_t = \mu + y_t \tag{1}$$

$$y_t = \alpha y_{t-1} + u_t \tag{2}$$

$t = 1, \ldots, T$, where $x_t \in \left[\underline{b}, \overline{b}\right]$ almost surely for all t, $y_0 = O_p(1)$, and $\left[\underline{b}, \overline{b}\right]$ denote the boundaries that affect the time series. The autoregressive parameter is set as $\alpha = \exp(-\kappa/T) \approx 1 - \kappa/T$, with $\kappa \geq 0$ being the non-centrality parameter, so that the model specification covers both the case in which the time series is a near-integrated process – i.e., a NI(1) process with $\kappa > 0$ – and an I(1) non-stationary process – when $\kappa = 0$. The disturbance term u_t is assumed to admit the decomposition:

$$u_t = \varepsilon_t + \underline{\xi_t} - \overline{\xi_t} \tag{3}$$

and the variables $\underline{\xi_t}$ and $\overline{\xi_t}$ are non-negative processes (regulators) such that $\underline{\xi_t} > 0$ if and only if $\alpha y_{t-1} + \varepsilon_t < \underline{b} - \mu$ and $\overline{\xi_t} > 0$ if and only if $\alpha y_{t-1} + \varepsilon_t > \overline{b} - \mu$. The following assumptions are assumed to be satisfied by the stochastic processes in (3).

Assumption 1: $\varepsilon_t = C(L) v_t$, where $C(L) = \sum_{j=0}^{\infty} c_j L^j$ with $\sum_{j=0}^{\infty} j^s |c_j| < \infty$, for some $s \geq 1$, and v_t is a martingale difference sequence adapted to the filtration $F_t = \sigma - field\{v_{t-j}; j \geq 0\}$. The LRV of ε_t is given by (a) $\sigma^2 = \lim_{T\to\infty} E[T^{-1}(\sum_{t=1}^{T} \varepsilon_t)^2] = \sigma_v^2 C(1)^2$, (b) $\sigma_v^2 = \lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} E(v_t^2) < \infty$ $\forall t$, and (c) $E|v_t^r| < \infty$ for some $r > 4$.

Assumption 2: $\{\underline{\xi_t}\}_{t=1}^{T}$and $\left\{\overline{\xi_t}\right\}_{t=1}^{T}$ satisfy restrictions to ensure that $\max_{t=1,\ldots,T} |\underline{\xi_t}| = o_p\left(T^{1/2}\right)$ and $\max_{t=1,\ldots,T} |\overline{\xi_t}| = o_p\left(T^{1/2}\right)$.

Assumption 3: $(\underline{b}-\mu)=\underline{c}\sigma T^{1/2}+o(1)$ and $(\overline{b}-\mu)=\overline{c}\sigma T^{1/2}+o(1)$, $\underline{c}\leq 0\leq\overline{c}$, $\underline{c}\neq\overline{c}$.

Based on these assumptions, we can define the (standardized) bounds that affect y_t as:

$$\left[\frac{(\underline{b}-\mu)}{\sigma T^{1/2}},\frac{(\overline{b}-\mu)}{\sigma T^{1/2}}\right]=[\underline{c},\overline{c}]+o(1) \tag{4}$$

with $\underline{c}\leq 0\leq\overline{c}$, $\underline{c}\neq\overline{c}$. Note that the model specification can be particularized to stochastic processes that are only limited below – i.e., $x_t\in[\underline{b},\infty]$ – or only limited above – i.e., $x_t\in\left[-\infty,\overline{b}\right]$ – but also covers the case of unbounded processes – i.e., $x_t\in[-\infty,\infty]$. For this near-integrated set-up the following theorem shows that the OLS estimator of α in the system defined by (1) and (2) is consistent.

Theorem 1 *Let $\{x_t\}_{t=1}^{T}$ be the bounded stochastic process given by (1) to (3). Under Assumptions 1 to 3 and as $T\to\infty$, the OLS estimator is:*

$$(\hat{\alpha}-\alpha)\xrightarrow{p}0$$

where $\xrightarrow{p}$ denotes convergence in probability.

The proof is given in the appendix. Although the OLS estimator is a consistent estimator of α when dealing with bounded stochastic processes, it is to be expected that some estimation bias would appear in finite samples. Thus, the goal of this paper is to study the extent of this estimation bias distortion and how some popular estimation bias correction procedures perform in practice.

Finally, it is worth emphasizing that the paper deals with NI(1) and I(1) stochastic processes, and that the bounds are defined according to this framework, i.e., $\left[\underline{b},\overline{b}\right]=O\left(T^{1/2}\right)$. In principle, it would be possible to design a framework for bounded I(0) stochastic processes with fixed bounds given by $\left[(\underline{b}-\mu)/\sigma,\left(\overline{b}-\mu\right)/\sigma\right]=[\underline{c},\overline{c}]+o(1)$, so that $\left[\underline{b},\overline{b}\right]=O(1)$. However, the OLS estimator of α would lead to inconsistent estimates.

III. Estimation bias correction methods

The estimation of autoregressive models is at the heart of popular practices in empirical economics such as order of integration analysis and the computation of shock persistence measures. However, it is well known that OLS estimation provides biased estimates in finite samples, although the bias disappears asymptotically. The literature has provided different estimation bias correction methods such as those found in Andrews (1993), Andrews and Chen (1994), Kilian (1998) and Hansen (1999) – which rely on simulation techniques – and Roy and Fuller (2001) and Perron and Yabu (2009a) – which apply a correction function to the OLS estimate. In what follows, we focus on two of these approaches: (i) the median-unbiased (MU) estimation procedure suggested in Andrews (1993) and Andrews and Chen (1994) and (ii) the truncated and super-efficient estimator advocated in Perron and Yabu (2009a). The selection of these approaches is driven by two reasons. First, preliminary simulations, not reported here to save space, reveal that the MU procedure outperforms the proposals in Kilian (1998) and Hansen (1999). Second, we take into account that the Perron and Yabu (2009a) method builds upon the Roy and Fuller (2001) approximation, so it can be seen as an enhanced estimation procedure. The discussion below focuses on the standard implementation of these proposals and their performance when they are applied to bounded stochastic processes.

The median-unbiased estimation method

The estimation bias correction approach in Andrews (1993) deals with AR(1) stochastic processes:

$$x_t = f(t) + \alpha x_{t-1} + e_t \tag{5}$$

where $f(t)$ denotes the deterministic component – i.e., $f(t) = 0$, $f(t) = \mu$ or $f(t) = \mu + \beta t$. The MU estimation technique is based on establishing a correspondence between the OLS estimation of the autoregressive parameter α in (5) – denoted as $\hat{\alpha}$ – and the median of the empirical distribution that is obtained by means of the Imhof routine under the assumption that $\alpha = \hat{\alpha}$. This defines the so-called median-unbiased autoregressive estimator of α – henceforth, $\hat{\alpha}_{MU}$. Andrews (1993) provides look-up tables for the AR(1) case that can be used to approximate $\hat{\alpha}_{MU}$. Andrews and Chen (1994) extend the procedure to AR(p) stochastic processes of the form:

$$x_t = f(t) + \alpha x_{t-1} + \sum_{j=1}^{p-1} \psi_j \Delta x_{t-j} + e_t \tag{6}$$

although in this case the look-up tables are time-series-dynamic-specific, and the authors suggest the use of a simple iterative procedure that yields an approximate $\hat{\alpha}_{MU}$ – see Pesavento and Rossi (2006) for further details. As can be seen, the MU-based approach is a computationally-intensive procedure, especially for high-order autoregressive processes.

Weighted symmetric least-squares estimation

The truncated and super-efficient estimator of Perron and Yabu (2009a) – henceforth, PY – accounts for the fact that the bias of the autoregressive parameter estimates in the near-integrated and I(1) non-stationary areas worsens when compared to the bias obtained for moderate persistent processes. Roy and Fuller (2001) and Perron and Yabu (2009a) suggest the use of the modified estimator – hereafter, the truncated weighted

(TW) estimator – given by:

$$\hat{\alpha}_{TW} = \hat{\alpha}_W + C\left(\hat{\tau}_W\right)\hat{\sigma}_W \tag{7}$$

where $\hat{\alpha}_W$ denotes the weighted symmetric least-squares (WSLS) estimate of the autoregressive parameter α in the model:

$$\begin{aligned} x_t &= (1-\alpha)\,\mu + \alpha x_{t-1} + \sum_{j=1}^{k} \psi_j \Delta x_{t-j} + e_{k,t} \\ \hat{x}_t &= \alpha \hat{x}_{t-1} + \hat{e}_{k,t} \end{aligned} \tag{8}$$

with $\hat{x} = Mx$ and $\hat{e}_k = Me_k$, with $M = I_{T-k} - Z\left(Z'Z\right)^{-1}Z'$ being the idempotent projection matrix defined with $Z = [z_{k+1}, \ldots, z_T]'$ and $z_t = (1, \Delta x_{t-1}, \ldots, \Delta x_{t-k}) = (1, \xi_t')$. The WSLS estimator proposed in Fuller (1996) is:

$$\hat{\alpha}_W = \frac{\sum_{t=k+2}^{T} \hat{x}_t \hat{x}_{t-1}}{\sum_{t=k+2}^{T-1} \hat{x}_t^2 + T^{-1} \sum_{t=k+1}^{T} \hat{x}_t^2}$$

and

$$\hat{\sigma}_W^2 = \frac{\sum_{t=k+2}^{T} \left(\hat{x}_t - \hat{\alpha}_W \hat{x}_{t-1}\right)^2}{(T-k-1)\left[\sum_{t=k+2}^{T-1} \hat{x}_t^2 + T^{-1} \sum_{t=k+1}^{T} \hat{x}_t^2\right]}$$

so that:

$$\hat{\tau}_W = \frac{\hat{\alpha}_W - 1}{\hat{\sigma}_W} \tag{9}$$

is the pseudo t-ratio statistic to test the null hypothesis that $\alpha = 1$ in (8). The limiting distribution of $\hat{\tau}_W$ under the null hypothesis of unit root can be found in Fuller (1996), pp. 570 and following.

The modification in (7) requires the definition of $C\left(\hat{\tau}_W\right)$ that is given by the discontinuous function:

$$C\left(\hat{\tau}_W\right) = \begin{cases} -\hat{\tau}_W & \text{if } \hat{\tau}_W > \tau_{pct} \\ I_{k+1}T^{-1}\hat{\tau}_W - 2\left[\hat{\tau}_W + K\left(\hat{\tau}_W + A\right)\right]^{-1} & \text{if } -A < \hat{\tau}_W \leq \tau_{pct} \\ I_{k+1}T^{-1}\hat{\tau}_W - 2\left[\hat{\tau}_W\right]^{-1} & \text{if } -\left(2T\right)^{1/2} < \hat{\tau}_W \leq -A \\ 0 & \text{if } \hat{\tau}_W \leq -\left(2T\right)^{1/2} \end{cases} \tag{10}$$

with $K = \left[\left(1 + I_{k+1}T^{-1}\right)\tau_{pct}\left(\tau_{pct} + A\right)\right]^{-1}\left[2 - I_{k+1}T^{-1}\tau_{pct}^2\right]$, $I_{k+1} = \lfloor (k+2)/2 \rfloor$, $\lfloor \cdot \rfloor$ being the integer part, while k denotes the order of the autoregressive correction in (8) – note that for an AR(p) process $k = p - 1$ – and τ_{pct} is a percentile of the limiting distribution of $\hat{\tau}_W$ when $\alpha = 1$. The percentile τ_{pct} is either set at the median (τ_{50}) or at the 85th percentile (τ_{85}) of the distribution of $\hat{\tau}_W$, which are reported in the last row of Table 1. Finally, the function K depends on the deterministic specification used in

(1) – i.e., a constant or a linear time trend.[1] The value of the constant A is empirically chosen in Roy and Fuller (2001) after conducting simulation experiments, which is set at $A = 5$ for unbounded stochastic processes.[2] Taking into account $\hat{\alpha}_{TW}$, Perron and Yabu (2009a) defined the super-efficient estimator:

$$\hat{\alpha}_{PY} = \begin{cases} \hat{\alpha}_{TW} & \text{if } |\hat{\alpha}_{TW} - 1| > T^{-1/2} \\ 1 & \text{if } |\hat{\alpha}_{TW} - 1| \leq T^{-1/2} \end{cases} \tag{11}$$

an estimator that is aimed at correcting the downward estimation bias of α when it is near one.

This type of estimator has become popular in the literature since it can be used for other purposes such as testing for a linear trend in the presence of autoregressive processes. Although this framework is outside the scope of the paper, it should be mentioned that Roy *et al.* (2004) build upon the truncated estimator in Roy and Fuller (2001) to test for a linear time trend in autoregressive processes. However, Perron and Yabu (2012) documented errors in both the theoretical and simulation results reported in Roy *et al.* (2004) that, in fact, were supporting the method suggested by Perron and Yabu (2009a). Extensions of the use of the super-efficient estimator to test for the presence of one or multiple structural breaks affecting a linear time trend can be found in Perron and Yabu (2009b) and Kejriwal and Perron (2010). Therefore, the contribution of this paper also shows that the super-efficient estimator can be useful when dealing with bounded nearly-integrated stochastic processes.

Performance of the standard bias correction methods for bounded stochastic processes

To motivate interest in the proposal developed in this paper, this section examines whether dealing with bounded stochastic processes presents any different features compared to unbounded situations. This is addressed through a simulation experiment with the DGP defined by (1) to (3) with $\mu = 0$, $\alpha = 1 - \kappa/T$, $\kappa = \{0, 1, 2, \ldots, 30\}$, $-\underline{c} = \overline{c} = \{0.3, 0.5, 0.7\}$, $x_0 = 0$, $\varepsilon_t \sim iidN(0, 1)$, $T = 200$ and 1,000 replications. Throughout the paper, we use the algorithm described in Cavaliere (2005) and Cavaliere and Xu (2014) to generate the bounded stochastic processes. Subfigure A in Figure 1 presents the mean of the OLS estimated autoregressive parameter for an AR(1) model for different values of κ and $\overline{c}$. The upper-straight solid line represents the true value of α. As can be seen, the smaller $\overline{c}$ and κ, the bigger the estimation bias. Furthermore, the

[1] See Roy and Fuller (2001) for the function that corresponds to the linear time trend. It is worth noting that Perron and Yabu (2009b) use the same function when testing for multiple shifts in the trend.

[2] Roy and Fuller (2001) also set $A = 5$ for the linear time trends, whereas Perron and Yabu (2009b) specify $A = 10$.

estimation bias does not reduce as κ increases.

The results of the application of the MU, TW and PY estimation bias correction procedures discussed above are depicted in subfigures B to D of Figure 1 for $\overline{c} = \{0.3, 0.5, 0.7\}$, respectively. The three procedures lead to an improvement of the OLS estimation of the autoregressive parameter, with MU being more conservative than the TW and PY methods for I(1) and NI(1) – with $\kappa \leq 10$ for $\overline{c} = 0.5$ and $\kappa \leq 5$ for $\overline{c} = 0.7$ – stochastic processes. However, MU tends to produce slightly better results than TW as $\overline{c}$ and κ increase – see the results with $\kappa > 10$ for $\overline{c} = 0.5$ and $\kappa > 5$ for $\overline{c} = 0.7$ – although the differences are small. In general, the PY estimator tends to over-estimate α as $\overline{c}$ and κ (≤ 20) increase, although all corrections produce almost the same outcome for $\kappa > 20$. These features indicate that the bounded nature of time series should be accounted for when estimating autoregressive model specifications. In what follows, we discuss how the bias correction procedures described above can be modified to consider that $x_t \in \left[\underline{b}, \overline{b}\right]$.

IV. Estimation bias correction methods for bounded stochastic processes

The extension of Andrews (1993) and Andrews and Chen (1994) to cover the case of bounded stochastic processes requires consideration of the restriction that $x_t \in \left[\underline{b}, \overline{b}\right]$ when the empirical distribution of α is approximated by means of simulation experiments. Thus, once $\hat{\alpha}$ is obtained from the OLS estimation of either (5) or (6), the Monte Carlo simulation used to compute the empirical distribution of α, under the assumption that $\alpha = \hat{\alpha}$, will use simulated stochastic processes that satisfy $x_t \in \left[\underline{b}, \overline{b}\right]$ – or, equivalently, $\sigma^{-1}T^{-1/2}\left(x_t - \mu\right) \in [\underline{c}, \overline{c}]$. This generates an intensive computational problem since look-up tables have to be obtained for different combinations of $[\underline{c}, \overline{c}]$ values, for different values of p and for specific values of T when working on finite samples. To solve this problem, a Matlab code is available from the authors to compute look-up tables for any combinations of set bounds, p and T values.

The WSLS-based estimation procedures of Roy and Fuller (2001) and Perron and Yabu (2009a) depend on the bias correction term $C\left(\hat{\tau}_W\right)$ defined in (10), which involves two important elements: (i) the percentile (τ_{pct}) of the distribution of $\hat{\tau}_W$ and (ii) the constant A. As for τ_{pct}, the limiting distribution of $\hat{\tau}_W$ was derived by Fuller (1996) for the case of unbounded stochastic processes so that τ_{pct} can be approximated by simulation. However, Cavaliere (2005) shows that the limiting distribution of unit root statistics depends on the presence of bounds – Cavaliere (2015) and Cavaliere and Xu (2014) analyze the unit root statistics in Phillips and Perron (1988) and Ng and Perron (2001). The following theorem provides the limiting distribution for the $\hat{\tau}_W$ statistic defined in (9) generalized to bounded NI(1) and I(1) stochastic processes.

Theorem 2 *Let $\{x_t\}_{t=1}^T$ be the stochastic process given by (1) to (3), with $\alpha = \exp(-\kappa/T)$, $\kappa \geq 0$. Following Chang and Park (2002), let k in (8) be chosen in such a way that $1/k + k^2/T \to 0$ as $T \to \infty$. The pseudo t-ratio statistic defined in (9) converges as $T \to \infty$ to:*

$$\hat{\tau}_W \Rightarrow \frac{\frac{1}{2}\left[V_{\underline{c}}^{\bar{c},\kappa}(1)^2 - 1\right] - \int_0^1 V_{\underline{c}}^{\bar{c},\kappa}(r)^2\, dr}{\sqrt{\int_0^1 V_{\underline{c}}^{\bar{c},\kappa}(r)^2\, dr}}$$

where $\Rightarrow$ denotes weak convergence towards the associated measure of probability and $V_{\underline{c}}^{\bar{c},\kappa}(r) = J_{\underline{c}}^{\bar{c},\kappa}(r) - r\int_0^1 J_{\underline{c}}^{\bar{c},\kappa}(s)\, ds$ defines a demeaned regulated Ornstein-Uhlenbeck process.

The proof can be found in the appendix. Table 1 summarizes selected percentiles of the distribution of $\hat{\tau}_W$ for different values of (symmetric) bound parameters under the null hypothesis of unit root ($\kappa = 0$). As mentioned above, the last row shows the percentiles for unbounded stochastic processes. As can be seen, the limiting distribution of $\hat{\tau}_W$ depends on the bounds. The more a limiting distribution is shifted to the left, the narrower the range of variation defined by the bounds. This clearly affects the definition of the bias correction term $C(\hat{\tau}_W)$ given in (10).

Let us now focus on the constant A that also appears in the bias correction term $C(\hat{\tau}_W)$ and consider the median of the distribution ($\tau_{0.5}$) as the percentile used in the bias correction. First, note that setting $A = 5$, as is done for the unbounded stochastic process case, does not cause incongruence for the definition of $C(\hat{\tau}_W)$ since $-A < \tau_{0.5}$.[3] However, Table 1 shows that $\tau_{0.5}$ moves away from -1.21 as the range of variation defined by the bounds decreases, which might produce a poor performance of the correction when $\bar{c} < 0.5$.[4] In this regard, an extensive simulation experiment has been conducted to assess the sensitivity of $\hat{\alpha}_{TW}$ to the constant $A = \{5, 6, \ldots, 15\}$. Results available upon request indicate that $\hat{\alpha}_{TW}$ shows a good performance when $A = 5$ and $\bar{c} > 0.1$, and only marginal differences are found for the other values of A. Besides, for small values of the bound parameter ($\bar{c} \leq 0.1$), we find that $A = 10$ gives good results.

Taking into account the $\hat{\alpha}_{TW}$ estimator generalized for bounded stochastic processes, we can proceed to compute the super-efficient estimator $\hat{\alpha}_{PY}$ in Perron and Yabu (2009a), defined in (11), to correct the downward estimation bias of α when it is near one.

[3] This is also valid for the linear time trend case, for which Roy and Fuller (2001) estimated $\tau_{0.5} = -1.96$ and $A = 5$, as mentioned above. Note that the consideration of slope trend shifts in Perron and Yabu (2009b) led them to specify $A = 10$ for the one break case – it is well known that the limiting distribution of $\hat{\tau}_W$ shifts to the left as the number of structural breaks increases.

[4] Our guess is based on the fact that Roy and Fuller (2001) define $A = 5$ for the linear time trend case, for which the median of the distribution of $\hat{\tau}_W$ is $\tau_{0.5} = -1.96$. Consequently, we might expect that $A = 5$ is also valid for cases where $\bar{c} \geq 0.5$, although it should be borne in mind that the K function involved in the correction depends on the deterministic specification.

Implementation of the estimation procedures

In practice, the empirical computation of the bias correction methods for bounded time series that we propose needs some additional steps. Given a time series with known theoretical limits $\underline{b}$ and $\overline{b}$, we can proceed to estimate the bounds as:

$$\left[\underline{\hat{c}}, \overline{\hat{c}}\right] = \left[\frac{\underline{b} - \hat{D}_t}{\hat{\sigma}T^{1/2}}, \frac{\overline{b} - \hat{D}_t}{\hat{\sigma}T^{1/2}}\right]$$

which requires an estimation of the deterministic component (D_t) and the long-run variance (σ^2). Following Cavaliere and Xu (2014), the deterministic component is estimated under the null hypothesis of unit root so that $\hat{D}_t = x_0$. The estimation of the long-run variance deserves further attention because its estimation might suffer from estimation bias problems through the autoregressive parameter estimates. To address this issue we suggest implementing the following iterative estimation method:

(i) Estimate the LRV ignoring the bounds. In this regard, we can use the parametric estimation method proposed in Ng and Perron (2001) and Perron and Qu (2007), which also allows us to select the optimal lag of the autoregressive model.

(ii) Compute an initial educated estimate of the bounds:

$$\left[\underline{\hat{c}}_0, \overline{\hat{c}}_0\right] = \left[\frac{\underline{b} - x_0}{\hat{\sigma}_0 T^{1/2}}, \frac{\overline{b} - x_0}{\hat{\sigma}_0 T^{1/2}}\right]$$

where the subscript 0 in $\underline{\hat{c}}_0$, $\overline{\hat{c}}_0$ and $\hat{\sigma}_0$ indicates the initial estimate of the corresponding quantity.

(iii) Estimate α according to one of these procedures:

  (a) For the MU-based procedure, compute the look-up tables corresponding to $\left[\underline{\hat{c}}_0, \overline{\hat{c}}_0\right]$ by simulation and obtain $\hat{\alpha}_{MU}$.

  (b) For the truncated WSLS-based procedure, compute the percentiles of the $\hat{\tau}_W$ distribution corresponding to $\left[\underline{\hat{c}}_0, \overline{\hat{c}}_0\right]$ by simulation – see Table 1 for the symmetric bounds case – and obtain $\hat{\alpha}_{TW}$ or $\hat{\alpha}_{PY}$ defined above.

(iv) Use $\hat{\alpha}_m$, $m = \{MU, TW, PY\}$, from the previous step to estimate the LRV again as follows,

$$y_t - \hat{\alpha}_m y_{t-1} = \mu + \sum_{j=1}^{k} \psi_j \Delta y_{t-j} + \varepsilon_t \tag{12}$$

$$\hat{\sigma}_1^2 = \frac{T^{-1}\sum_{t=1}^{T}\hat{\varepsilon}_t^2}{(1-\hat{\alpha}_m)^2}$$

where $\hat{\alpha}_m$ is imposed in (12), but the other parameters are freely estimated by OLS.

(v) Update bounds estimates as:

$$\left[\hat{\underline{c}}_1, \hat{\overline{c}}_1\right] = \left[\frac{\underline{b} - x_0}{\hat{\sigma}_1 T^{1/2}}, \frac{\overline{b} - x_0}{\hat{\sigma}_1 T^{1/2}}\right]$$

(vi) Iterate until $\left|\sum_{t=1}^{T} \hat{\varepsilon}_{t,l}^2 - \sum_{t=1}^{T} \hat{\varepsilon}_{t,l-1}^2\right| < Tol$, where Tol is the desired level of tolerance and l the step of iteration.

The implementation of the procedure can be done without performing iterations (steps (i) to (iii), i.e., non-iterative scheme) or carrying out multiple iterations (steps (i) to (vi)) – we denote by $\hat{\alpha}_m^*$, $m = \{MU, TW, PY\}$, the resulting estimators. The potential gain of the multiple iterative estimation method is assessed through Monte Carlo simulation experiments in the next section.

V. Monte Carlo simulations

This section analyzes the performance of the different bias correction methods discussed above using a DGP based on equations (1) to (3):

$$\begin{aligned} x_t &= \mu + y_t \\ y_t &= \alpha y_{t-1} + \psi \Delta y_{t-1} + u_t + \theta u_{t-1} \end{aligned} \tag{13}$$

with $\mu = 0$, $\alpha = 1 - \kappa/T$, $\kappa = \{0, 1, 2, \ldots, 30\}$ and $\varepsilon_t \sim iid\ N(0, 1)$. The simulation exercise focuses on three stochastic processes: (i) AR(1) when imposing $\psi = \theta = 0$ in (13), (ii) AR(2) when setting $\psi = 0.5$ and $\theta = 0$ in (13) and (iii) ARMA(1,1) when fixing $\psi = 0$ and $\theta = \{-0.8, -0.4\}$ in (13). Symmetric bounds are defined as $[\underline{c}, \overline{c}] = [-\overline{c}, \overline{c}]$, $\overline{c} = \{0.3,$ $0.5, 0.7, 0.9\}$, and we consider the general case of unknown p so that k in (8) is estimated using the BIC information criterion specifying a maximum of $k_{\max} = \left\lfloor 12(T/100)^{1/4} \right\rfloor$ lags.[5] This implies that the MU estimation procedure implemented is the one described in Andrews and Chen (1994). Two sample sizes are used, $T = \{50, 200\}$, and 1,000 replications are conducted throughout all simulation experiments. The discussion is organized according to whether $[\underline{c}, \overline{c}]$ are treated as known or unknown – note, however, that in the latter case $\left[\underline{b}, \overline{b}\right]$ are assumed to be known.

The AR(1) case

Figure 2.a presents the mean of the empirical distribution of the different estimators studied for the AR(1) case when $[\underline{c}, \overline{c}]$ are known. The result based on the iterative

[5] This maximum number of lags is set throughout the simulation experiment section. Simulations available upon request also assessed the MAIC information criterion proposed in Ng and Perron (2001), although the use of the BIC gives a better overall performance.

algorithm to estimate $[\underline{c}, \overline{c}]$ appears below.[6] Note that all figures include two solid straight lines that designate the true value of α for each T. The general conclusion is that the downward estimation bias shown by $\hat{\alpha}$ is clearly corrected by $\hat{\alpha}_{MU}$, $\hat{\alpha}_{TW}$ and $\hat{\alpha}_{PY}$, although their relative performance depends on $\overline{c}$, κ and T.

First, for $T = 50$ we can see that $\hat{\alpha}_{MU}$ leads to higher bias correction than $\hat{\alpha}_{TW}$ and $\hat{\alpha}_{PY}$ when $\overline{c} = 0.3$, whereas the opposite is found when using $\hat{\alpha}_{PY}$ for $\overline{c} > 0.3$. If we compare the estimators, $\hat{\alpha}_{MU}$ and $\hat{\alpha}_{TW}$ produce similar results, but with $\hat{\alpha}_{TW}$ outperforming $\hat{\alpha}_{MU}$ in the unit root case. However, $\hat{\alpha}_{MU}$ seems to feature a mild (increasing) over-estimation distortion for $\kappa \geq 20$, while $\hat{\alpha}_{TW}$ and $\hat{\alpha}_{PY}$ are only slightly above α. Additional simulations using $\alpha = \{0, 0.1, 0.2, ..., 1\}$, not reported here to save space, confirm these results when dealing with less persistent processes.

Second, $\hat{\alpha}_{MU}$ outperforms $\hat{\alpha}_{TW}$ and $\hat{\alpha}_{PY}$ for $T = 200$ and $\overline{c} = 0.3$, although the former tends to over-estimate α when $\kappa \geq 10$. The converse situation is found for $\overline{c} > 0.3$, with $\hat{\alpha}_{PY}$ almost ($\overline{c} = 0.5$) and fully ($\overline{c} > 0.5$) correcting the estimation bias for $\kappa = 0$, and leading to mild over-estimates of α for $0 < \kappa \leq 20$. $\hat{\alpha}_{MU}$ and $\hat{\alpha}_{TW}$ behave in a similar way. Interestingly, $\hat{\alpha}_{PY}$ tends to be located slightly below α for $\kappa > 20$, whereas $\hat{\alpha}_{MU}$ shows the over-estimation distortions mentioned above. These features are also found for smaller values of α.

Figure 2.b shows the simulation results when $[\underline{c}, \overline{c}]$ are unknown. For ease of comparison, we also include the results that assume that $[\underline{c}, \overline{c}]$ are known.[7] At first sight, we can establish a clear distinction in the performance of the estimators depending on whether $\overline{c} = 0.3$ or $\overline{c} > 0.3$. First, for $\overline{c} = 0.3$ and $T = 50$, both estimators lead to lower bias corrections compared to the known $\overline{c}$ situation when $\kappa \leq 5$. $\hat{\alpha}^*_{TW}$ and $\hat{\alpha}^*_{PY}$ follow the same pattern and show an over-estimation distortion for $\kappa > 15$, whereas this feature is found for $\hat{\alpha}^*_{MU}$ when $\kappa > 25$. The performance improves for $T = 200$, with $\hat{\alpha}^*_{MU}$ giving lower estimates of α than $\hat{\alpha}_{MU}$, whereas $\hat{\alpha}^*_{PY}$ improves with respect to $\hat{\alpha}_{PY}$. Note that $\hat{\alpha}^*_{TW}$ is encompassed by $\hat{\alpha}^*_{PY}$ when $\kappa \leq 15$, and they become equivalent for $\kappa > 15$. If we compare $\hat{\alpha}^*_{MU}$ and $\hat{\alpha}^*_{PY}$, we conclude that $\hat{\alpha}^*_{MU}$ outperforms $\hat{\alpha}^*_{PY}$ when $\kappa \leq 15$. However, $\hat{\alpha}^*_{MU}$ shows a mild over estimation bias for $\kappa > 15$, whereas $\hat{\alpha}^*_{PY}$ (and $\hat{\alpha}^*_{TW}$) is placed below α.

As $\overline{c}$ increases, the predominance of $\hat{\alpha}^*_{PY}$ over $\hat{\alpha}^*_{MU}$ becomes evident when $T = 50$, since $\hat{\alpha}^*_{PY}$ provides similar ($\overline{c} = 0.5$) or better ($\overline{c} > 0.5$) bias corrections than $\hat{\alpha}^*_{MU}$ when $\kappa \leq 10$. The two estimators can hardly be distinguished in the range $10 < \kappa \leq 25$, but $\hat{\alpha}^*_{PY}$ (and $\hat{\alpha}^*_{TW}$) outperforms $\hat{\alpha}^*_{MU}$ when $\kappa > 25$. For $T = 200$ the dominance of $\hat{\alpha}^*_{PY}$ is clear when $\kappa \leq 5$, although it tends to over-estimate α when $5 < \kappa \leq 15$, a range

[6] Similar results were obtained with the use of $\tau_{pct} = \tau_{85}$ so that, in order to save space, in what follows we only focus on τ_{50}. Consequently, unless required, we remove the reference to the percentile in the supscript to simplify the notation.

[7] The TW estimator is excluded to reduce the number of curves in the figures, although some comments are provided – the complete set of figures is available in the appendix.

in which $\hat{\alpha}^*_{MU}$ produces better results. Finally, $\hat{\alpha}^*_{MU}$ leads to a mild over-estimation distortion, while $\hat{\alpha}^*_{PY}$ (and $\hat{\alpha}^*_{TW}$) does not.

All in all, and regardless of whether $[\underline{c}, \overline{c}]$ are assumed to be known or unknown, the MU approach outperforms the TW and PY approaches for narrow ranges of variation ($\overline{c} = 0.3$), even though $\hat{\alpha}_{MU}$ leads to mild over-estimates when $\kappa \geq 20$. TW and PY outperform MU when $\overline{c} > 0.3$, although PY tends to over-estimate α for $0 < \kappa \leq 20$, a distortion that is clearly corrected for $\kappa > 20$. This phenomenon can be attributed to the inherent super-efficient correction device introduced by the PY procedure. Finally, the results for the AR(2) case are qualitatively similar to those obtained for the AR(1) model. They are available from the authors upon request.

The ARMA(1,1) case

Figures 3 and 4 show the simulation results for the ARMA (1,1) case. As can be seen, the introduction of a MA(1) component causes a downward bias of the α estimate, regardless of $\overline{c}$ and the estimation method. The estimation bias is more severe, first, the smaller θ is and, second, as $-\theta$ approaches α. The latter is to be expected since in these cases we are near to the common factor situation. Note that there is a common factor when $\alpha = -\theta$, in which case the estimated autoregressive coefficient should approach zero. It is worth studying this situation with more detail using the "nearly white noise - nearly integrated" framework described in Nabeya and Perron (1994), with the DGP given by (1), but with a slight modification of (2) to accommodate a local to unit MA root:[8]

$$y_t = \alpha y_{t-1} + u_t + \theta_T u_{t-1} \tag{14}$$

where $\alpha = \exp(-\kappa/T)$, $\kappa \geq 0$, $\theta_T = -1+\delta/T^{1/2}$, $\delta \geq 0$, with $y_0 = O_p(1)$ and $u_0 = O_p(1)$. The following theorem provides the limiting distribution of the OLS estimator in this case.

Theorem 3 *Let $\{x_t\}_{t=1}^T$ be the stochastic process given by (1), (3) and (14), with $\alpha = \exp(-\kappa/T)$, $\kappa \geq 0$, and $\theta_T = -1 + \delta/T^{1/2}$, $\delta \geq 0$. Then, as $T \to \infty$*

$$\hat{\alpha} \Rightarrow \frac{\gamma_{u,1} + \delta^2\sigma^2 \int_0^1 V_{\underline{c}}^{\overline{c},\kappa}(r)^2\, dr}{\sigma_u^2 + \delta^2\sigma^2 \int_0^1 V_{\underline{c}}^{\overline{c},\kappa}(r)^2\, dr}$$

The proof is given in the appendix. As can be seen, the OLS estimator converges towards a random variable, unless $\delta = 0$. However, even in such a case $\hat{\alpha} \xrightarrow{p} \gamma_{u,1}/\sigma_u^2 \neq 0$ when $\delta = 0$, with $\gamma_{u,1}$ and σ_u^2 being the first autocovariance and variance of u_t, respectively. Therefore, under the near common factor situation, $\hat{\alpha}$ is not a consistent

[8] While the use of the "nearly white noise" concept in Nabeya and Perron (1994) is fine, it is not appropriate in our framework since, in general, u_t is not white noise when dealing with bounded stochastic processes.

estimator of α. Besides, $\hat{\alpha} \xrightarrow{p} 1$ when $\delta \to \infty$ regardless of whether the stochastic process is bounded, which explains why $\hat{\alpha}$ is a consistent estimator of α in the NI(1) and I(1) scenarios. These features can help to interpret the performance of the different estimators investigated.

Consider first the known $[\underline{c}, \overline{c}]$ case, for which Figure 3.a depicts the results when $\theta = -0.4$. Since $\hat{\alpha}_{TW}$ and $\hat{\alpha}_{PY}$ give almost equivalent results, in what follows we only refer to $\hat{\alpha}_{PY}$. For $T = 50$ and except for $\overline{c} = 0.3$, $\hat{\alpha}_{PY}$ outperforms $\hat{\alpha}_{MU}$ when $\kappa = 0$, whereas the opposite is found when $\kappa > 0$. For $T = 200$ the bias correction of both methods improves, with $\hat{\alpha}_{PY}$ being superior to $\hat{\alpha}_{MU}$ for $0 < \kappa \leq 10$, and the other way around for $\kappa > 10$. Note that with $\theta = -0.4$ we are far from the common factor situation so that all estimators are close to α in the unit root region, as predicted by Theorem 3. Both methods produce poor results when $\theta = -0.8$, although now we have to bear in mind that we are close to the common factor situation for high values of α – see Figure 4.a. For $T = 50$, $\hat{\alpha}_{PY}$ gives better results than $\hat{\alpha}_{MU}$ when $\kappa = 0$, but it tends to deviate more than $\hat{\alpha}_{MU}$ from α as κ increases. The performance of both methods improves as T increases, with $\hat{\alpha}_{PY}$ giving better results than $\hat{\alpha}_{MU}$ for small values of κ – say, $\kappa \leq 5$ – whereas the converse situation is found for $\kappa > 5$.

Let us now focus on the unknown $[\underline{c}, \overline{c}]$ case with results depicted in Figures 3.b and 4.b for $\theta = -0.4$ and $\theta = -0.8$, respectively. Again, the performance of the statistics depends on θ, $\overline{c}$ and T. When $\theta = -0.4$ and $T = 50$, both $\hat{\alpha}^*_{MU}$ and $\hat{\alpha}^*_{PY}$ produce lower bias corrections than $\hat{\alpha}_{MU}$ and $\hat{\alpha}_{PY}$, although these differences disappear when $T = 200$. $\hat{\alpha}^*_{MU}$ is superior to $\hat{\alpha}^*_{PY}$ when $T = 50$ and $\overline{c} \leq 0.5$, although the converse is found for $\overline{c} > 0.5$. As T increases to $T = 200$, $\hat{\alpha}^*_{PY}$ is equally good ($\overline{c} = 0.3$) or clearly outperforms ($\overline{c} > 0.3$) $\hat{\alpha}^*_{MU}$ when $\kappa \leq 5$. There is a range of κ values for which $\hat{\alpha}^*_{MU}$ and $\hat{\alpha}^*_{PY}$ render equivalent results, but $\hat{\alpha}^*_{MU}$ starts dominating $\hat{\alpha}^*_{PY}$ when $\kappa > 15$.

In general, similar conclusions are found for $\theta = -0.8$, although now the differences between $\hat{\alpha}_{MU}/\hat{\alpha}_{PY}$ and $\hat{\alpha}^*_{MU}/\hat{\alpha}^*_{PY}$ persist even when $T = 200$, $\hat{\alpha}_{MU}$ and $\hat{\alpha}_{PY}$ showing the better results. This might be due to the fact that now we are close to the common factor situation, which might imply a poorer estimation of the long-run variance required to approximate $[\underline{c}, \overline{c}]$. Contrary to what has been found for $\theta = -0.4$, now $\hat{\alpha}^*_{MU}$ seems to be superior to $\hat{\alpha}^*_{PY}$ when $T = 50$ and $\kappa \leq 10$, although its behavior deteriorates for $\kappa > 10$, with a performance that is worse than that shown by the OLS estimator. $\hat{\alpha}^*_{PY}$ is equally good or even better than the OLS estimator. Note that this feature is found regardless of $\overline{c}$. A neater picture is reached when T increases to $T = 200$, since now $\hat{\alpha}^*_{PY}$ outperforms $\hat{\alpha}^*_{MU}$ when $\kappa = 0$. Further, we can observe a region defined by small values of κ where both estimators are equally good. Finally, $\hat{\alpha}^*_{MU}$ clearly gives better results than $\hat{\alpha}^*_{PY}$ for $\kappa \geq 5$.

All these elements lead to the conclusion that, in large samples and regardless of whether $[\underline{c}, \overline{c}]$ are assumed to be known or unknown, the PY procedure corrects bias distortions more satisfactorily than the MU for small values of κ – in particular, for $\kappa = 0$. As we move away from the unit root, the MU approach seems to outperform the PY, although the former over-estimates α as κ increases – this characteristic has become evident for AR(1) and AR(2) stochastic processes. Besides, $\hat{\alpha}_{PY}$ does not suffer from this problem, providing very good estimates for large T in those cases.

VI. Empirical illustration

The estimation methods proposed above have been applied in a study of the persistence in the unemployment rate of the G7 economies. Annual harmonized unemployment rates covering the period from 1955 (or later, depending on the country) to 2018 have been obtained from the OECD Stat database – see Figure 5 for a visual inspection. Although there is a flurry of papers in the economic literature addressing unemployment persistence, the available results are scarcely robust. Due to the high inertia of unemployment rates and their limited nature, these series are especially suitable for illustrating the estimation bias for bounded near-integrated processes. Furthermore, with the exception of Cavaliere (2005), most of the empirical contributions ignore the potential presence of bounds, which might lead to misleading conclusions about the degree of unemployment persistence. The model that has been estimated for each of the G7 economies is given by:

$$unem_{i,t} = \mu_i + \alpha_i unem_{i,t-1} + \sum_{j=1}^{k_i} \psi_{i,j} \Delta unem_{i,t-j} + \varepsilon_{i,t} \tag{15}$$

$i = 1, \ldots, 7$ and $t = T_0, \ldots, 2018$, where T_0 aims for the initial year that varies depending on the country. The order of the autoregressive correction in (15) has been selected using the BIC with a maximum of $k_{\max} = \left[(12(T/100)^{1/4}\right]$ lags.

The results presented in Table 2 reflect heterogeneous estimates for the G7 countries. When the bounds are ignored, we can observe that the range of variation of the OLS estimates goes from 0.71 (USA) to 0.96 (Japan). The use of the MU estimators slightly raises the estimated persistence and the range now goes from 0.74 (USA) to 1 (Japan and Germany). Finally, similar results are obtained for $\hat{\alpha}^*_{PY}$, which provides estimates between 0.75 (USA) and 0.97 (Japan).[9] As a consequence, the omission of bounds would lead us to consider that the persistence of the unemployment rate is relatively low, except for Germany and Japan.

The picture changes when the presence of bounds is accounted for, since unemployment persistence clearly increases in all cases. It should be mentioned that we have conducted the analysis using two different sets of boundaries $\left[\underline{b}, \overline{b}\right]$. First, we have the boundaries that derive from the unemployment rate definition, i.e., $\left[\underline{b}, \overline{b}\right] = [0, 100]$. The

[9] We only report the results using τ_{50} since they are equivalent to the ones based on τ_{85}.

second set of boundaries is established following the strategy set out in Herwartz and Xu (2008), who consider a potential set of boundaries arising from the observed country-specific minimum ($\underline{b}_i = \min(unem_{i,t})$) and maximum ($\overline{b}_i = \max(unem_{i,t})$) values of $unem_t$ – which defines $[\underline{\hat{c}}_i, \overline{\hat{c}}_i] = \left[(\underline{b}_i - \hat{D}_{i,t})/(\hat{\sigma}_i T^{1/2}), (\overline{b}_i - \hat{D}_{i,t})/(\hat{\sigma}_i T^{1/2})\right]$, $i = 1, \ldots, 7$. In addition, following Herwartz and Xu (2008), we increase this initial range up to 300 per cent in absolute value – i.e., $[\underline{\hat{c}}_i - \delta\omega_i/2, \overline{\hat{c}}_i + \delta\omega_i/2]$, $\omega_i = |\overline{\hat{c}}_i - \underline{\hat{c}}_i|$ and $\delta = \{0, 0.1, 0.2, 0.3, \ldots, 1, 1.5, 2, 2.5, 3\}$ – so that the robustness of the analysis can be tested using different sets of bounds. The key issue here is how to select among these values of bounds. The suggestion in Herwartz and Xu (2008) is based on the p-values of the augmented Dickey-Fuller (ADF) unit root test, so that the bounds are selected in such a way that the p-values of the ADF statistic with and without bounds equalize – the so-called "break-even" boundaries, which are denoted by $[\underline{b}_i, \overline{b}_i] = [\underline{b}_i^*, \overline{b}_i^*]$, warrant a minimum range under which the standard ADF unit root test does not suffer from over-sizing.

In any event, the results are quite similar regardless of the boundaries used. We observe that the USA unemployment rate again exhibits the lowest persistence, with the estimation of the autoregressive parameter never exceeding 0.78. The opposite cases are Germany and Japan, for which the estimations of the autoregressive parameters are always 1. The rest of the cases show mixed results in that $\hat{\alpha}^*_{PY}$ also equals one. By contrast, the use of the median-unbiased corrections provides lower estimated values. Consequently, these results are in line with the evidence described in the previous sections and, basically, illustrate that ignoring the bounded nature of some economic variables reduces the estimated degree of persistence. When these bounds are taken into account the results clearly change, leading to higher persistence estimates.

VII. Conclusions

This paper addresses the issue of the estimation of autoregressive models when the stochastic process being studied is influenced by the presence of bounds that regulate its evolution. We consider standard techniques proposed in the literature aimed at correcting the finite sample estimation bias of autoregressive parameters. Initial motivating simulation experiments show that the presence of bounds clearly distorts the performance of these types of estimators when bounds are ignored. The more limited the stochastic process – i.e., the narrower the fluctuation bands – the higher the estimation bias distortion. In order to remove this effect, we have modified the methods proposed by Andrews (1993), Andrews and Chen (1994), and Perron and Yabu (2009a) to account for the bounded nature of time series.

Simulation experiments have evidenced that these extensions are quite helpful in order to appropriately determine shock persistence for bounded stochastic processes. All the procedures investigated in the paper outperform the OLS estimation. Although the differences among the bounds-generalized procedures are minimal, we have found that the Perron and Yabu (2009a) proposal produces better results for highly persistent stochastic processes, whereas the median-unbiased approach tends to be preferred when we move away from the unit root neighborhood. Finally, we have applied these new methods to the analysis of the unemployment shock persistence for the G7 countries. Our results show that the use of the proposed methods improves our knowledge of the stochastic properties of the variables under study, allowing us to carry out more accurate shock persistence analysis.

Appendix A: Mathematical appendix

Lemma 1 *Let* $\{y_t\}_{t=1}^{T}$ *be a stochastic process generated according to (2) and (3) with* $\alpha = \exp(-\kappa/T)$, $\kappa \geq 0$, *and satisfying Assumptions 1 to 3. As* $T \to \infty$, $\sigma^{-1}T^{-1/2}y_t \Rightarrow J_{\underline{c}}^{\bar{c},\kappa}(r)$, *with* $\underline{c} \leq 0 \leq \bar{c}$, $\underline{c} \neq \bar{c}$, *where* $J_{\underline{c}}^{\bar{c},\kappa}(r) = J^{\kappa}(r) + L(r) - U(r)$ *denotes a standard regulated Ornstein-Uhlenbeck (OU) process being* $J^{\kappa}(r) = \int_0^r \exp(-\kappa(r-s))\, dB(s)$ *a standard OU process,* $B(r)$ *a standard Brownian motion, and* $L(r) = -\{0 \wedge \inf_{0 \leq r' \leq r}$ $(J^{\kappa}(r') - \underline{c})\}$ *and* $U(r) = \{0 \wedge \inf_{0 \leq r' \leq r}(\bar{c} - J^{\kappa}(r'))\}$ *the two side regulator processes.*

See Theorems 1 and 4 in Cavaliere (2005) for the proof.

Proof of Theorem 1

The model given in (1) and (2) can be written as:

$$\begin{aligned} x_t - \mu &= \alpha(x_{t-1} - \mu) + u_t \\ x_t &= (1-\alpha)\mu + \alpha x_{t-1} + u_t \end{aligned}$$

so that the OLS estimator of α is given by $\hat{\alpha} = \left(\hat{x}'_{-1}\hat{x}_{-1}\right)^{-1}\hat{x}'_{-1}\hat{x}$, where $\hat{x} = x - \bar{x}$, $\hat{x}_{-1} = x_{-1} - \bar{x}_{-1}$, $x_{-1} = (x_1, \ldots, x_{T-1})'$ and $x = (x_2, \ldots, x_T)'$. In terms of estimation bias we have:

$$\begin{aligned} \hat{\alpha} - \alpha &= \left(\hat{x}'_{-1}\hat{x}_{-1}\right)^{-1}\hat{x}'_{-1}\hat{u} \\ &= \left(\hat{x}'_{-1}\hat{x}_{-1}\right)^{-1}\hat{x}'_{-1}(\hat{\varepsilon} + \hat{\underline{\xi}} - \hat{\bar{\xi}}) \qquad \text{(Appendix A:.1)} \end{aligned}$$

Under Assumptions 1 to 3, and using Lemma 1 and the Functional Central Limit Theorem (FCLT), we can see that:

$$\begin{aligned} T^{-1/2}\hat{x}_t &= T^{-1/2}\left(x_t - \bar{x}\right) \\ &\Rightarrow \sigma\left(J_{\underline{c}}^{\bar{c},\kappa}(r) - r\int_0^1 J_{\underline{c}}^{\bar{c},\kappa}(s)\,ds\right) \\ &\equiv \sigma V_{\underline{c}}^{\bar{c},\kappa}(r) \end{aligned}$$

with $V_{\underline{c}}^{\bar{c},\kappa}(r) = J_{\underline{c}}^{\bar{c},\kappa}(r) - r\int_0^1 J_{\underline{c}}^{\bar{c},\kappa}(s)\,ds$ being a demeaned regulated OU process. Similarly, $T^{-2}\sum_{t=2}^{T-1}\hat{x}_t^2 \Rightarrow \sigma^2\int_0^1 V_{\underline{c}}^{\bar{c},\kappa}(r)^2\,dr$ and $T^{-1}\sum_{t=2}^{T}\hat{x}_{t-1}\hat{u}_t \Rightarrow \sigma^2\int_0^1 V_{\underline{c}}^{\bar{c},\kappa}(r)\,dV_{\underline{c}}^{\bar{c},\kappa}(r)$, so that

$$T\left(\hat{\alpha} - \alpha\right) = O_p(1)$$

and hence $(\hat{\alpha} - \alpha) \xrightarrow{p} 0$, with $\xrightarrow{p}$ denoting convergence in probability. □

Proof of Theorem 2

Following Fuller (1996), we can derive the expression of $\hat{\tau}_W$ in (8) working with the moments of the projected variable $\hat{x} = Mx$.[10] Note that:

$$\begin{aligned} \hat{x}_t &= x_t - z_t\left(Z'Z\right)^{-1}Z'x \\ &= x_t - z_t\left(T^{-1}Z'Z\right)^{-1}T^{-1}Z'x \end{aligned}$$

with $z_t = (1, \Delta x_{t-1}, \ldots, \Delta x_{t-k})$ so that:

$$T^{-1}Z'Z = \begin{bmatrix} 1 & T^{-1}\sum_{t=k}^{T}\Delta x_{t-1} & \cdots & T^{-1}\sum_{t=k}^{T}\Delta x_{t-k+1} \\ & T^{-1}\sum_{t=k}^{T}\Delta x_{t-1}^2 & \cdots & T^{-1}\sum_{t=k}^{T}\Delta x_{t-1}\Delta x_{t-k+1} \\ & & \ddots & \vdots \\ & & & T^{-1}\sum_{t=k}^{T}\Delta x_{t-k+1}^2 \end{bmatrix} = \begin{bmatrix} 1 & O_p\left(T^{-1/2}\right) \\ & \Sigma_{\Delta x} \end{bmatrix}$$

where $\Sigma_{\Delta x} = T^{-1}\xi'\xi$ denotes the covariance matrix of the first k lags of Δx, and:

$$T^{-3/2}Z'x = \begin{bmatrix} T^{-3/2}\sum_{t=k+1}^{T}x_t \\ T^{-3/2}\sum_{t=k+1}^{T}\Delta x_{t-1}x_t \\ \vdots \\ T^{-3/2}\sum_{t=k+1}^{T}\Delta x_{t-k+1}x_t \end{bmatrix} = \begin{bmatrix} O_p(1) \\ O_p\left(T^{-1/2}\right) \\ \vdots \\ O_p\left(T^{-1/2}\right) \end{bmatrix}$$

[10] Although we could work with the model given in (8), derivations using the orthogonal projected variable are neater. In any event, both ways are asymptotically equivalent as mentioned in Fuller (1996, pp. 416) when dealing with the unknown constant case.

since $T^{-3/2}\sum_{t=k+1}^{T} x_t \Rightarrow \sigma \int_0^1 J_{\underline{c}}^{\bar{c},\kappa}(r)\, dr$. Using these elements and the FCLT, we can see that:

$$\begin{aligned} T^{-1/2}\hat{x}_t &= T^{-1/2}x_t - z_t\left(T^{-1}Z'Z\right)^{-1} T^{-3/2}Z'x \\ &\Rightarrow \sigma\left(J_{\underline{c}}^{\bar{c},\kappa}(r) - r\int_0^1 J_{\underline{c}}^{\bar{c},\kappa}(s)\, ds\right) \\ &\equiv \sigma V_{\underline{c}}^{\bar{c},\kappa}(r) \end{aligned}$$

The numerator of the $\hat{\tau}_W$ statistic is given by:

$$\begin{aligned} \hat{\alpha}_W - 1 &= \frac{\sum_{t=k+2}^{T} \hat{x}_t \hat{x}_{t-1}}{\sum_{t=k+2}^{T-1} \hat{x}_t^2 + T^{-1}\sum_{t=k+1}^{T} \hat{x}_t^2} - 1 \\ &= \frac{\sum_{t=k+2}^{T} \hat{x}_{t-1}^2 + \sum_{t=k+2}^{T} \hat{x}_{t-1}\hat{e}_{k,t}}{\sum_{t=k+2}^{T-1} \hat{x}_t^2 + T^{-1}\sum_{t=k+1}^{T} \hat{x}_t^2} - 1 \\ &= \frac{\sum_{t=k+2}^{T} \hat{x}_{t-1}^2 + \sum_{t=k+2}^{T} \hat{x}_{t-1}\hat{e}_{k,t} - \sum_{t=k+2}^{T-1} \hat{x}_t^2 - T^{-1}\sum_{t=k+1}^{T} \hat{x}_t^2}{\sum_{t=k+2}^{T-1} \hat{x}_t^2 + T^{-1}\sum_{t=k+1}^{T} \hat{x}_t^2} \end{aligned}$$

which, once it has been properly rescaled, in the limit converges to:

$$T\left(\hat{\alpha}_W - 1\right) \Rightarrow \frac{\int_0^1 V_{\underline{c}}^{\bar{c},\kappa}(r)\, dV_{\underline{c}}^{\bar{c},\kappa}(r) - \int_0^1 V_{\underline{c}}^{\bar{c},\kappa}(r)^2\, dr}{\int_0^1 V_{\underline{c}}^{\bar{c},\kappa}(r)^2\, dr} = \frac{\frac{1}{2}\left[V_{\underline{c}}^{\bar{c},\kappa}(1)^2 - 1\right] - \int_0^1 V_{\underline{c}}^{\bar{c},\kappa}(r)^2\, dr}{\int_0^1 V_{\underline{c}}^{\bar{c},\kappa}(r)^2\, dr}$$

provided that $T^{-2}\sum_{t=k+2}^{T-1} \hat{x}_t^2 \Rightarrow \sigma^2 \int_0^1 V_{\underline{c}}^{\bar{c},\kappa}(r)^2\, dr$ and $T^{-1}\sum_{t=k+2}^{T} \hat{x}_{t-1}\hat{e}_{k,t} \Rightarrow \sigma^2 \int_0^1 V_{\underline{c}}^{\bar{c},\kappa}(r)$ $dV_{\underline{c}}^{\bar{c},\kappa}(r)$. If we now focus on the variance of $\hat{\alpha}_W$:

$$\hat{\sigma}_W^2 = \frac{\sum_{t=k+2}^{T} \left(\hat{x}_t - \hat{\alpha}_W \hat{x}_{t-1}\right)^2}{(T-k-1)\left[\sum_{t=k+2}^{T-1} \hat{x}_t^2 + T^{-1}\sum_{t=k+1}^{T} \hat{x}_t^2\right]}$$

we have that:

$$T^2\hat{\sigma}_W^2 \Rightarrow \left[\int_0^1 V_{\underline{c}}^{\bar{c},\kappa}(r)^2\, dr\right]^{-1}$$

provided that $(T-k-1)^{-1}\sum_{t=k+2}^{T} \left(\hat{x}_t - \hat{\alpha}_W \hat{x}_{t-1}\right)^2 \xrightarrow{p} \sigma^2$. Consequently, in the limit the $\hat{\tau}_W$ statistic converges to:

$$\begin{aligned} \hat{\tau}_W &= \frac{\hat{\alpha}_W - 1}{\hat{\sigma}_W} = \frac{T\left(\hat{\alpha}_W - 1\right)}{\sqrt{T^2\hat{\sigma}_W^2}} \\ &\Rightarrow \frac{\frac{1}{2}\left[V_{\underline{c}}^{\bar{c},\kappa}(1)^2 - 1\right] - \int_0^1 V_{\underline{c}}^{\bar{c},\kappa}(r)^2\, dr}{\sqrt{\int_0^1 V_{\underline{c}}^{\bar{c},\kappa}(r)^2\, dr}} \end{aligned}$$

which proofs the theorem. $\square$

Proof of Theorem 3

We can write the model defined by (1) and (14) in equivalent terms as:

$$x_t = (1-\alpha)\,\mu + \alpha x_{t-1} + u_t + \theta_T u_{t-1}$$

so that:

$$\begin{aligned}
x_t &= \left(1-\alpha^t\right)\mu + \alpha^t x_0 + \sum_{j=1}^{t} \exp\left(-\kappa\,(t-j)\,/T\right)\left(u_j - u_{j-1} + \delta T^{-1/2} u_{j-1}\right) \\
&= \left(1-\alpha^t\right)\mu + \alpha^t x_0 + \left(1-\delta T^{-1/2}\right)\exp\left(\kappa/T\right) u_t \\
&\quad + \left(1-\exp\left(\kappa/T\right)\left(1-\delta T^{-1/2}\right)\right)\sum_{j=1}^{t}\exp\left(-\kappa\,(t-j)\,/T\right) u_j \\
&= \mu + (x_0-\mu)\,\alpha^t + a_T u_t + b_T X_t
\end{aligned}$$

with $a_T = \left(1-\delta T^{-1/2}\right)\exp\left(\kappa/T\right)$, $b_T = \left(1-\exp\left(\kappa/T\right)\left(1-\delta T^{-1/2}\right)\right)$ – that, as can be seen, $a_T \to 1$ and $T^{1/2} b_T \to \delta$ as $T\to\infty$ – and $X_t = \sum_{j=1}^{t}\exp\left(-\kappa\,(t-j)\,/T\right)u_j$. Note that in this case in the limit:

$$T^{-1}\sum_{t=1}^{T} x_t \Rightarrow \mu + (x_0-\mu)\,\bar{\alpha} + M_u + \sigma\int_0^1 J_{\underline{c}}^{\bar{c},\kappa}(r)\,dr$$

with $\bar{\alpha} = T^{-1}\sum_{t=1}^{T}\alpha^t \in [0,1]$ for $\kappa \geq 0$. Consider the second order moment of $\hat{x}_t = x_t - \bar{x}$:

$$\begin{aligned}
T^{-1}\sum_{t=1}^{T}(x_t-\bar{x})^2 &= T^{-1}\sum_{t=1}^{T}\left((x_0-\mu)\left(\alpha^t-\bar{\alpha}\right) + a_T\,(u_t-\bar{u}) + b_T\left(X_t-\bar{X}\right)\right)^2 \\
&= T^{-1}\sum_{t=1}^{T}\left((x_0-\mu)\left(\alpha^t-\bar{\alpha}\right)\right)^2 + a_T^2 T^{-1}\sum_{t=1}^{T}(u_t-\bar{u})^2 \\
&\quad + T b_T^2 T^{-2}\sum_{t=1}^{T}\left(X_t-\bar{X}\right)^2 \\
&\quad + 2\,(x_0-\mu)\,a_T T^{-1}\sum_{t=1}^{T}\left(\alpha^t-\bar{\alpha}\right)(u_t-\bar{u}) \\
&\quad + 2\,(x_0-\mu)\,T^{1/2} b_T T^{-3/2}\sum_{t=1}^{T}\left(\alpha^t-\bar{\alpha}\right)\left(X_t-\bar{X}\right) \\
&\quad + 2 a_T T^{1/2} b_T T^{-3/2}\sum_{t=1}^{T}\left(X_t-\bar{X}\right)(u_t-\bar{u})
\end{aligned}$$

where $T^{-1}\sum_{t=1}^{T}\left(\alpha^t-\bar{\alpha}\right)^2 \to 0$, since $\alpha\to 1$ and $\bar{\alpha}\to 1$ as $T\to\infty$, $a_T^2 T^{-1}\sum_{t=1}^{T}(u_t-\bar{u})^2 \xrightarrow{p} \sigma_u^2$, and $T^{-2}\sum_{t=1}^{T}\left(X_t-\bar{X}\right)^2 \Rightarrow \sigma^2\int_0^1 V_{\underline{c}}^{\bar{c},\kappa}(r)^2\,dr$. Further, note that $\sum_{t=1}^{T}\left(\alpha^t-\bar{\alpha}\right)(u_t-\bar{u})$

$= O_p\left(T^{1/2}\right)$ and $\sum_{t=1}^{T}\left(\alpha^t-\bar{\alpha}\right)\left(X_t-\bar{X}\right)=O_p\left(T\right)$ and $\sum_{t=1}^{T}\left(X_t-\bar{X}\right)\left(u_t-\bar{u}\right)=O_p\left(T\right)$. Therefore, in the limit we have:

$$T^{-1}\sum_{t=1}^{T}\left(x_t-\bar{x}\right)^2 \Rightarrow \sigma_u^2+\delta^2\sigma^2\int_0^1 V_{\underline{c}}^{\bar{c},\kappa}\left(r\right)^2 dr$$

Now consider the cross-product $A=T^{-1}\sum_{t=1}^{T}\left(x_{t-1}-\bar{x}_{-1}\right)\left(\left(u_t-\bar{u}\right)-\left(u_{t-1}-\bar{u}_{-1}\right)\right.$ $\left.+\,\delta T^{-1/2}\left(u_{t-1}-\bar{u}_{-1}\right)\right)$:

$$\begin{aligned} A &= T^{-1}\sum_{t=1}^{T}\Big(\left(x_0-\mu\right)\left(\alpha^{t-1}-\bar{\alpha}\right)+a_T\left(u_{t-1}-\bar{u}_{-1}\right)+b_T\left(X_{t-1}-\bar{X}_{-1}\right) \\ &\quad \left(\left(u_t-\bar{u}\right)-\left(u_{t-1}-\bar{u}_{-1}\right)+\delta T^{-1/2}\left(u_{t-1}-\bar{u}_{-1}\right)\right)\Big) \end{aligned}$$

that is,

$$\begin{aligned} A &= \left(x_0-\mu\right)T^{-1}\sum_{t=1}^{T}\left(\alpha^{t-1}-\bar{\alpha}\right)\left(u_t-\bar{u}\right) \\ &\quad -\left(x_0-\mu\right)\left(1-\delta T^{-1/2}\right)T^{-1}\sum_{t=1}^{T}\left(\alpha^{t-1}-\bar{\alpha}\right)\left(u_{t-1}-\bar{u}_{-1}\right) \\ &\quad +a_T T^{-1}\sum_{t=1}^{T}\left(u_{t-1}-\bar{u}_{-1}\right)\left(u_t-\bar{u}\right)-a_T\left(1-\delta T^{-1/2}\right)T^{-1}\sum_{t=1}^{T}\left(u_{t-1}-\bar{u}_{-1}\right)^2 \\ &\quad +b_T T^{-1}\sum_{t=1}^{T}\left(X_{t-1}-\bar{X}_{-1}\right)\left(u_t-\bar{u}\right) \\ &\quad -b_T\left(1-\delta T^{-1/2}\right)T^{-1}\sum_{t=1}^{T}\left(X_{t-1}-\bar{X}_{-1}\right)\left(u_{t-1}-\bar{u}_{-1}\right) \end{aligned}$$

Taking into account the previous elements, the first two terms on the right hand side of A are $O_p\left(T^{1/2}\right)$, whereas $T^{-1}\sum_{t=1}^{T}\left(u_{t-1}-\bar{u}_{-1}\right)\left(u_t-\bar{u}\right)\xrightarrow{p}\gamma_{u,1}$, $T^{-1}\sum_{t=1}^{T}\left(u_{t-1}-\bar{u}_{-1}\right)^2\xrightarrow{p}\sigma_u^2$, $\sum_{t=1}^{T}\left(X_{t-1}-\bar{X}_{-1}\right)\left(u_t-\bar{u}\right)=O_p\left(T\right)$, $T^{-1}\sum_{t=1}^{T}\left(X_{t-1}-\bar{X}_{-1}\right)\left(u_{t-1}-\bar{u}_{-1}\right)=O_p\left(T\right)$. Since $a_T\to 1$ and $b_T\to 0$ as $T\to\infty$, we have in the limit:

$$T^{-1}\sum_{t=1}^{T}\left(x_{t-1}-\bar{x}_{-1}\right)\left(\left(u_t-\bar{u}\right)-\left(u_{t-1}-\bar{u}_{-1}\right)+\delta T^{-1/2}\left(u_{t-1}-\bar{u}_{-1}\right)\right)\xrightarrow{p}\gamma_{u,1}-\sigma_u^2$$

Since the OLS estimator is given by:

$$\hat{\alpha}=\alpha+\frac{\sum_{t=1}^{T}\left(x_{t-1}-\bar{x}_{-1}\right)\left(\left(u_t-\bar{u}\right)-\left(u_{t-1}-\bar{u}_{-1}\right)+\delta T^{-1/2}\left(u_{t-1}-\bar{u}_{-1}\right)\right)}{\sum_{t=1}^{T}\left(x_{t-1}-\bar{x}\right)^2}$$

we have that in the limit:

$$\hat{\alpha} \Rightarrow \frac{\gamma_{u,1} + \delta^2 \sigma^2 \int_0^1 V_{\underline{c}}^{\bar{c},\kappa}(r)^2 \, dr}{\sigma_u^2 + \delta^2 \sigma^2 \int_0^1 V_{\underline{c}}^{\bar{c},\kappa}(r)^2 \, dr}$$

provided that $\alpha = 1$, an expression that would be equivalent to the one obtained in Theorem 1 of Nabeya and Perron (1994) if the stochastic process were to be unbounded, with $\mu = 0$ and $\varepsilon_t \sim iid\,(0, \sigma_\varepsilon^2)$. However, whereas in Nabeya and Perron (1994) $\hat{\alpha} \xrightarrow{p} 0$ when $\delta = 0$, here we have $\hat{\alpha} \xrightarrow{p} \gamma_{u,1}/\sigma_u^2 \neq 0$. Besides, $\hat{\alpha} \xrightarrow{p} 1$ when $\delta \to \infty$ regardless of whether the stochastic process is bounded. In this case, this result coincides with the one in Nabeya and Perron (1994). □

References

[1] Andrews, D. W. K. (1993): Exactly Median-Unbiased Estimation of First Order Autoregressive /Unit Root Models. *Econometrica*, Vol. 61, pp. 139-165.

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

[3] Balli, H. O., Murray, Ch. J. and Papell, D. H. (2014): Median-unbiased estimation of structural change models: an application to real exchange rate persistence. *Journal of Applied Economics*, Vol. 46, pp. 3300-3311.

[4] Cavaliere G. (2005): Limited time series with a unit root. *Econometric Theory* 21, pp. 907-945.

[5] Cavaliere, G. and Xu, F. (2014): Testing for Unit Roots in Bounded Nonstationary Time Series. *Journal of Econometrics*, Vol. 178, pp. 259-272.

[6] Chang, Y. and Park, G.Y. (2002): On the Asymptotics of ADF Tests for Unit Roots. *Econometric Reviews*, Vol. 21, pp. 431-447.

[7] Fuller, W. A. (1996): *Introduction to Statistical Time Series*, 2nd ed. Wiley, New York, NY.

[8] Hansen, B. E. (1999): Bootstrapping the Autoregressive Model. *Review of Economics and Statistics*, Vol. 81, pp. 594–607.

[9] Herwartz, H. and Xu, F. (2008): Reviewing the Sustainability/Stationarity of Current Account Imbalances with Tests for Bounded Integration. *The Manchester School*, Vol. 76, pp. 267-278.

[10] Kejriwal, M. and Perron, P. (2010): A Sequential Procedure to Determine the Number of Breaks in Trend with an Integrated or Stationary Noise Component. *Journal of Time Series Analysis*, Vol. 31, pp. 305-328.

[11] Kilian, L. (1998): Small-Sample Confidence Intervals for Impulse Response Functions. *Review of Economics and Statistics*, Vol. 80, pp. 218-230.

[12] Nabeya, S. and Perron P. (1994): Local Asymptotic Distribution Related to the AR(1) Model with Dependent Errors. *Journal of Econometrics*, Vol. 62, pp. 229-264.

[13] Nelson, C. R. and Plosser, C. I. (1982): Trends and Random Walks in Macroeconomic Time Series. *Journal of Monetary Economics*, Vol. 10, pp. 139-162.

[14] Ng S. and Perron P. (2001): Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power. *Econometrica*, Vol. 69, pp. 1519-1554.

[15] Perron P. and Qu Z. (2007): A Simple Modification to Improve the Finite Sample Properties of Ng and Perron's Unit Root Tests. *Economics Letters*, Vol. 94, pp. 12-19.

[16] Perron, P. and Yabu, T. (2009a): Estimating Deterministic Trends with an Integrated or Stationary Noise Component. *Journal of Econometrics, Vol.* 151, pp. 56-69.

[17] Perron, P. and Yabu, T. (2009b): Testing for Shifts in Trend with an Integrated or Stationary Noise Component. *Journal of Business & Economics Statistics*, Vol. 27, pp. 369-396.

[18] Perron, P. and Yabu, T. (2012): Testing for Trend in the Presence of Autoregressive Error: A Comment. *Journal of the American Statistical Association*, Vol. 107, pp. 844-844.

[19] Pesavento, E. and Rossi, B. (2006): Small-sample Confidence Intervals for Multivariate Impulse Response Functions at Long Horizons. *Journal of Applied Econometrics*, Vol. 21, pp. 1135-1155.

[20] Phillips, P. C. B. and Perron, P. (1988): Testing for a Unit Root in Time Series Regression. *Biometrika*, Vol. 75, pp. 335-346.

[21] Roy, A. and Fuller, W. A. (2001): Estimation for Autoregressive Time Series with a Root Near 1. *Journal of Business & Economics Statistics*, Vol. 19, pp. 482-493.

[22] Roy, A., Falk, B. and Fuller, W. A. (2004): Testing for Trend in the Presence of Autoregressive Errors. *Journal of the American Statistical Association*, Vol. 99, pp. 1082-1091.

[23] Watson, M. W. (2014): Inflation Persistence, the NAIRU, and the Great Recession. *American Economic Review*, Vol. 104, pp. 31-36.

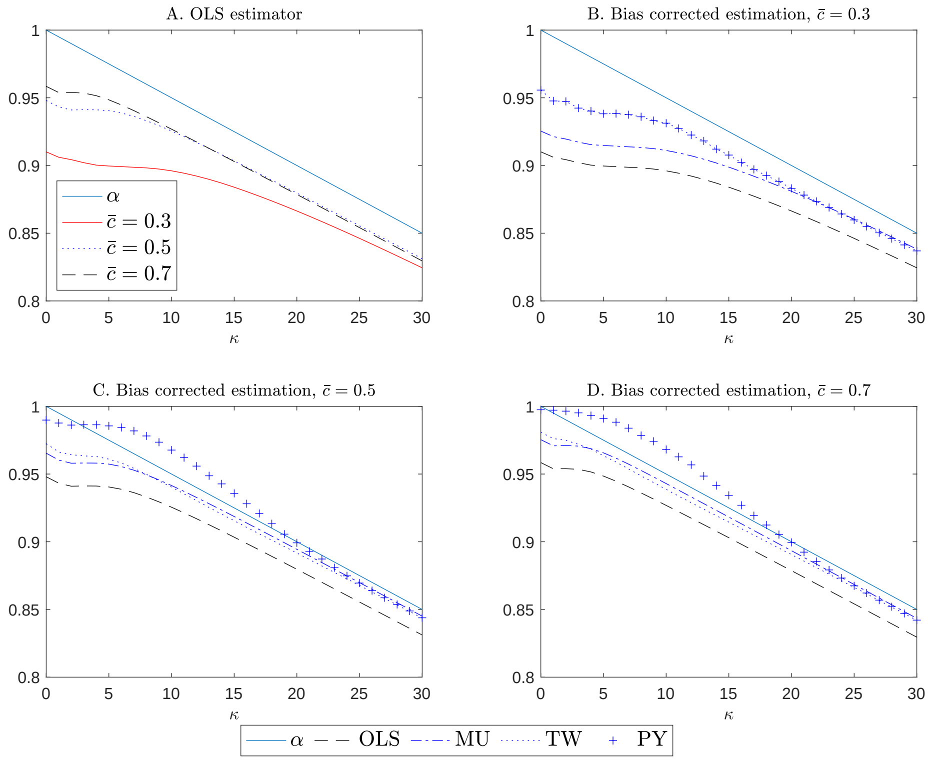
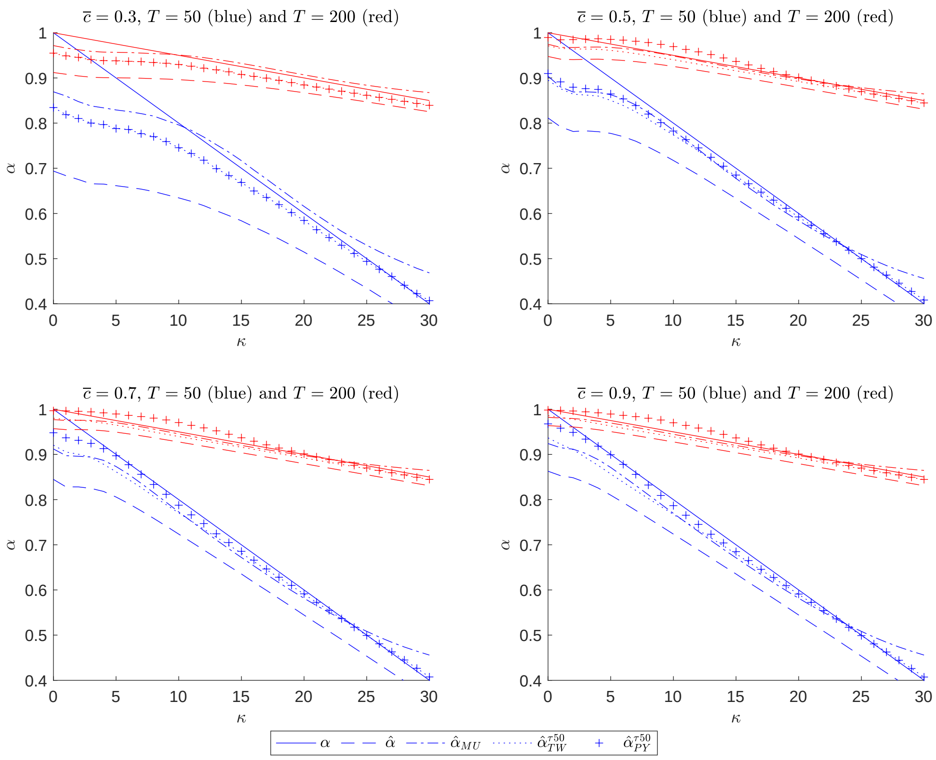


Figure 1. OLS and standard bias corrected autoregressive parameter estimates for unattended bounded processes

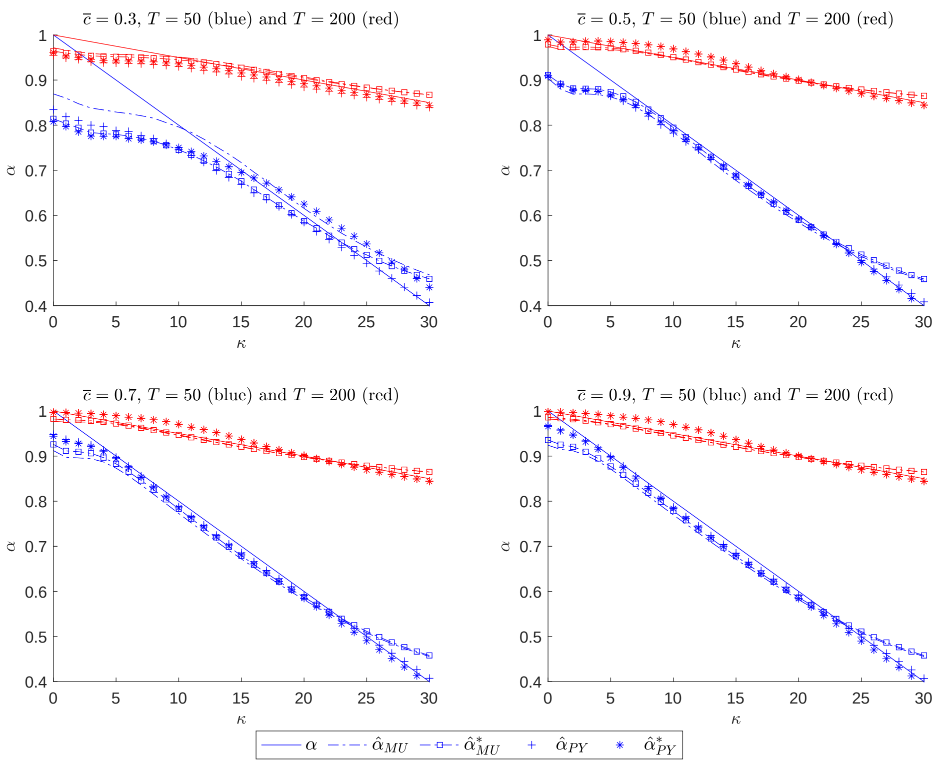
TABLE 1

Percentiles of the limiting distribution of $\hat{\tau}_W$ for different (symmetric) bounds

$(\underline{c}, \overline{c})$	*1%*	*2.5%*	*5%*	*7%*	*7.5%*	*10%*	*15%*	*50%*	*85%*
$(-0.1, 0.1)$	-9.16	-9.01	-8.88	-8.82	-8.80	-8.74	-8.64	-8.25	-7.89
$(-0.2, 0.2)$	-5.39	-5.18	-5.02	-4.94	-4.93	-4.86	-4.76	-4.38	-4.07
$(-0.3, 0.3)$	-4.58	-4.21	-3.94	-3.82	-3.79	-3.70	-3.56	-3.11	-2.80
$(-0.4, 0.4)$	-4.17	-3.85	-3.58	-3.44	-3.41	-3.28	-3.09	-2.52	-2.17
$(-0.5, 0.5)$	-3.75	-3.49	-3.27	-3.15	-3.13	-3.02	-2.85	-2.22	-1.79
$(-0.6, 0.6)$	-3.37	-3.14	-2.95	-2.86	-2.84	-2.74	-2.60	-2.04	-1.56
$(-0.7, 0.7)$	-3.15	-2.89	-2.70	-2.61	-2.59	-2.50	-2.38	-1.89	-1.42
$(-0.8, 0.8)$	-3.11	-2.81	-2.56	-2.45	-2.43	-2.33	-2.20	-1.74	-1.32
$(-0.9, 0.9)$	-3.10	-2.79	-2.54	-2.40	-2.38	-2.26	-2.08	-1.59	-1.20
$(-1.0, 1.0)$	-3.14	-2.81	-2.55	-2.40	-2.38	-2.25	-2.06	-1.46	-1.08
$(-1.5, 1.5)$	-3.13	-2.81	-2.53	-2.39	-2.36	-2.23	-2.03	-1.20	-0.49
$(-\infty, \infty)$	-3.12	-2.80	-2.53	-2.39	-2.37	-2.24	-2.04	-1.21	-0.24

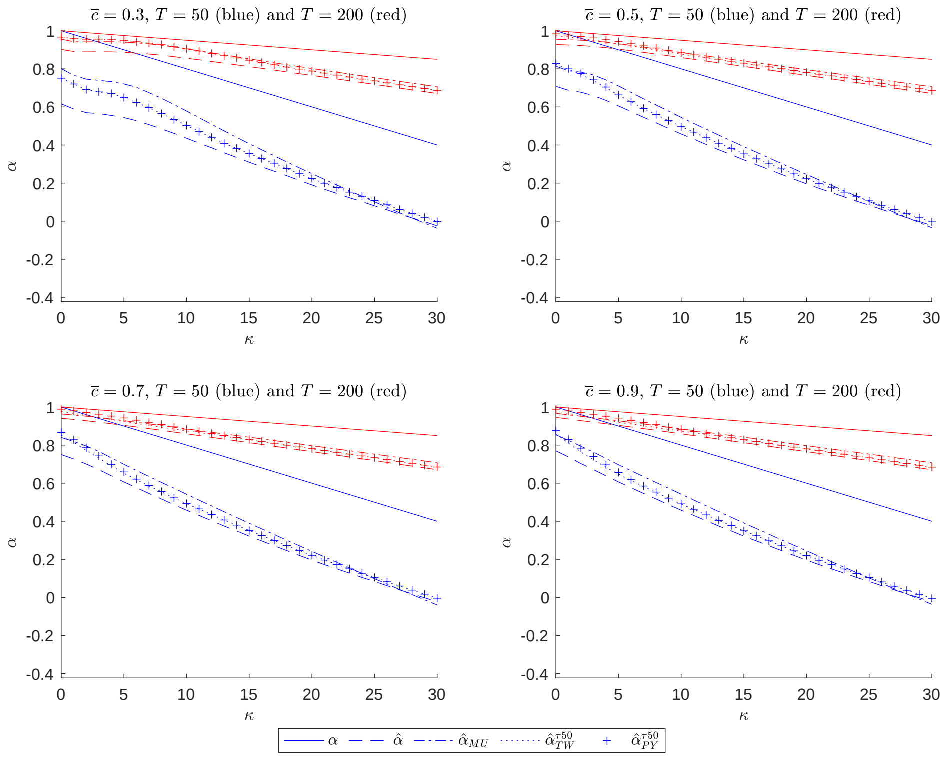


(a) Known $[\underline{c}, \bar{c}]$

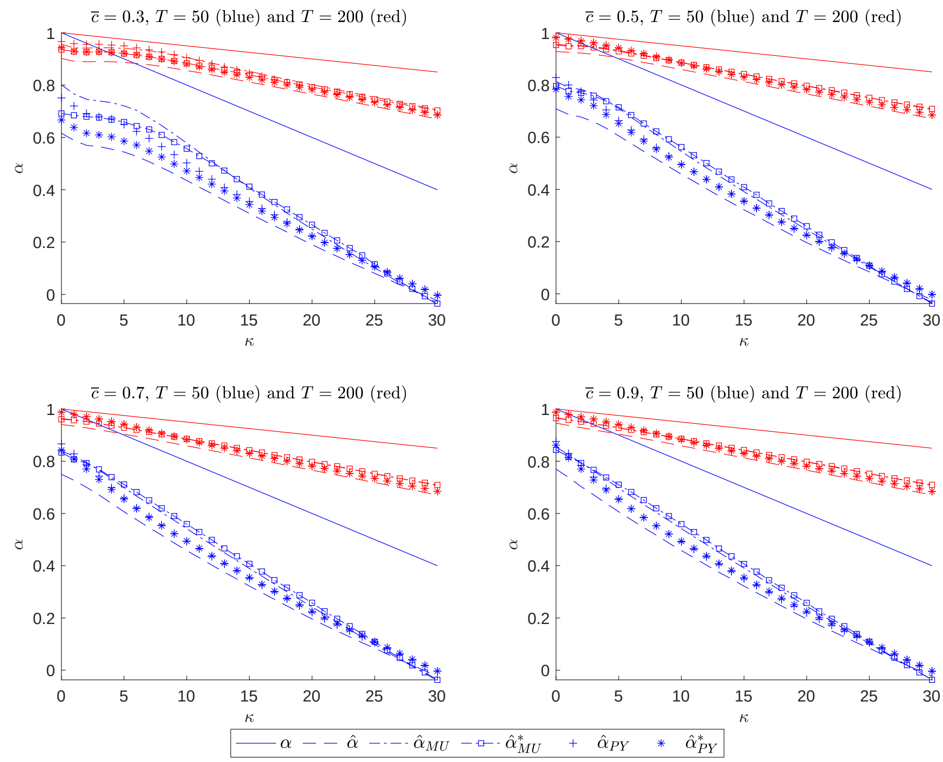


(b) Iterative approach

Figure 2. OLS and bias corrected autoregressive parameter estimates for bounded processes. AR(1) case

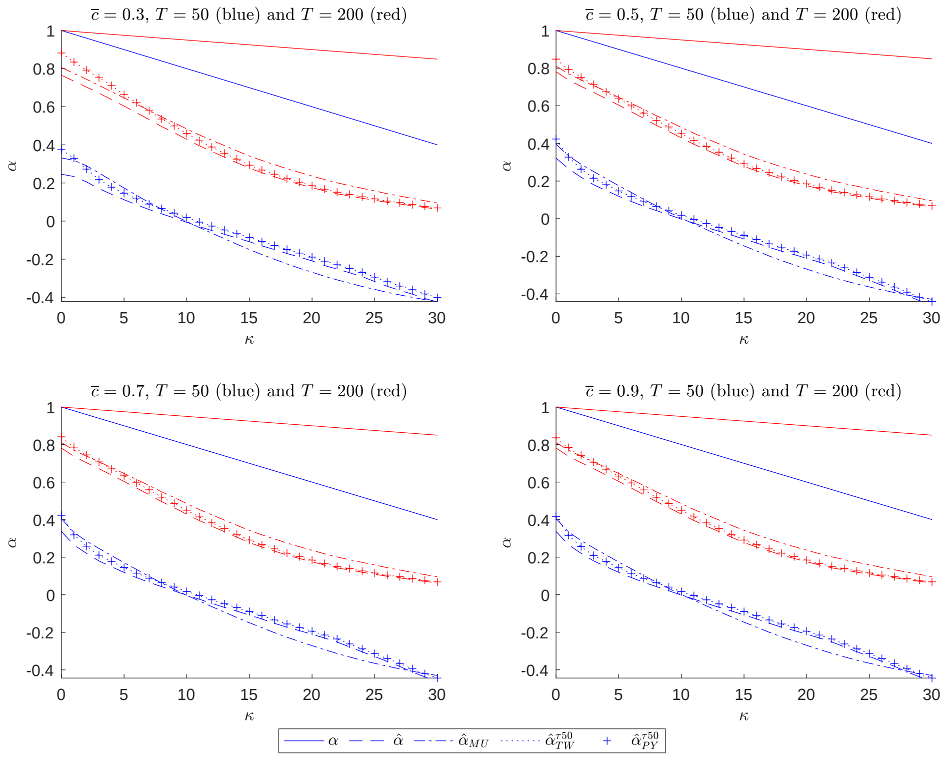


(a) Known $[\underline{c}, \overline{c}]$

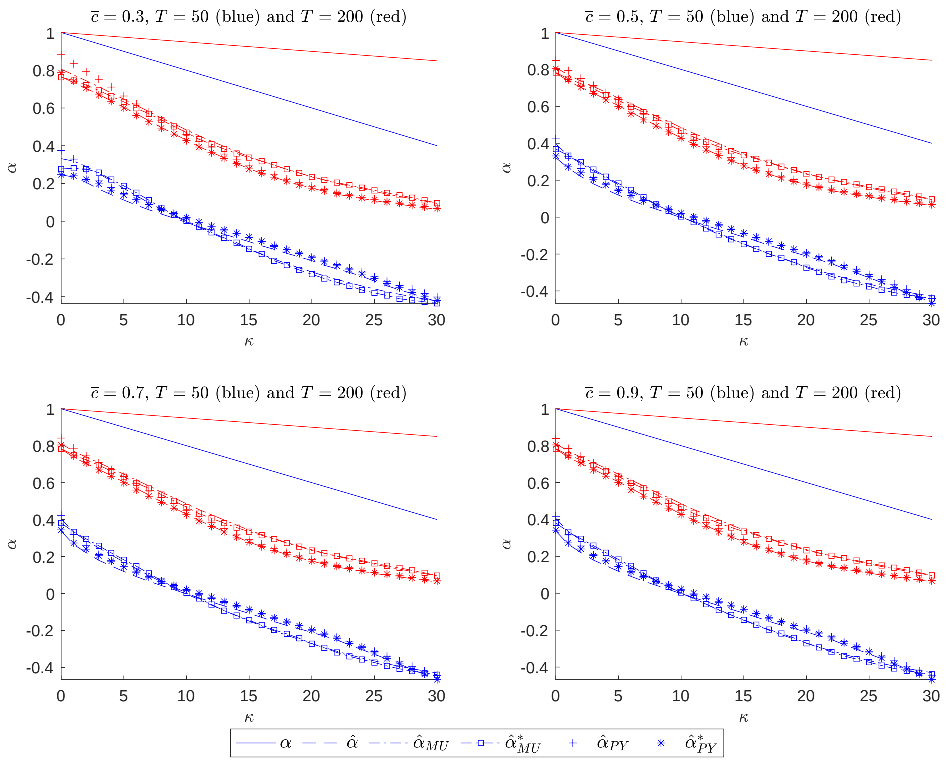


(b) Iterative approach

Figure 3. OLS and bias corrected autoregressive parameter estimates for bounded processes. ARMA(1,1) case with $\theta = -0.4$



(a) Known $[\underline{c}, \overline{c}]$



(b) Iterative approach

Figure 4. OLS and bias corrected autoregressive parameter estimates for bounded processes. ARMA(1,1) case with $\theta = -0.8$

TABLE 2

Unemployment rate persistence for the G7 countries

	Bounds ignored				Bounds considered					
		$[-\infty, \infty]$			$[0, 100]$		$[\underline{b}^{\min}, \overline{b}^{\max}]$		$[\underline{b}^*, \overline{b}^*]$	
	$\hat{k}$	$\hat{\alpha}$	$\hat{\alpha}_{MU}$	$\hat{\alpha}_{PY}$	$\hat{\alpha}_{MU}$	$\hat{\alpha}^*_{PY}$	$\hat{\alpha}_{MU}$	$\hat{\alpha}^*_{PY}$	$\hat{\alpha}_{MU}$	$\hat{\alpha}^*_{PY}$
Canada	2	0.84	0.89	0.89	0.92	1	0.95	1	0.92	1
France	2	0.82	0.86	0.85	0.94	1	1	1	0.94	1
Germany	2	0.92	1	0.95	1	1	1	1	1	1
Italy	2	0.83	0.86	0.86	0.94	1	0.97	1	0.94	1
Japan	2	0.96	1	0.97	1	1	1	1	1	1
UK	2	0.87	0.91	0.93	0.99	1	1	1	0.99	1
USA	2	0.71	0.74	0.75	0.76	0.75	0.77	0.75	0.76	0.75

Note: The computation of $\hat{\alpha}_{PY}$ is based on the τ_{50} percentile. $[\underline{b}, \overline{b}] = [\underline{b}^{\min}, \overline{b}^{\max}]$ denote the bounds defined by the minimum and maximum of the observed values of the $unem_{i,t}$ time series. $[\underline{b}, \overline{b}] = [\underline{b}^*, \overline{b}^*]$ denote the "break-even" bounds.

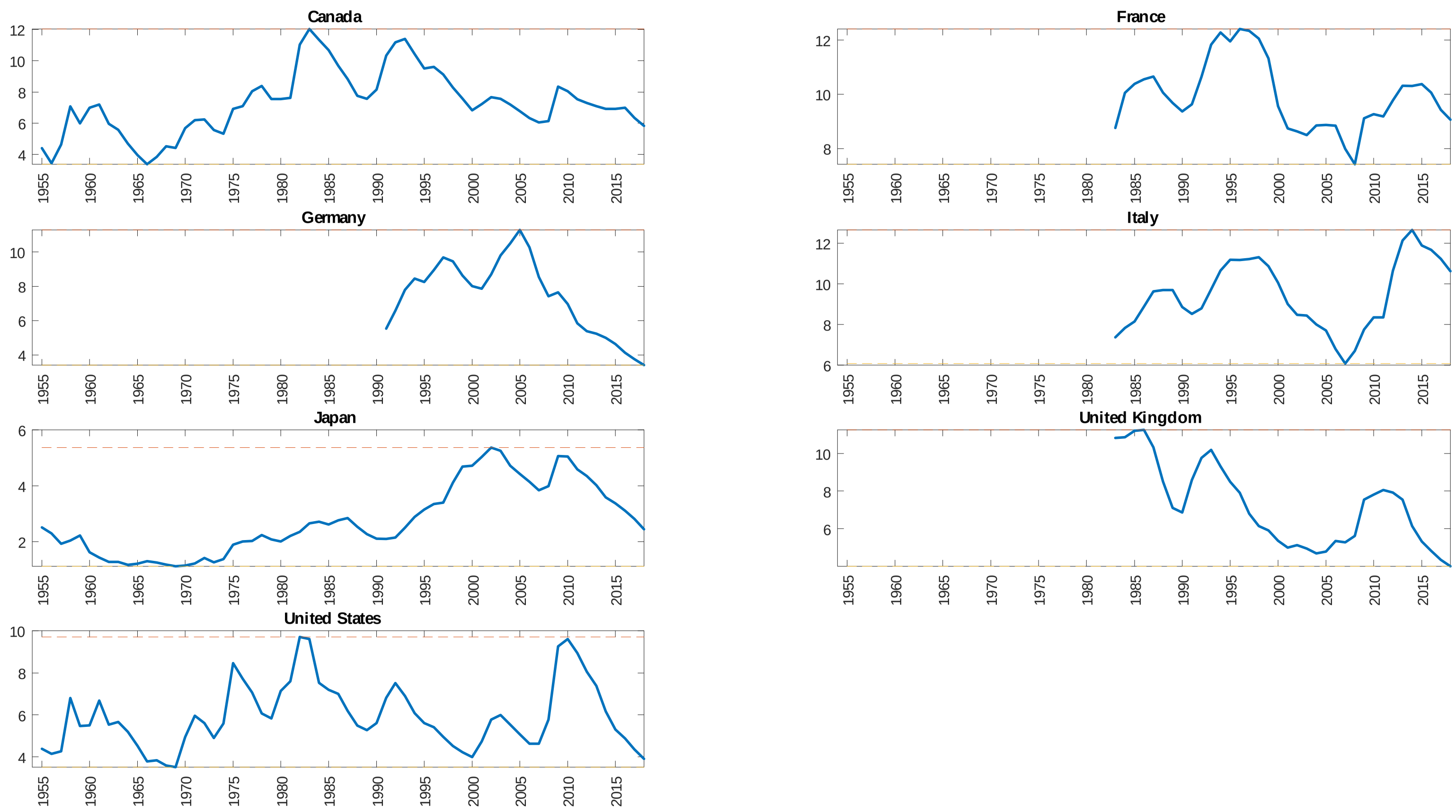


Figure 5. Unemployment rate for the G7 countries

Appendix B: Supplementary material. Additional Monte Carlo simulations results

This section provides simulation results for the different estimation bias correction procedures that consider the bounded nature of time series. Instead of focusing on the near-integrated area, these results allow us to analyze the performance of the generalized methods using a discrete set of values for the autoregressive parameter given by $\alpha = \{0, 0.1, 0.2, ..., 1\}$ for the AR(1) case. The rest of the simulation set-up is described in the paper.

TABLE B.1

Bias corrected estimator using median-unbiased estimator. The AR(1) case

		Andrews $\hat{\alpha}_{MU}$								Andrews and Chen $\hat{\alpha}_{MU}$, k known								Andrews and Chen $\hat{\alpha}_{MU}$, $k(BIC)$							
T	$\alpha\backslash\bar{c}$	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
50	0	0.07	0.06	0.07	0.06	0.06	0.06	0.06	0.06	0.08	0.07	0.07	0.07	0.07	0.08	0.07	0.07	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
	0.1	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
	0.2	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.22	0.22	0.22	0.22	0.22	0.22	0.22	0.22	0.23	0.22	0.23	0.22	0.23	0.23	0.22	0.22
	0.3	0.31	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.32	0.31	0.31	0.31	0.31	0.32	0.31	0.31	0.32	0.32	0.32	0.32	0.32	0.32	0.31	0.32
	0.4	0.41	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.42	0.41	0.41	0.41	0.41	0.41	0.41	0.41	0.43	0.41	0.41	0.41	0.41	0.42	0.41	0.41
	0.5	0.51	0.50	0.49	0.49	0.50	0.50	0.49	0.49	0.54	0.52	0.51	0.51	0.51	0.52	0.51	0.51	0.54	0.52	0.51	0.51	0.51	0.52	0.51	0.51
	0.6	0.61	0.60	0.59	0.59	0.59	0.59	0.59	0.59	0.66	0.63	0.62	0.62	0.61	0.62	0.62	0.62	0.66	0.63	0.62	0.62	0.61	0.62	0.62	0.62
	0.7	0.72	0.71	0.69	0.69	0.69	0.69	0.69	0.69	0.78	0.75	0.73	0.73	0.72	0.72	0.72	0.72	0.78	0.75	0.73	0.73	0.72	0.72	0.72	0.72
	0.8	0.80	0.81	0.80	0.80	0.79	0.79	0.79	0.79	0.88	0.87	0.85	0.85	0.84	0.83	0.83	0.83	0.88	0.87	0.85	0.84	0.84	0.83	0.83	0.83
	0.85	0.82	0.84	0.85	0.85	0.84	0.84	0.84	0.84	0.91	0.92	0.91	0.90	0.89	0.89	0.89	0.89	0.91	0.91	0.91	0.90	0.89	0.89	0.89	0.88
	0.9	0.84	0.86	0.87	0.89	0.89	0.89	0.89	0.89	0.92	0.95	0.95	0.95	0.94	0.94	0.94	0.93	0.92	0.95	0.95	0.94	0.94	0.94	0.93	0.93
	0.95	0.86	0.87	0.88	0.90	0.91	0.92	0.92	0.92	0.93	0.96	0.97	0.98	0.97	0.97	0.97	0.97	0.93	0.96	0.97	0.97	0.97	0.97	0.97	0.97
	1	0.89	0.91	0.91	0.92	0.92	0.93	0.94	0.94	0.94	0.97	0.99	0.99	0.99	0.99	0.99	0.99	0.94	0.97	0.98	0.99	0.99	0.99	0.99	0.99
200	0	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06
	0.1	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13
	0.2	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.22	0.22	0.21	0.22	0.22	0.21	0.22	0.21
	0.3	0.30	0.30	0.30	0.30	0.30	0.30	0.29	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31
	0.4	0.40	0.40	0.40	0.40	0.40	0.40	0.39	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.41	0.41	0.41	0.41	0.41	0.41	0.41	0.41
	0.5	0.50	0.50	0.50	0.50	0.50	0.50	0.49	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.51	0.51	0.51	0.51	0.51	0.50	0.51	0.50
	0.6	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60
	0.7	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70
	0.8	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.81	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.81	0.80	0.80	0.80	0.80	0.80	0.80	0.80
	0.85	0.85	0.85	0.84	0.84	0.84	0.84	0.84	0.84	0.86	0.86	0.86	0.85	0.85	0.85	0.85	0.85	0.86	0.86	0.86	0.85	0.85	0.85	0.85	0.85
	0.9	0.91	0.90	0.90	0.90	0.90	0.89	0.90	0.90	0.93	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.93	0.91	0.91	0.91	0.91	0.91	0.91	0.91
	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.98	0.97	0.97	0.97	0.96	0.96	0.96	0.96	0.98	0.97	0.97	0.97	0.96	0.96	0.96	0.96
	1	0.97	0.98	0.98	0.98	0.98	0.98	0.99	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00

TABLE B.2

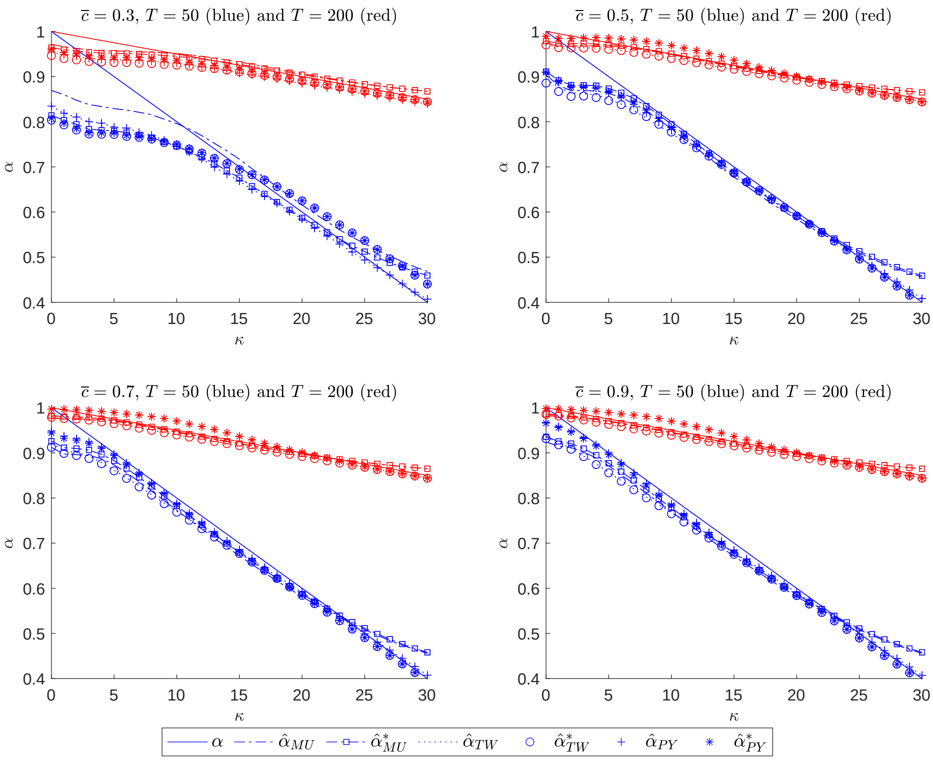
Mean of the distribution of $\hat{\alpha}_{TW}^{\tau_{50}}$. The AR(1) case

		k known								k (BIC)							
T	$\alpha\backslash\bar{c}$	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
50	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.04	0.04	0.04	0.03	0.03	0.03	0.03
	0.1	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.14	0.13	0.13	0.13	0.13	0.13	0.13	0.13
	0.2	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.23	0.22	0.22	0.22	0.22	0.22	0.22	0.22
	0.3	0.29	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.32	0.31	0.31	0.31	0.31	0.31	0.31	0.31
	0.4	0.38	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.41	0.41	0.40	0.40	0.40	0.40	0.40
	0.5	0.48	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.50	0.50	0.50	0.50	0.50	0.50	0.50
	0.6	0.56	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.58	0.59	0.59	0.59	0.59	0.59	0.59	0.59
	0.7	0.66	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.67	0.68	0.68	0.68	0.68	0.68	0.68	0.68
	0.8	0.74	0.77	0.77	0.77	0.77	0.77	0.77	0.77	0.75	0.77	0.77	0.77	0.77	0.77	0.77	0.77
	0.85	0.77	0.81	0.82	0.82	0.82	0.82	0.81	0.81	0.77	0.81	0.82	0.82	0.82	0.81	0.81	0.81
	0.9	0.78	0.83	0.85	0.86	0.87	0.86	0.86	0.86	0.79	0.83	0.85	0.86	0.86	0.86	0.86	0.85
	0.95	0.80	0.84	0.87	0.89	0.90	0.90	0.90	0.90	0.81	0.84	0.87	0.89	0.90	0.90	0.90	0.90
	1	0.83	0.87	0.89	0.91	0.92	0.93	0.94	0.94	0.83	0.87	0.89	0.91	0.92	0.93	0.94	0.94
200	0	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
	0.1	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14
	0.2	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.23	0.23	0.23	0.23	0.23	0.23	0.23	0.23
	0.3	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32
	0.4	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42
	0.5	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51
	0.6	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60
	0.7	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70
	0.8	0.79	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.79	0.80	0.80	0.80	0.80	0.80	0.80	0.80
	0.85	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84
	0.9	0.88	0.89	0.89	0.89	0.89	0.89	0.89	0.89	0.88	0.89	0.89	0.89	0.89	0.89	0.89	0.89
	0.95	0.93	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.93	0.94	0.94	0.94	0.94	0.94	0.94	0.94
	1	0.95	0.97	0.97	0.98	0.98	0.98	0.99	0.99	0.95	0.97	0.97	0.98	0.98	0.98	0.99	0.99

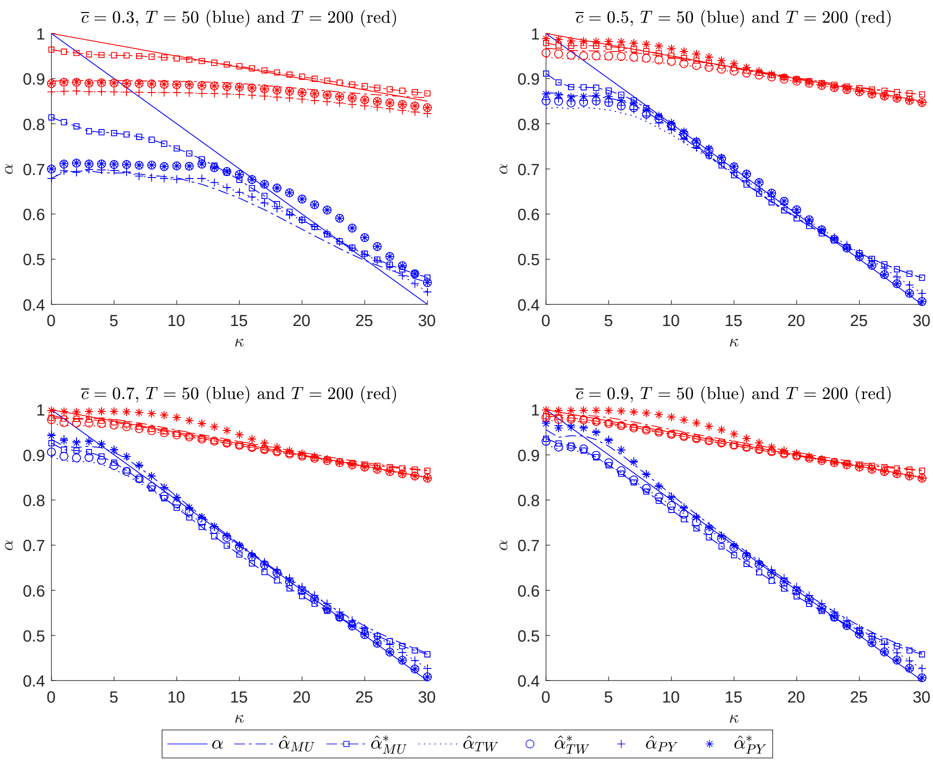
TABLE B.3

Mean of the distribution of $\hat{\alpha}_{PY}^{\tau_{50}}$. The AR(1) case

		k known								k (BIC)							
T	$\alpha \backslash \bar{c}$	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
50	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.04	0.04	0.04	0.03	0.03	0.03	0.03
	0.1	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.14	0.13	0.13	0.13	0.13	0.13	0.13	0.13
	0.2	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.23	0.22	0.22	0.22	0.22	0.22	0.22	0.22
	0.3	0.29	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.32	0.31	0.31	0.31	0.31	0.31	0.31	0.31
	0.4	0.38	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.41	0.41	0.40	0.40	0.40	0.40	0.40
	0.5	0.48	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.50	0.50	0.50	0.50	0.50	0.50	0.50
	0.6	0.56	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.58	0.59	0.59	0.59	0.59	0.59	0.59	0.59
	0.7	0.66	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.67	0.68	0.68	0.68	0.68	0.68	0.68	0.68
	0.8	0.74	0.77	0.78	0.79	0.79	0.79	0.79	0.79	0.75	0.77	0.78	0.79	0.79	0.79	0.79	0.79
	0.85	0.77	0.81	0.83	0.85	0.85	0.85	0.85	0.85	0.77	0.81	0.83	0.84	0.85	0.85	0.84	0.84
	0.9	0.78	0.83	0.87	0.89	0.90	0.90	0.90	0.90	0.79	0.83	0.87	0.89	0.90	0.90	0.90	0.90
	0.95	0.80	0.84	0.88	0.91	0.93	0.94	0.94	0.94	0.81	0.84	0.88	0.91	0.93	0.94	0.94	0.94
	1	0.83	0.87	0.91	0.93	0.95	0.96	0.97	0.97	0.83	0.87	0.91	0.93	0.95	0.96	0.97	0.97
200	0	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
	0.1	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14
	0.2	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.23	0.23	0.23	0.23	0.23	0.23	0.23	0.23
	0.3	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32
	0.4	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42
	0.5	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51
	0.6	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60
	0.7	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70
	0.8	0.79	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.79	0.80	0.80	0.80	0.80	0.80	0.80	0.80
	0.85	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84
	0.9	0.88	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.88	0.90	0.90	0.90	0.90	0.90	0.90	0.90
	0.95	0.93	0.96	0.97	0.97	0.97	0.97	0.97	0.97	0.93	0.96	0.97	0.97	0.97	0.97	0.97	0.97
	1	0.95	0.98	0.99	0.99	1.00	1.00	1.00	1.00	0.95	0.98	0.99	0.99	1.00	1.00	1.00	1.00

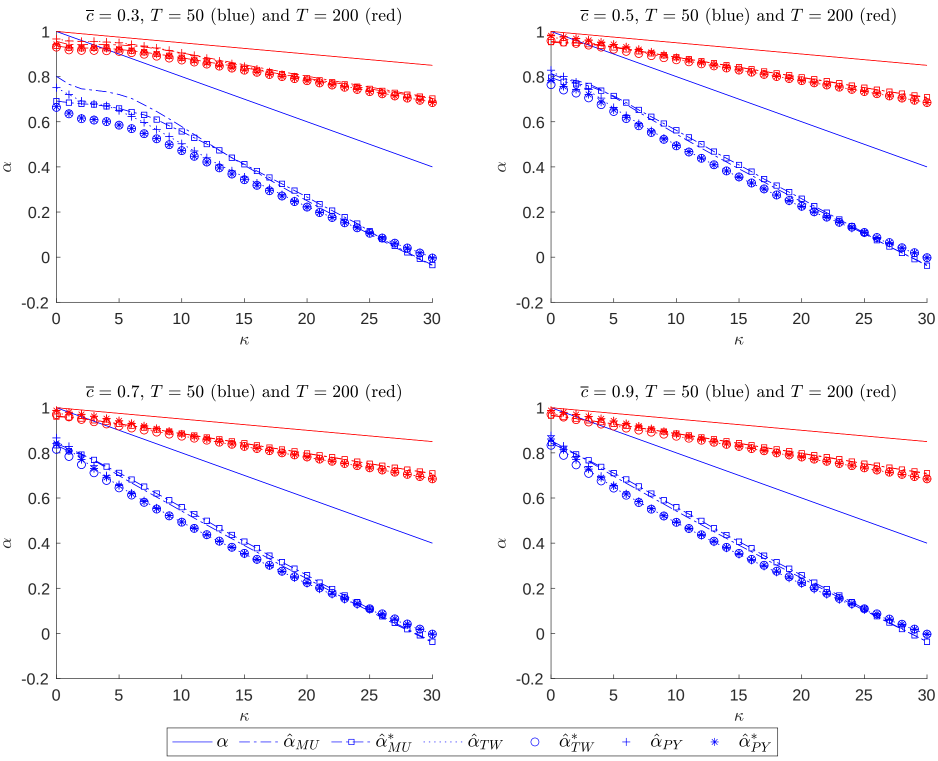


(a) AR(1), known $[\underline{c}, \bar{c}]$

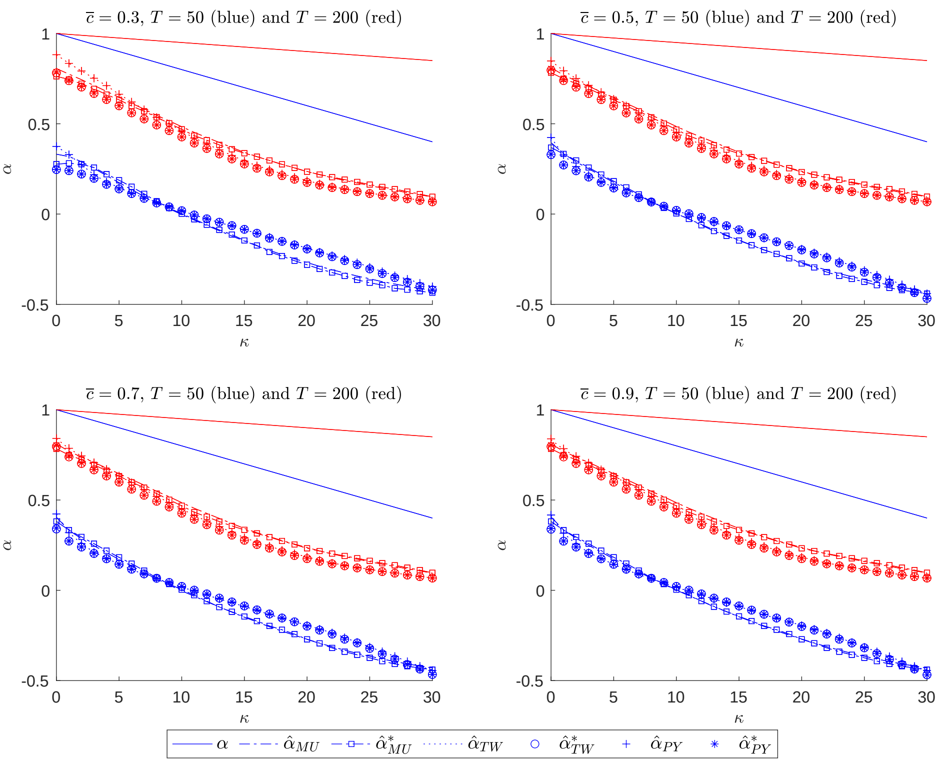


(b) AR(2), known $[\underline{c}, \bar{c}]$

Figure B.1. OLS and bias corrected autoregressive parameter estimates for bounded processes. Iterative estimation for the AR(1) and AR(2) cases



(a) ARMA(1,1), $\theta = -0.4$



(b) ARMA(1,1), $\theta = -0.8$

Figure B.2. OLS and bias corrected autoregressive parameter estimates for bounded processes. Iterative estimation for the ARMA(1,1) case