The Impact of Integrated Measurement Errors on Modelling Long-run Macroeconomic Time Series

James A. Duffy and David F. Hendry*
Economics Department and Institute for New Economic Thinking
at the Oxford Martin School, University of Oxford, UK

Abstract

Data spanning long time periods, such as that over 1860–2012 for the UK, seem likely to have substantial errors of measurement that may even be integrated of order one, but which are probably cointegrated for cognate variables. We analyze and simulate the impacts of such measurement errors on parameter estimates and tests in a bivariate cointegrated system with trends and location shifts which reflect the many major turbulent events that have occurred historically. When trends or shifts therein are large, cointegration analysis is not much affected by such measurement errors, leading to conventional stationary attenuation biases dependent on the measurement-error variance, unlike the outcome when there are no offsetting shifts or trends.

JEL classifications: C51, C22.
KEYWORDS: Integrated Measurement Errors; Location Shifts; Long-run Data; Cointegration.

1 Introduction

Data spanning long time periods, such as the period 1860–2012 analyzed for UK real wages in Castle and Hendry (2014), and for many macroeconomic time series in Hendry (2015), are bound to be inaccurately measured. The magnitudes of mis-measurements are unknown, but if large, would attenuate, and could potentially camouflage, empirical links between variables. However, the logs of many real and most nominal macroeconomic variables seem integrated of at least first order, denoted $\mathsf{I}(1)$, and their (unlogged) levels have usually changed by orders of magnitude, often exploding exponentially: nominal wages in the UK have risen by almost 70,000% since 1860. There is also considerable evidence of abrupt shifts in many macroeconomic time series due to the turbulence of the past 150 years that witnessed two world wars and other major events like oil crises and the ensuing inflation.

It seems plausible that errors of measurement would on average be roughly proportional to the levels of time series, or else data would become increasingly accurate over time, whereas the flaws recently highlighted in UK data do not suggest a high level of accuracy even in 2014. Of course, measurement error variances may well be larger during turbulent periods or when new technologies emerge, and smaller when greater resources are devoted to data measurement. To investigate a 'worst case scenario' of the impact of measurement errors on parameter estimates and cointegration tests, we treat measurement errors as $\mathsf{I}(1)$ for log levels, but to preclude indefinite 'drift' between measures of cognate variables (such as real wages and productivity), we assume the measurement errors are at least 'nearly' cointegrated. One

*Financial support from the Robertson Foundation (award 9907422), Institute for New Economic Thinking (grant 20029822) and Statistics Norway through Research Council of Norway Grant 236935, are all gratefully acknowledged, as are helpful comments from Jennifer L. Castle, Jurgen A. Doornik, Neil R. Ericsson, Andrew B. Martinez, John N.J. Muellbauer, Felix Pretis and three anonymous referees. The simulations and graphs used *PcNaive*: see Doornik and Hendry (2013).

possible 'model' for $\mathsf{I}(1)$, but cointegrated, measurement errors is that changes in the logs of macroeconomic variables (i.e., growth rates) are estimated with random errors by statistical agencies, so cumulate for log levels.

Denote the measured (log) value by y_t and the correct but unknown log value by y_t^* where lower-case denotes logs and capitals the original series, then:

$$\Delta y_t = \Delta y_t^* + v_t \text{ where } v_t \sim \mathsf{ID}[0, \sigma^2] \tag{1.1}$$

so $y_t = y_t^* + \sum_{s=0}^{t} v_{t-s}$. However, measured values are likely to be subject to periodic rebasing and are constrained by National Accounts identities, which thereby induces cointegrated measurement errors.

We provide two empirical examples of macroeconomic data affected by large measurement errors in Section 2.1, one pertaining to the late 19th century and the other to the 21st, to emphasise that the problem has not greatly lessened over time. Both examples arise from major data revisions that reveal the presence of measurement error in at least one of the vintages. Denoting the initial measured variable by $y_{1,t}$ and its revised value by $y_{2,t}$, then:

$$y_{1,t} = y_t^* + u_{1,t} \text{ and } y_{2,t} = y_t^* + u_{2,t}$$

so that the magnitude of the revision is:

$$y_{2,t} - y_{1,t} = u_{2,t} - u_{1,t} \tag{1.2}$$

Thus, if the revision differential in (1.2) is $\mathsf{I}(1)$, then at least one of the measurements must have had an $\mathsf{I}(1)$ error. Hendry (1995, Ch.12) analysed the impacts of revisions to $\mathsf{I}(1)$ data, where the original and revised series do not cointegrate, an outcome that would arise from white-noise revisions to log changes when y_t^* in (1.1) is $\mathsf{I}(1)$. He presented several empirical examples of major US and UK revisions consistent with that interpretation, proposed an encompassing test of which vintage was the more accurate, and examined the properties of log transforms of linearly aggregated data. However, the impacts on estimator distributions facing $\mathsf{I}(1)$ measurement errors and location shifts to growth rates were not derived, so are the focus here.

A notable characteristic of long-term macroeconomic data, when expressed in growth rates, is the presence of large location shifts that reflect the many major turbulent events that have occurred historically (world wars, oil crises, pandemics etc.), and lead to broken trends in log levels as seen in panels (a), (b) and (c) of Figure 2 below. A location shift at time T_0 in a growth rate is a change of the form:

$$\Delta y_t = \theta + \lambda \mathbf{1}_{\{T_0 \leq t \leq T_1\}}$$

where $\mathbf{1}_{\{T_0 \leq t \leq T_1\}}$ is an indicator equal to zero except for unity over the period shown, and $\lambda \neq 0$. (Throughout this paper, the term 'location shift' refers to a shift in the mean growth rate of a variable.) Several such shifts are illustrated by the time series for UK inflation in Figure 2(d). If a shift did not revert to zero at a later date (denoted T_1), that would correspond to a permanent change in the growth rate from θ to $\theta + \lambda$ – such has certainly occurred historically, e.g., in productivity growth. A single shift in a time series of $T = 100$ observations is used below as a baseline, since if that sufficed to 'reveal' the underlying process, then more shifts and larger shifts would clearly do so as well. Most such shifts can be identified with historically turbulent events, so are a feature of the true underlying series, and increase the ratio of 'signal' to 'noise' in the observed series. Below, we show by both analysis and simulation that when there are strong trends, and/or shifts therein are sufficiently large, then cointegration analysis is not much affected by (near) cointegrated $\mathsf{I}(1)$ measurement errors. However, conventional $\mathsf{I}(0)$ attenuation biases remain, with a magnitude dependent on measurement-error variances.

The structure of the paper is as follows. As motivation, Section 2 illustrates both serious measurement errors (from comparing revisions) and strong, shifting trends in some UK macroeconomic data over the long historical period spanning 1860–2012 (the UK is one of only a few economies that has continuous data over 150 years). Based on empirical models of real wages and productivity for that period, Section 3 describes the formulation of the cointegrated data generating process (DGP) we use for Monte Carlo simulations and asymptotic analyses to illustrate the baseline outcomes when there are no measurement errors. Section 4 considers nearly cointegrated $\mathsf{I}(1)$ measurement errors for trending data, and compares those findings with outcomes when the data are also subject to location shifts (in growth rates). Section 5 then allows the measurement errors themselves to be fully cointegrated, as seems likely with cognate macroeconomic data. Section 6 concludes. The Appendix provides mathematical derivations.

2 Measurement errors, trends and shifts in UK macroeconomic data

This section motivates the analysis of measurement errors in the presence of trends and breaks, and also provides a guide to an appropriate DGP for the simulation studies. Thus, the theory and associated simulations have a real-world grounding and emphasize the advantages of using the largest possible sample even when subsamples within the data set appear disparate, as large shifts can be helpful as well as detrimental when not modeled (as in Hendry and Mizon, 2011).

Section 2.1 looks at some examples of past and recent measurement errors revealed by large revisions, indicating that at least one vintage must be affected by mis-measurement. Then Section 2.2 illustrates some of the strong trends and and large shifts present in historical UK macroeconomic data.

2.1 Measurement errors in macroeconomic data: two examples

Some indication of the possible magnitudes of measurement errors in older macroeconomic data is provided by UK unemployment data from 1870–1913. This series was substantively revised by Boyer and Hatton (2002), who used archival information to improve upon the previous measure based on Trades Unions statistics, as analyzed in e.g., Castle and Hendry (2014). Table 1 records the means and standard deviations of the old (denoted $U_{r,O}$) and new ($U_{r,N}$) unemployment rates, and the revisions defined as the difference between the two series, $R_r = U_{r,N} - U_{r,O}$, as percentages.

Table 1: Means and standard deviations of unemployment measures and revisions (%)

	$U_{r,O}$	$U_{r,N}$	R_r
Mean	4.17	5.37	1.20
Std. dev.	2.25	1.94	1.25

Panels (a) and (b) of Figure 1 display $U_{r,O}$ and $U_{r,N}$ and the revisions. The new unemployment series is less volatile, but has a higher mean (by more than 1%). Although not integrated, the revisions (viewed as proxies for the previous unknown measurement errors) have an autocorrelation of 0.55, and their standard deviation is more than half that of the revised series, with which the revisions are uncorrelated.

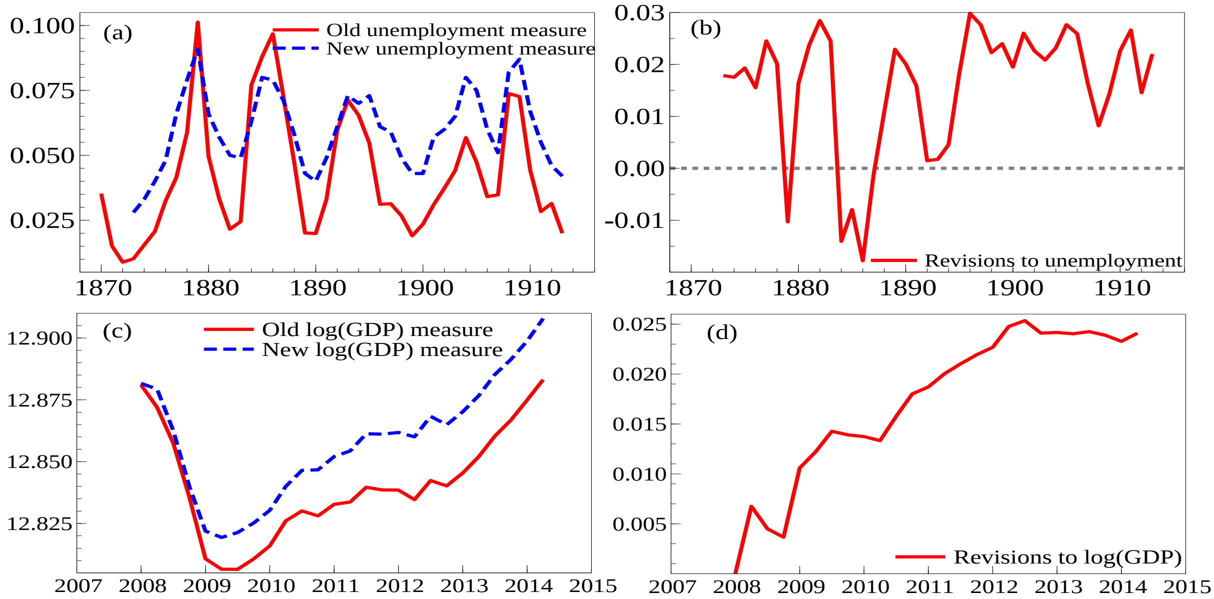
Table 2: Unemployment coefficients and standard errors

Variable	$\widehat{\beta}_{U_{r,O}}$	$\mathsf{SE}_{\widehat{\beta}_{U_{r,O}}}$	$\widehat{\beta}_{U_{r,N}}$	$\mathsf{SE}_{\widehat{\beta}_{U_{r,N}}}$
$(U_{r,t} - 0.05)$	-0.17	0.03	-0.20	0.04
$(U_{r,t} - 0.05)^2$	2.68	0.68	2.92	0.66
$\Delta_2 U_{r,t}$	-0.13	0.05	-0.19	0.05
$\widehat{\sigma}; \mathsf{R}^2$	1.04	0.82	1.01	0.83

Nevertheless, as shown in Davidson, Hendry, Srba, and Yeo (1978), measurement errors need to be large to greatly distort estimates in 'informative' time series. Using the revised measure, we re-estimated the final real-wage model in Castle and Hendry (2014), in which unemployment is a key determinant. As seen in Table 2, the use of the revised data induces only relatively small changes in the estimated coefficients and standard errors of their three functions of unemployment, and in their measures of goodness of fit. However, all the estimates on the old data were closer to the origin than on the revised, with a larger error standard deviation, $\widehat{\sigma}$, consistent with the former being mis-measured.

Next, early 2016 revisions to UK log(GDP) are shown in Figure 1. Panel (c) plots the latest vintages of the original and revised time series over 2008Q1 to 2014Q1, with the revisons shown in panel (d).

Figure 1: (a) Old and new measures for UK unemployment from 1870–1913, and (c) log GDP over 2008Q1 to 2014Q1, with their revisions in (b) and (d).



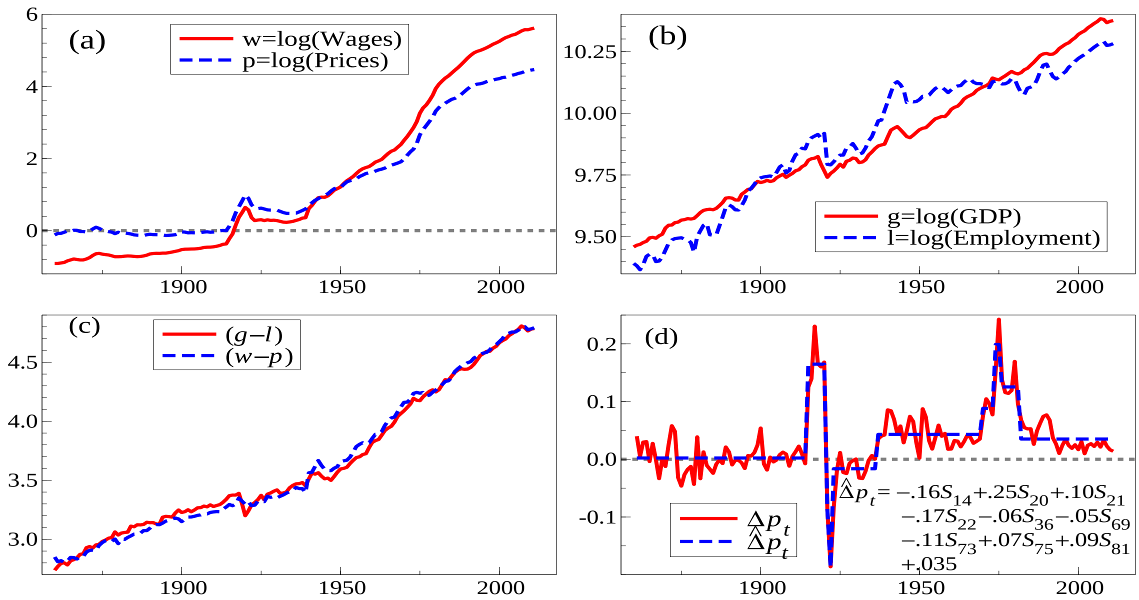
The largest revisions exceed the average annual growth rate of UK GDP. Using an augmented Dickey and Fuller (1981) test on the revisions with up to two lagged differences (neither significant), constant and trend, a unit root cannot be rejected. Thus, the possibilty of I(1) revisions cannot be excluded even for recent data, albeit only over a short time period, so from (1.2), one of the series must have I(1) measurement errors.

2.2 Trends and shifts in UK macroeconomic data

Figure 2 illustrates UK macroeconomic time series over 1860–2012 for the logs of annual average weekly earnings, w, and prices, p, in panel (a); for logs of annual GDP, g, and employment, l, in (b) (matched to the same means and ranges for viewing); real wages, $w - p$, and productivity, $y - l$, in (c); and inflation Δp in (d), showing the location shifts selected by step-indicator saturation (SIS) at 0.1% significance, where e.g., S_{xx} denotes a step indicator that is unity till $19xx$ and zero thereafter, to highlight some of the major events: see Castle, Doornik, Hendry, and Pretis (2015).

The log-levels of all these series, in both nominal and constant prices, exhibit strong but changing trends, with first-differences that are also nonstationary due to location shifts. Inflation has a drift of 3.5% per annum and nine location shifts, the largest of which is 25%, compared to a residual standard deviation of 2.3%. Such features are apparent in many long-run UK time series.

Figure 2: Trends and breaks in UK time series



2.3 Empirical real-wage and productivity relationships

This section demonstrates that the trends and shifts manifest in the data illustrated in Figure 2 persist in empirical macro relationships. The analysis is motivated by the range of single-equation macroeconomic relationships estimated for the UK over 1860–2012 reported in Hendry (2015), re-estimated here using data updated to 2014. The equations show the prevalence of drift and of shifts in growth rates that are external to the variables in the model. We consider a simplified version of a 2-equation real-wage and productivity system for the data shown in Figure 2(c). The economic variables were retained without selection in levels representations, and only the step shifts selected by SIS at 0.25%, then transformed to equilibrium-correction model (EqCM) or differenced formulations and any insignificant variables eliminated (here, and below, the figures in parentheses are estimated coefficient standard errors):

$$\begin{aligned} \Delta\widehat{(w-p)}_t \;=\;& \underset{(0.06)}{0.45}\,\Delta(g-l)_t - \underset{(0.04)}{0.17}\,(w-p-g+l-\widehat{\mu})_{t-1} - \underset{(0.02)}{0.13}\,S_{39} \\ & + \underset{(0.02)}{0.14}\,S_{40} - \underset{(0.02)}{0.06}\,S_{41} + \underset{(0.01)}{0.04}\,S_{43} + \underset{(0.002)}{0.014} \end{aligned} \tag{2.1}$$

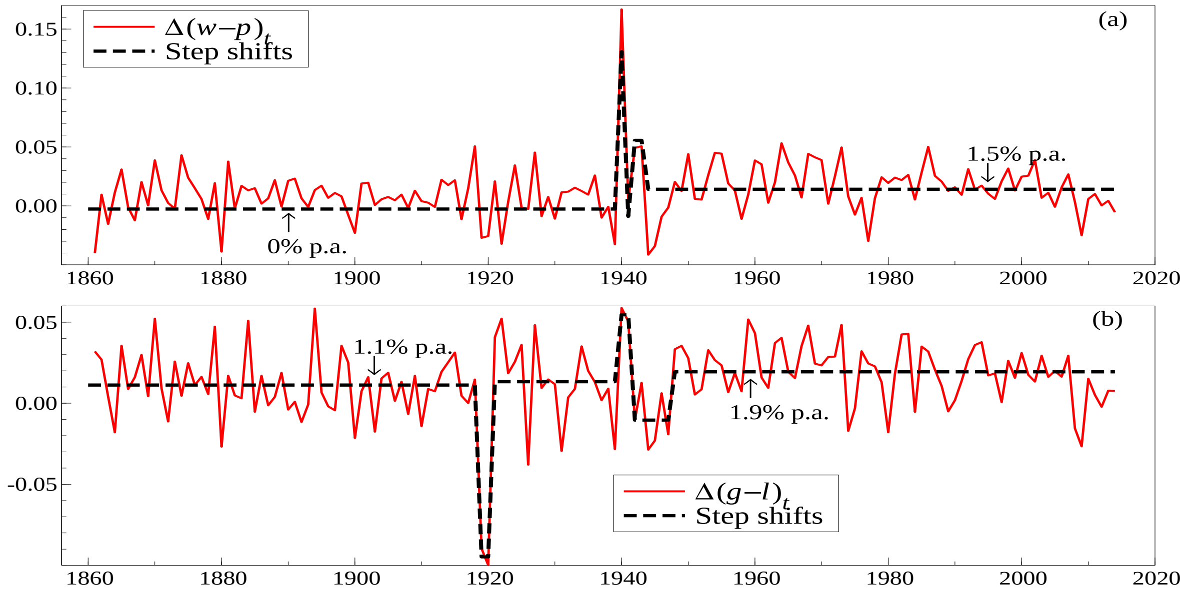
with a residual standard deviation of $\widehat{\sigma}_{wp} = 1.5\%$, which is very small compared to the location shifts of up to 14% ($\widehat{\mu} = 1.85$ is the sample mean of $(w-p-g+l)$). Compared to the 9 large location shifts in Figure 2(d), there is considerable co-breaking (see Hendry and Massmann, 2007) in real wages. Next:

$$\Delta\widehat{(g-l)}_t = \underset{(0.002)}{0.02} + \underset{(0.01)}{0.11}\,S_{18} - \underset{(0.01)}{0.11}\,S_{20} - \underset{(0.01)}{0.04}\,S_{39} + \underset{(0.02)}{0.07}\,S_{41} - \underset{(0.01)}{0.03}\,S_{47} \tag{2.2}$$

The residual standard deviation of $\widehat{\sigma}_{gl} = 1.9\%$ is slightly larger than the average annual growth rate of 1.7%, but much smaller than the location shifts of more than 10%.

Figure 3 shows the explanation provided by the selected location shifts for the models of $\Delta(w-p)_t$ (panel (a)) and $\Delta(g-l)_t$ (panel (b)). For $\Delta(w-p)_t$, although the average annual growth rate was approximately 0.6% before World War II, this is explained by the coefficient of 0.45 on productivity

Figure 3: (a) Changes in UK log real wages and (b) changes in UK log GDP, both over 1860–2014 with selected location shifts.



improvements and the falling level of $(w - p - g + l - \widehat{\mu})_{t-1}$, so no contribution is required from the selected step indicators. However, $\Delta(w-p)_t$ grows at about 1.7% p.a. after 1945, consistent with the dramatic increases in living standards in the post-war period, but most is now explained by the step shift as the economic regressors offset each other. For productivity growth, the post-war jump is a smaller magnitude (about 0.8% p.a.), but occurs at about the same time. The crash after World War I is very large, and importantly has little effect on $\Delta(w-p)_t$, validating the contemporaneous conditioning in (2.1). These location shifts in first-differences translate to trend shifts in the log-levels, which are clearly visible in Figure 2 for all the time series in panels (a)–(c).

3 A cointegrated data generation process for analyses and simulations

Given the above evidence of possibly integrated measurement errors, the earlier results on I(1) data revisions, and the empirical models in the previous section, we consider a DGP of two I(1) cointegrated series, denoted y_t and z_t. y_t is determined by a first-order autoregressive-distributed lag equation (denoted AD(1,1)), and z_t is a unit-root process with a drift θ and a location shift λ therein as in (2.2):

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 z_t + \beta_3 z_{t-1} + \epsilon_t \quad \text{where} \quad \epsilon_t \sim \mathsf{IN}\left[0, \sigma_\epsilon^2\right] \tag{3.1}$$

$$z_t = \theta + z_{t-1} + \lambda 1_{\{T_1 \le t \le T_2\}} + e_t \quad \text{where} \quad e_t \sim \mathsf{IN}\left[0, \sigma_e^2\right] \tag{3.2}$$

(3.1) can be equivalently expressed as the EqCM in (2.1):

$$\Delta y_t = \beta_2 \Delta z_t - (1 - \beta_1)(y_{t-1} - \kappa z_{t-1}) + \epsilon_t \tag{3.3}$$

where $\kappa := (\beta_2 + \beta_3) / (1 - \beta_1)$. The DGP parameter values in the simulations are $\beta_0 = 0$, $\beta_1 = 0.8$, $\beta_2 = 0.5$, $\beta_3 = -0.3$, $\sigma_\epsilon = \sigma_e = 1$, $T = 100$, $T_1 = 50$ and $T_2 = 70$; with $M = 10000$ Monte Carlo replications. The values of $\theta \in \{0, 0.25\}$ and $\lambda \in \{0, 0.1, 0.5\}$ are varied across 10 simulation designs (S0–S9), in the manner described below. Relative to the innovation variances (of unity), the trends and shifts corresponding to these parameter values are smaller than those which occurred in the UK historical time series shown above.

However, to provide an initial illustration of how location shifts can mitigate the biases that would be otherwise be introduced by measurement errors, Figure 4 presents the limiting distributions of OLS estimates of β_1 and a test for cointegration in the following scenarios:

(i) no drift and no shifts ($\theta = 0$ and $\lambda = 0$), with $\mathsf{I}(1)$ measurement errors (see Section 4 below);

(ii) as in (i), but with $\theta = 0.25$ and $\lambda = 0.5$;

(iii) as in (i), but without measurement errors.

Figure 4: Densities of $\widehat{\beta}_1$ for $\beta_1 = 0.8$ and of the cointegration tests in the three cases (i)–(iii).

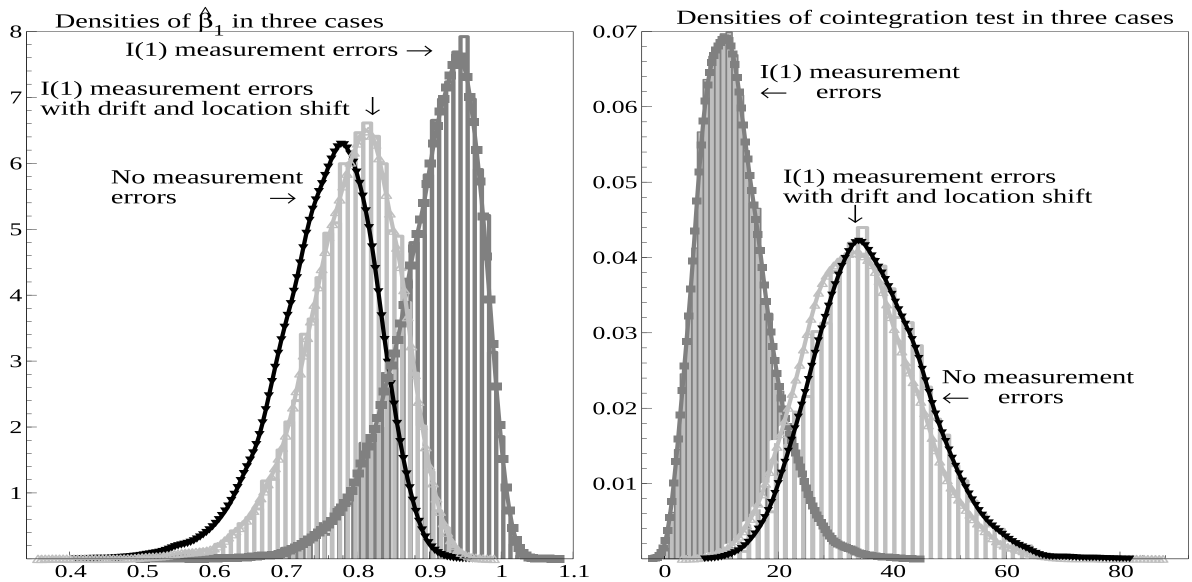


Figure 4 shows that the distributions of both statistics in (ii) are close to the corresponding (unrealistic) scenario with no measurement errors in (iii). In the remainder of the paper, we further investigate these distributions when there is either just drift, so $\theta = 0.25$ and $\lambda = 0$, or just a location shift to the growth rate, so $\theta = 0$ and $\lambda \neq 0$, first with no measurement errors, and then when these are present and are $\mathsf{I}(1)$.

3.1 Limiting distribution of the least squares estimator (OLS)

This section introduces the notation and presents baseline analytical results when there are no measurement errors, which will be useful in the later analyses. Section 4 adds $\mathsf{I}(1)$ measurement errors.

3.1.1 Drift with no location shifts

Let $\widehat{\boldsymbol{\beta}}' := (\widehat{\beta}_1, \widehat{\beta}_2, \widehat{\beta}_3)$ denote the OLS estimates of the slope coefficients in (3.1), when a constant term is also included in the regression. If z_t contains only a drift, which in (3.2) corresponds to the case where $\theta \neq 0$ and $\lambda = 0$, then $\widehat{\boldsymbol{\beta}}$ is asymptotically normal, but its limiting variance matrix is only of rank 2, owing to the faster rate at which the signal accumulates in the direction of the cointegrating vector. This can be seen most clearly by rewriting (3.1) as:

$$y_t = \beta_0 + \beta_1 \mu_{t-1} + \beta_2 \Delta z_t + \kappa z_{t-1} + \epsilon_t, \tag{3.4}$$

where $\mu_t := y_t - \kappa z_t$ is the equilibrium error, which follows:

$$\mu_t = \beta_1 \mu_{t-1} + (\beta_2 - \kappa)\Delta z_t + \epsilon_t,$$

and we have used $\kappa = \kappa\beta_1 + \beta_2 + \beta_3$. In (3.4), $(\mu_{t-1}, \Delta z_t)$ are $\mathsf{I}(0)$, whereas z_{t-1} is $\mathsf{I}(1)$.

Proposition 3.1. *If $\theta \neq 0$ and $\lambda = 0$ in (3.2), then*

$$T^{1/2}\begin{bmatrix}\widehat{\beta}_1 - \beta_1 \\ \widehat{\beta}_2 - \beta_2\end{bmatrix} \rightsquigarrow \mathsf{N}\left[0,\ \sigma_\epsilon^2 \begin{bmatrix}\sigma_\mu^{-2} & 0 \\ 0 & \sigma_e^{-2}\end{bmatrix}\right] \tag{3.5}$$

and

$$(\widehat{\beta}_3 - \beta_3) = -(\widehat{\beta}_2 - \beta_2) - \kappa(\widehat{\beta}_1 - \beta_1) + o_p(T^{-1/2}).$$

As is noted in the proof (given in Appendix A.1, as are the proofs of all subsequent results), this asymptotic independence of $\widehat{\beta}_1$ and $\widehat{\beta}_2$ does not hold generally, and for instance would fail here if e_t were allowed to be serially correlated. (Note that this would still be consistent with z_t being strongly exogenous.) On the other hand, the reduced rank of the limiting distribution of $\widehat{\boldsymbol{\beta}}$ *is* a general consequence of the stochastic trend in z_t: see Section 2 in Watson (1994) for a further discussion. It will also be observed that the magnitude of the drift θ does not appear in (3.5): it affects only the limiting variance of the OLS estimate of κ in (3.4).

3.1.2 Location shifts

When location shifts to the drift are instead present ($\theta = 0$ and $\lambda \neq 0$), the behaviour of $\widehat{\boldsymbol{\beta}}$ is not substantively altered, unless the shifts are 'large'—which will be modelled here by allowing their size to grow as $T \to \infty$. In view of panel (d) in Figure 2, in which the displayed shifts are considerably larger than the ordinary variation in the series, this seems to be a reasonable assumption. Moreover, when measurement errors are introduced into the model in Section 4 below, this assumption will allow us to determine the *maximum* extent to which location shifts afford robustness to such errors, and will facilitate the derivation of simpler, and more readily interpretable, asymptotic bias formulae.

For the purpose of deriving the large-sample approximations given in Propositions 3.2 and 5.1 below, we shall accordingly modify (3.2) to:

$$\Delta z_t := \theta_t + e_t \quad \text{where} \quad \theta_t := T^{\tau}\sum_{i=1}^{K}\lambda_i \mathbf{1}\{t \geq \lfloor Tr_i \rfloor\} \tag{3.6}$$

for some $\tau \in (0, \frac{1}{2})$. In this case, μ_t and Δz_t will be asymptotically collinear, owing to the dominance of the location shift introduced into both by θ_t; it is therefore convenient to rewrite the model as:

$$y_t = \beta_1 \rho_{t-1} + \gamma\Delta z_t + \kappa z_{t-1} + \epsilon_t \tag{3.7}$$

where $\gamma := \alpha\beta_1 + \beta_2$, $\alpha := (\beta_2 - \kappa)/(1 - \beta_1)$, and $\rho_t := \mu_t - \alpha\Delta z_{t+1}$ which follows:

$$\rho_t = \beta_1 \rho_{t-1} - \alpha\Delta\theta_{t+1} - \alpha(e_{t+1} - e_t) + \epsilon_t \tag{3.8}$$

Since $\Delta\theta_t = 0$ for all but finitely many t, ρ_t behaves similarly to an $\mathsf{I}(0)$ process (in large samples). Under (3.6), *both* Δz_t and z_{t-1} will be of larger stochastic order than a stationary regressor, and so $\widehat{\boldsymbol{\beta}}$ will enjoy a faster rate of convergence along two directions, in the sense that

$$\mathbf{A}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) := \begin{bmatrix}\kappa & 1 & 1 \\ \alpha & 1 & 0\end{bmatrix}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = o_p(T^{-1/2}) \tag{3.9}$$

(see the proof of the next result). The direction $(\kappa, 1, 1)$ is associated with the signal from the (broken) stochastic trend in z_t, and the direction $(\alpha, 1, 0)$ with the signal from the location shift in Δz_t.

The limiting distribution of $\widehat{\boldsymbol{\beta}}$ remains jointly normal, but with an asymptotic covariance matrix that now is only of rank 1. For the statement of the next result, let $\overline{x}_t := x_t - T^{-1}\sum_{s=1}^{T} x_s$, and $\sigma_\rho^2 := (1-\beta_1)^{-1}(\sigma_\epsilon^2 + 2\alpha^2\sigma_e^2)$.

Proposition 3.2. *Under (3.6),*

$$T^{1/2}(\widehat{\beta}_1 - \beta_1) = \frac{T^{-1/2}\sum_{t=1}^{T}\overline{\rho}_{t-1}\overline{\epsilon}_t}{T^{-1}\sum_{t=1}^{T}\overline{\rho}_{t-1}^2} + o_p(1) \rightsquigarrow \mathsf{N}[0,\ \sigma_\epsilon^2\sigma_\rho^{-2}],$$

and

$$T^{1/2}\begin{bmatrix}\widehat{\beta}_2 - \beta_2 \\ \widehat{\beta}_3 - \beta_3\end{bmatrix} = \begin{bmatrix}-\alpha \\ \alpha - \kappa\end{bmatrix} T^{1/2}(\widehat{\beta}_1 - \beta_1) + o_p(1).$$

It will be observed that the limiting distribution of $\widehat{\beta}_1$ depends on the variability of the (essentially) $\mathsf{I}(0)$ component ρ_t, but not on the magnitudes $\{\lambda_i\}$ of the shifts, indicating that there is a definite limit to the extent to which location shifts may render the OLS estimator robust to measurement errors.

3.2 Simulating the AD(1,1) model without measurement errors

In this section, we record the small-sample behavior of the OLS estimator and associated test statistics in the absence of measurement errors, and no drift or shift (so $\theta = \lambda = 0$ in (3.1); denoted S0), then allowing for a drift of $\theta = 0.25$ but no shift (S1), and finally a location shift of magnitude $\lambda = 0.5$ when $\theta = 0$ (S2). This provides a benchmark against which to assess the performance of OLS in the presence of $\mathsf{I}(1)$ measurement errors, as will be considered subsequently.

Table 3: Monte Carlo simulations: no measurement errors

	S0		S1		S2	
	$\theta = 0.0, \lambda = 0.0$		$\theta = 0.25, \lambda = 0.0$		$\theta = 0.0, \lambda = 0.5$	
	Test rejection frequencies					
Sig. level	1%	5%	1%	5%	1%	5%
$\mathsf{t}_{\beta_1=0}$	1.00	1.00	1.00	1.00	1.00	1.00
$\mathsf{t}_{\beta_2=0}$	0.98	1.00	0.98	1.00	0.99	1.00
$\mathsf{t}_{\beta_3=0}$	0.40	0.65	0.40	0.65	0.41	0.65
$\mathsf{F}_{AR(1)}$	0.01	0.05	0.01	0.05	0.01	0.05
F_{Het}	0.02	0.05	0.02	0.06	0.02	0.05
$\mathsf{H}_{r=0}$	1.00	1.00	1.00	1.00	1.00	1.00
	OLS estimates					
	mean	MCSD	mean	MCSD	mean	MCSD
$\widehat{\beta}_1$	0.77	0.06	0.77	0.06	0.77	0.06
$\widehat{\beta}_2$	0.50	0.10	0.50	0.10	0.50	0.10
$\widehat{\beta}_3$	-0.27	0.12	-0.27	0.12	-0.27	0.11
$\mathsf{ESE}[\widehat{\beta}_1]$	0.06	0.01	0.06	0.01	0.06	0.01
$\mathsf{ESE}[\widehat{\beta}_2]$	0.10	0.01	0.10	0.01	0.10	0.01
$\mathsf{ESE}[\widehat{\beta}_3]$	0.12	0.01	0.12	0.01	0.12	0.01
$\widehat{\sigma}_\epsilon^2$	1.00	0.14	1.00	0.14	1.00	0.14
R^2	0.92	0.07	0.98	0.04	0.96	0.05

The results of these exercises are displayed in Table 3. In the upper panel, $\mathsf{t}_{\beta_1=0}$ denotes the t-test for the null hypothesis that the coefficient β_1 of y_{t-1} is zero, and so on. Thus, that is rejected 100% of the time at both 1% and (hence also) 5%, irrespective of the scenario. Similarly, $\mathsf{H}_0 : \beta_2 = 0$ and $\mathsf{H}_0 : \beta_3 = 0$ are respectively rejected in 99% and 40% of samples at a 1% significance level, with the last rising to 65% at the 5% level.

Next, $\mathsf{F}_{AR(1)}$ and F_{Het} report the rejection rates on mis-specification tests for first-order residual autocorrelation and for residual heteroskedasticity, which are roughly equal to the nominal significance. The bottom row, $\mathsf{H}_{r=0}$, refers to testing the null of no cointegration based on the EqCM 'PcGive test' (see e.g., Ericsson and MacKinnon, 2002), which correctly rejects almost 100% of the time.

The means and Monte Carlo Standard Deviations (MCSDs) of the parameter estimates are shown in the lower panel of the table, along with their Estimated Standard Errors (ESEs): see e.g., Hendry (1995) for definitions of these and the MCSDs of the ESEs. The mean estimated parameter values are close to their respective DGP values of $\beta_1 = 0.8$, $\beta_2 = 0.5$, $\beta_3 = -0.3$, so least-squares estimation of the AD(1,1) works well in this cointegrated setting. The theoretical justification for ESEs relies on manifold assumptions (including IN errors, constant parameters, accurate data, etc.), and only when all of those assumptions are satisfied will the ESE equal the corresponding MCSD, which happens here.

Finally, the respective outcomes are reported for $\widehat{\sigma}^2_\epsilon$ (where the DGP value is unity) and R^2, where the very high value reflects the data being I(1). Overall, there is little impact of drift or a location shift in $\{z_t\}$ on the distribution of any statistic when the data are accurate, a finding which accords entirely with Propositions 3.1 and 3.2.

4 The DGP with I(1) measurement errors

The DGP in (3.1)–(3.2) of two I(1) cointegrated series, y_t and z_t, is now augmented by each having an I(1) measurement error, w_t and u_t respectively:

$$\begin{aligned} w_t &= w_{t-1} + v_t \quad \text{where} \quad v_t \sim \mathsf{IN}\left[0, \sigma_v^2\right] \\ u_t &= u_{t-1} + \eta_t \quad \text{where} \quad \eta_t \sim \mathsf{IN}\left[0, \sigma_\eta^2\right] \end{aligned} \tag{4.1}$$

so the observed data are the mis-measured values:

$$y_t^* = y_t + w_t \quad \text{and} \quad z_t^* = z_t + u_t. \tag{4.2}$$

Thus the estimated model is:

$$y_t^* = \beta_0 + \beta_1 y_{t-1}^* + \beta_2 z_t^* + \beta_3 z_{t-1}^* + \xi_t \tag{4.3}$$

where, letting $\delta_t := w_t - \kappa u_t$, the induced error is:

$$\xi_t = (1 - \beta_1)\,\delta_{t-1} + v_t - \beta_2 \eta_t + \epsilon_t$$

which for the parameter values typical in the simulations below – that is, $\sigma_v = \sigma_\eta = 0.25$ – leads to:

$$\xi_t = 0.2\delta_{t-1} + v_t - 0.5\eta_t + \epsilon_t. \tag{4.4}$$

The EqCM for the observed process (cf. (3.3) above) becomes:

$$\Delta y_t^* = \beta_2 \Delta z_t^* - (1 - \beta_1)\left(y_{t-1}^* - \kappa z_{t-1}^*\right) + \xi_t \tag{4.5}$$

Differencing y_t^* and z_t^* to Δy_t^* and Δz_t^* will difference the measurement errors to Δw_t and Δu_t, which thereby become I(0). However, $(y_t^* - \kappa z_t^*)$ will remain I(1) unless the measurement errors also cointegrate with coefficient κ, as we will consider in Section 5.

The properties of $\{\xi_t\}$ depend on all the AD(1,1) parameter values and the two error variances σ_v^2 and σ_η^2. However, ξ_t does not depend on θ (nor on λ when that is non-zero), so the 'signal-noise' ratio of the information from z_t relative to the noise from ξ_t is changed by the magnitudes and lengths of location shifts, and drift, in the strongly exogenous variable. This is most easily seen when estimating the long-run equation directly:

$$y_t^* = (\beta_2 - \kappa)\theta + \kappa z_t^* + \varsigma_t \tag{4.6}$$

where:

$$\varsigma_t = \beta_1(\mu_t + \delta_t) + (\beta_2 - \kappa)(e_t + \eta_t) + \xi_t.$$

However, while this entails that the estimation of κ will be robust to the presence of measurement error, this robustness is not inherited by the estimates of the $\boldsymbol{\beta}$ coefficients in (3.1). This should already be apparent from Propositions 3.1 and 3.2 above, in particular from the fact that neither θ nor λ appear in the limiting distribution of $\widehat{\boldsymbol{\beta}}$. Nonetheless, the results below do indicate that OLS estimates of $\boldsymbol{\beta}$ still perform well when measurement errors are of a plausible magnitude.

4.1 Asymptotic analysis of OLS in the presence of drift and I(1) measurement errors

To provide some insight into the behaviour of OLS estimates of $\boldsymbol{\beta}$ in the presence of I(1) measurement errors, we have derived expressions for the first-order bias of the OLS estimator. To prevent the resultant biases from completely dominating the estimator, it is assumed that:

$$T^{1/2}(w_t - \kappa u_t) = T^{1/2}\delta_t =: \pi_t \tag{4.7}$$

has incremental variance σ_π^2, such that $T^{-1/2}\pi_{\lfloor nr \rfloor} \rightsquigarrow B_\pi(r)$ on $D[0,1]$, where B_π denotes a Brownian motion with variance σ_π^2. This may be regarded as corresponding to the case where w_t and u_t 'almost' cointegrate, a possibility that seems empirically plausible, and which is further explored in the next section. Under this parameterization, we obtain:

Proposition 4.1. *If $\theta \neq 0$ and $\lambda = 0$ in (3.2), and (4.2) and (4.7) hold, then*

$$\widehat{\beta}_1 \xrightarrow{d} \beta_1 + \frac{1-\beta_1}{1+R} \qquad\qquad \widehat{\beta}_2 \xrightarrow{p} \beta_2 \frac{\sigma_e^2}{\sigma_e^2 + \sigma_\eta^2} \tag{4.8}$$

where $R := \sigma_\mu^2/\sigma_\pi^2 \int \widetilde{W}^2$, and $\widetilde{W}$ denotes a demeaned and linearly detrended standard Brownian motion; and

$$\widehat{\beta}_3 - \beta_3 = -(\widehat{\beta}_2 - \beta_2) - \kappa(\widehat{\beta}_1 - \beta_1) + o_p(1). \tag{4.9}$$

As was noted in respect of Proposition 3.1 above, the particularly simple expressions that obtain here are partly a consequence of our assumption that the innovation e_t in z_t is i.i.d. While it is also possible to determine the asymptotic bias of the OLS estimator in the presence of location shifts of the form (3.6), the resulting expression is not easily interpretable, as it involves the inverse of a 3×3 matrix.

Proposition 4.1 indicates that measurement errors will bias estimates of $\widehat{\beta}_2$ towards zero, and so replicates the sort of attenuation bias familiar from models with stationary variables. On the other hand, $\widehat{\beta}_1$ is biased towards unity: together these imply that measurement errors will push estimates of the EqCM model (3.3) towards a pure (i.e. non-cointegrated) random walk model for y_t.

4.2 Simulating the AD(1,1) model with I(1) measurement errors

The next experiment simulates estimation and inference in the mis-measured AD(1,1) model with either a drift or a location shift, to investigate how often cointegration between y_t and z_t is still found, any biases in the estimated parameters of the AD(1,1), and whether there are signs of mis-specification such as residual autocorrelation and non-constancy in the estimated coefficients. The DGP is the same as (3.1) – there is either a drift in z_t ($\theta \neq 0$) or a location shift in Δz_t of size λ – augmented by (4.1). Four alternative simulation designs (S3–S6) are considered. S3 has a drift ($\theta = 0.25$) but no location shift, while S4–S6 have no drift ($\theta = 0$) but location shifts of varying magnitudes (0.5, 0.1 and 0.1 respectively). In each of S3–S5, the error innovation variances are fixed at 0.25 (for both σ_v and σ_η) whereas in S6, these are both increased to unity. The results are collected in Table 4.

Table 4: Monte Carlo simulations: $\mathsf{I}(1)$ measurement errors facing drift or location shifts

	S3		S4		S5		S6	
	$\theta = 0.25, \lambda = 0.0$		$\theta = 0, \lambda = 0.5$		$\theta = 0, \lambda = 0.1$		$\theta = 0, \lambda = 0.1$	
	$\sigma_v = \sigma_\eta = 0.25$		$\sigma_v = \sigma_\eta = 0.25$		$\sigma_v = \sigma_\eta = 0.25$		$\sigma_v = \sigma_\eta = 1$	
	Test rejection frequencies							
Sig. level	1%	5%	1%	5%	1%	5%	1%	5%
$\mathsf{t}_{\beta_1=0}$	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$\mathsf{t}_{\beta_2=0}$	0.97	0.99	0.97	0.99	0.96	0.99	0.38	0.63
$\mathsf{t}_{\beta_3=0}$	0.54	0.76	0.64	0.82	0.62	0.81	0.28	0.51
$\mathsf{F}_{AR(1)}$	0.01	0.05	0.01	0.05	0.01	0.06	0.01	0.05
F_{Het}	0.02	0.06	0.02	0.06	0.02	0.07	0.02	0.06
$\mathsf{H}_{r=0}$	1.00	1.00	0.98	1.00	0.97	0.99	0.32	0.58
	OLS estimates							
	mean	MCSD	mean	MCSD	mean	MCSD	mean	MCSD
$\widehat{\beta}_1$	0.84	0.07	0.86	0.07	0.86	0.07	0.95	0.05
$\widehat{\beta}_2$	0.48	0.11	0.48	0.11	0.47	0.11	0.25	0.11
$\widehat{\beta}_3$	-0.31	0.12	-0.34	0.13	-0.34	0.13	-0.22	0.11
$\mathsf{ESE}[\widehat{\beta}_1]$	0.05	0.01	0.05	0.01	0.05	0.01	0.03	0.01
$\mathsf{ESE}[\widehat{\beta}_2]$	0.10	0.01	0.10	0.01	0.11	0.01	0.11	0.01
$\mathsf{ESE}[\widehat{\beta}_3]$	0.11	0.01	0.11	0.01	0.11	0.01	0.11	0.01
$\widehat{\sigma}^2_\epsilon$	1.10	0.17	1.14	0.17	1.13	0.17	2.23	0.33
R^2	0.98	0.04	0.97	0.05	0.93	0.07	0.95	0.07

When $\sigma_v = \sigma_\eta = 0.25$ (i.e. for S3–S5), the $\mathsf{I}(1)$ measurement errors have surprisingly little impact on the outcomes. The t-tests reject about the same (or more often) as when there are no measurement errors (compare with Table 3 above). There is no sign of mis-specification, and non-cointegration is still rejected almost all the time. Similarly, the coefficient estimates in designs S3–S5 remain close to their DGP values. Some of the ESEs differ from their corresponding MCSDs, tending to underestimate the true variability of the estimators. $\widehat{\sigma}^2_\epsilon$ is estimated as being 10–14% larger than its population value of unity. Overall, the empirical investigator would not necessarily realize, nor obtain results greatly affected by, the $\mathsf{I}(1)$ measurement errors for the present parameter configuration.

On the other hand, the results from design S6 show that larger values for σ_v and σ_η, in combination with smaller values for λ, can be more detrimental. While reducing λ from 0.5 to 0.1 – i.e., moving from S4 to S5 – has little effect on the estimators and associated tests, subsequently increasing σ_v and σ_η to unity – as in design S6 – leads to large coefficient estimator biases, non-cointegration only being rejected 30% of the time (at a 1% significance level), and an increase in the mean of $\widehat{\sigma}^2_\epsilon$ to 2.2. Indeed, the $\widehat{\beta}_i$ almost correspond to a differenced representation, although increasing λ to 1 delivers $\widehat{\boldsymbol{\beta}}' = (0.90, 0.30, 0.21)$, and 100% rejection of non-cointegration.

While Proposition 4.1 indicated that $\widehat{\beta}_1$ and $\widehat{\beta}_3$ converge in distribution to random variables, the means of those limiting variates may be computed (via simulation, replacing R by its sample analogue $R_n := \sum_{t=1}^n \widetilde{\mu}_t^2 / \sum_{t=1}^n \widetilde{\delta}_t^2$), and compared to the outcomes of Monte Carlo simulations in Table 4.

For the parameter values in design S3, the means of the limiting variates given by (4.8) and (4.9) are $(0.84, 0.47, -0.31)$, which agree almost exactly with the figures displayed in Table 4. This suggests that these asymptotic bias formulae provide a good approximation to the finite-sample bias of the OLS estimators in the presence of I(1) measurement errors, even when $T = 100$.

5 Cointegrated I(1) measurement errors

It is likely that w_t and u_t will be cointegrated by construction, as would occur with say, measurement errors on aggregate consumer expenditure and disposable income. In this case, ξ_t in (4.4) will be I(0), despite the measurement errors in themselves being I(1). To illustrate this, suppose that:

$$w_t = (1 - \psi_1)w_{t-1} + \psi_1 \kappa u_{t-1} + v_t \tag{5.1}$$

so that w_t and u_t are cointegrated when $0 < \psi_1 < 1$:

$$\Delta w_t = -\psi_1 \delta_t + v_t \tag{5.2}$$

where $\delta_t = w_t - \kappa u_t$. Consequently, when $\psi_1 = 0.25$, ξ_t in (4.4) becomes I(0):

$$\xi_t = -0.8\Delta w_t + 5v_t - 0.5\eta_t. \tag{5.3}$$

5.1 Asymptotic bias of OLS

When (5.1) holds, and location shifts of the form (3.6) are present, it is possible to derive explicit formulae for the bias of the OLS estimators of $\boldsymbol{\beta}$.

Proposition 5.1. *Under (3.6), (4.2) and (5.1),*

$$\widehat{\beta}_1 - \beta_1 = \frac{\sum_{t=1}^{n} \overline{\rho}_{t-1}^{*} \overline{\xi}_t}{\sum_{t=1}^{n} (\overline{\rho}_{t-1}^{*})^2} + o_p(1) \xrightarrow{p} \frac{(1 - \beta_1 - \psi_1)\sigma_\delta^2 + \alpha\beta_2\sigma_\eta^2}{\sigma_\rho^2 + \sigma_\delta^2 + \alpha^2\sigma_\eta^2}, \tag{5.4}$$

and

$$\begin{bmatrix} \widehat{\beta}_2 - \beta_2 \\ \widehat{\beta}_3 - \beta_3 \end{bmatrix} = \begin{bmatrix} -\alpha \\ \alpha - \kappa \end{bmatrix} (\widehat{\beta}_1 - \beta_1) + o_p(1).$$

It is evident from the proof of the preceding that, analogously to (3.9) above,

$$\mathbf{A}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) := \begin{bmatrix} \kappa & 1 & 1 \\ \alpha & 1 & 0 \end{bmatrix} (\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = o_p(1), \tag{5.5}$$

indicating that $\widehat{\boldsymbol{\beta}}$ remains consistent along two directions in the presence of cointegrated measurement errors. Along the remaining direction, however, $\widehat{\boldsymbol{\beta}}$ is *not* consistent, and the resultant bias infects the OLS estimates of *all* coefficients in the EqCM.

To provide some indication of the quality of the approximations provided by these formulae, to the finite-sample bias of $\widehat{\beta}_1$, Table 5 compares the bias $\widehat{\beta}_1 - \beta_1$ as computed through Monte Carlo simulation (with 10000 replications), with that the bias given by the r.h.s. of (5.4). $\lambda \in \{0.1, 1, 5\}$ varies as indicated in the table, while $\theta = 0$, $\sigma_v = \sigma_\eta = 1$, and all other parameters are set to their baseline values. (When $T = 1000$, the dates for the location shift are proportionately adjusted to $T_1 = 500$ and $T_2 = 700$.)

The table additionally records the mean of $\mathbf{A}\widehat{\boldsymbol{\beta}}$; for the parameter values chosen here, this should be close to $\mathbf{A}\boldsymbol{\beta} = (1, -1.5)$. While the first component of this vector is always close to unity, the second component is only close to its theoretical value when λ is sufficiently large. However, even for such 'large' values of λ, there remains a substantial bias in $\widehat{\beta}_1$, consistent with our earlier remark that there is a definite limit to the extent to which location shifts may mitigate the effects of measurement errors.

Table 5: Bias in $\widehat{\beta}_1$: Cointegrated measurement errors facing a location shift

T	100	100	100	1000	1000
λ	0.1	1	5	1	5
MC bias	-0.067	-0.072	-0.071	-0.039	-0.064
Eq. (5.4)	-0.073	-0.072	-0.051	-0.072	-0.069
$(\mathbf{A}\widehat{\boldsymbol{\beta}})_1$	0.995	0.997	0.998	1.000	1.000
$(\mathbf{A}\widehat{\boldsymbol{\beta}})_2$	-1.582	-1.559	-1.482	-1.641	-1.525

5.2 Simulating the AD(1,1) model with shifts and cointegrated measurement errors

The final three simulations designs, S7–S9, correspond exactly to the designs S4–S6 (see Table 4), except that now the $\mathsf{I}(1)$ measurement errors u_t and w_t cointegrate, as per (5.1) with $\psi_1 = 0.25$.

Table 6: Monte Carlo simulations: cointegrated $\mathsf{I}(1)$ measurement errors facing location shifts

	S7		S8		S9	
	$\theta = 0, \lambda = 0.5$		$\theta = 0, \lambda = 0.1$		$\theta = 0, \lambda = 0.1$	
	$\sigma_v = \sigma_\eta = 0.25$		$\sigma_v = \sigma_\eta = 0.25$		$\sigma_v = \sigma_\eta = 1$	
	Test rejection frequencies					
Sig. level	1%	5%	1%	5%	1%	5%
$\mathsf{t}_{\beta_1=0}$	1.00	1.00	1.00	1.00	1.00	1.00
$\mathsf{t}_{\beta_2=0}$	0.97	0.99	0.97	0.99	0.41	0.66
$\mathsf{t}_{\beta_3=0}$	0.30	0.54	0.31	0.54	0.01	0.05
$\mathsf{F}_{AR(1)}$	0.01	0.05	0.01	0.05	0.01	0.05
F_{Het}	0.01	0.05	0.02	0.05	0.02	0.05
$\mathsf{H}_{r=0}$	1.00	1.00	1.00	1.00	0.99	1.00
	OLS estimates					
	mean	MCSD	mean	MCSD	mean	MCSD
$\widehat{\beta}_1$	0.77	0.06	0.77	0.06	0.75	0.06
$\widehat{\beta}_2$	0.47	0.10	0.47	0.10	0.25	0.11
$\widehat{\beta}_3$	-0.24	0.12	-0.24	0.12	0.00	0.12
$\mathsf{ESE}[\widehat{\beta}_1]$	0.06	0.01	0.06	0.01	0.06	0.01
$\mathsf{ESE}[\widehat{\beta}_2]$	0.10	0.01	0.10	0.01	0.10	0.01
$\mathsf{ESE}[\widehat{\beta}_3]$	0.12	0.01	0.12	0.01	0.12	0.01
$\widehat{\sigma}^2_\epsilon$	1.08	0.16	1.08	0.16	2.13	0.31
R^2	0.96	0.05	0.92	0.06	0.91	0.07

The results of these simulations are displayed in Table 6. $\mathsf{H}_0 : \beta_3 = 0$ is rejected much less often than with accurate data, or with non-cointegrated measurement errors, but the null of non-cointegration is almost always rejected. Matching the rejection frequencies, the mean estimate of β_3 is now much smaller, and $\widehat{\beta}_1$ is slightly downward biased, whereas the mean of $\widehat{\beta}_2$ is little altered except for large measurement errors. Consequently, the estimated model ends up being close to a cointegrated AD(1,0) even with large, but cointegrated, $\mathsf{I}(1)$ measurement errors and small location shifts. Compared to S6 in Table 4, cointegrated measurement errors move the $\widehat{\beta}_i$ away from a near differenced representation.

The apparent insensitivity of $\widehat{\beta}_1$ to measurement errors—as distinct from the estimates of the other two coefficients—seems difficult to explain *a priori*, and may be specific to the chosen parametrization. Nonetheless, it is consistent with the formulae given in Proposition 5.1: for the final scenario depicted in the table (S9), the implied probability limit of $\widehat{\beta}$ is $(0.73, 0.32, -0.04)$, which is quite close to the means of the simulations. Bigger values of λ in S9 deliver mean parameter estimates closer to their DGP values, but much less so than even S8, so values of σ_{υ} and σ_{η} as large as σ_{ϵ} seem problematic. Despite that, drift or location shifts therein allow correct rejection of the null of no cointegration most of the time.

6 Conclusions

Key characteristics of long-run macroeconomic time series are that they almost certainly have measurement errors, are often integrated with drift, and have location shifts in first differences. This paper considered the effects of measurement errors that are I(1) and nearly (or fully) cointegrated on a 2-variable cointegrated data generation process with drift or location shifts herein, and derived asymptotic biases and their simulation counterparts for OLS estimates of the parameters of the first AD(1,1) equation.

Overall, the analysis and the simulation evidence suggests that drift or location shifts in exogenous variables can help mitigate some of the adverse consequences of integrated, but nearly cointegrated, measurement errors on long-run time series. Very large variances of cointegrated measurement errors greatly increase the uncertainty (e.g., the equation residual variance doubles in S9), and lead to substantial conventional I(0) attenuation biases, with a magnitude dependent on measurement-error variances. The resulting biases in parameter estimates can drive the AD(1,1) DGP to an AD(1,0) model, although even large measurement-error variances do not camouflage cointegration given drift or shifts. However, very large variances of non-cointegrated measurement errors drive the estimates to a near differenced representation.

Many large location shifts have occurred historically on strongly trending data, and intuitively, highlight the 'genuine' connections between variables that could otherwise be masked by mis-measurements. Thus, despite long runs of data almost certainly being mis-measured, approximated here by a 'worst case scenario' of being I(1), trends and shifts therein can help sustain reasonable estimates. Near cointegration between the measurement errors is essential to avoid them dominating asymptotically, but cointegrated measurement errors seem a natural requirement empirically for cognate series, and probably holds for the main series in National Income Accounts.

Many aspects remain to be explored including the possibility that measurement errors are worse when the economy experiences shifts (as from financial crises); and the role of co-breaking in the face of large measurement errors.

References

Boyer, G. R. and T. J. Hatton (2002). New estimates of British unemployment, 1870–1913. *Journal of Economic History 62*, 643–675.

Castle, J. L., J. A. Doornik, D. F. Hendry, and F. Pretis (2015). Detecting location shifts during model selection by step-indicator saturation. *Econometrics 3(2)*, 240–264.

Castle, J. L. and D. F. Hendry (2014). Semi-automatic non-linear model selection. In N. Haldrup, M. Meitz, and P. Saikkonen (Eds.), *Essays in Nonlinear Time Series Econometrics*, pp. 163–197. Oxford: Oxford University Press.

Davidson, J. E. H., D. F. Hendry, F. Srba, and J. S. Yeo (1978). Econometric modelling of the aggregate time-series relationship between consumers' expenditure and income in the United Kingdom. *Economic Journal 88*, 661–692.

Dickey, D. A. and W. A. Fuller (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica 49*, 1057–1072.

Doornik, J. A. and D. F. Hendry (2013). *Interactive Monte Carlo Experimentation in Econometrics using PcNaive: OxMetrics 7.10: Volume IV*. London: Timberlake Consultants Press.

Ericsson, N. R. and J. G. MacKinnon (2002). Distributions of error correction tests for cointegration. *Econometrics Journal 5*, 285–318.

Hall, P. and C. C. Heyde (1980). *Martingale Limit Theory and Its Application*. New York (USA): Academic Press.

Hendry, D. F. (1995). *Dynamic Econometrics*. Oxford: Oxford University Press.

Hendry, D. F. (2015). *Introductory Macro-econometrics: A New Approach*. London: Timberlake Consultants. http://www.timberlake.co.uk/macroeconometrics.html.

Hendry, D. F. and M. Massmann (2007). Co-breaking: Recent advances and a synopsis of the literature. *Journal of Business and Economic Statistics 25*, 33–51.

Hendry, D. F. and G. E. Mizon (2011). Econometric modelling of time series with outlying observations. *Journal of Time Series Econometrics 3 (1)*, DOI: 10.2202/1941–1928.1100.

Watson, M. W. (1994). Vector autoregressions and cointegration. *Handbook of Econometrics 4*, 2843–2915.

A Appendix

A.1 Proofs of propositions

For a series $\{x_t\}_{t=1}^T$, let $\overline{x}_t := x_t - T^{-1}\sum_{s=1}^T x_s$. For the statement of the next result, $\{\theta_t\}$ is as defined in (3.6) above, except that we shall also permit $\tau = 0$. Thus

$$n^{-\tau}\theta_{\lfloor nr\rfloor} \to \theta(r) := \sum_{i=1}^{K} \lambda_i \mathbf{1}\{r \geq r_i\}$$

in $D[0,1]$; let $\Theta(r) := \int_0^r \theta(s)\,\mathrm{d}s$, $\overline{\theta}(r) := \theta(r) - \int_0^1 \theta(s)\,\mathrm{d}s$, and $\overline{\Theta}(r) := \Theta(r) - \int_0^1 \Theta(s)\,\mathrm{d}s$. Throughout the following, $\int$ denotes $\int_0^1$, whenever limits of integration are left unspecified.

Lemma A.1. *Suppose that $w_t := (u_t, v_t)'$ is a linear process defined by*

$$w_t := \sum_{k=0}^{\infty} \Phi_k \epsilon_{t-k}$$

where $\Phi_k \in \mathbb{R}^{2\times 2}$, and $\epsilon_t \in \mathbb{R}^2$ is i.i.d. with $\mathbb{E}\epsilon_0 = \mu_\epsilon$, $\mathbb{E}\epsilon_0^2 = I_2$; and $\sum_{k=0}^{\infty} k\|\Phi_k\| < \infty$. Define $U_t := \sum_{s=1}^T u_t$, $V_t := \sum_{s=1}^T v_t$, and $Z_t := \sum_{s=1}^t \theta_s + U_t$. Then

(i) $T^{-3-2\tau}\sum_{t=1}^T \overline{Z}_t^2 \xrightarrow{p} \int \overline{\Theta}^2$

(ii) $T^{-2-2\tau}\sum_{t=1}^T \overline{\Delta Z}_t \overline{Z}_{t-1} \xrightarrow{p} \int \overline{\Theta\theta}$

(iii) $T^{-1-2\tau}\sum_{t=1}^T \overline{\Delta Z}_t \overline{\Delta Z}_{t-1} \xrightarrow{p} \int \overline{\theta}^2$

(iv) $T^{-5/2-\tau}\sum_{t=1}^T \overline{V}_t \overline{Z}_t \rightsquigarrow \int \overline{B}_V(r)\overline{\Theta}(r)\,\mathrm{d}r$

(v) $T^{-3/2-\tau}\sum_{t=1}^T \overline{V}_t \overline{\theta}_t \rightsquigarrow \int \overline{B}_V(r)\overline{\theta}(r)\,\mathrm{d}r$

(vi) $T^{-2}\sum_{t=1}^T \overline{U}_t \overline{V}_t \rightsquigarrow \int \overline{B}_U(r)\overline{B}_V(r)\,\mathrm{d}r$

(vii) $T^{-1}\sum_{t=1}^T \overline{u}_t \overline{V}_t = O_p(1)$

(viii) $T^{-3/2-\tau}\sum_{t=1}^{T}\overline{v}_t\overline{Z}_t = O_p(1)$

(ix) $T^{-1/2-\tau}\sum_{t=1}^{T}\overline{v}_t\overline{\theta}_t = O_p(1)$

(x) $T^{-1}\sum_{t=1}^{T}\overline{u}_t\overline{v}_t \xrightarrow{p} \operatorname{cov}(u_0, v_0)$

where $T^{-1/2}(U_{\lfloor nr_1\rfloor}, V_{\lfloor nr_2\rfloor}) \rightsquigarrow (B_U(r_1), B_V(r_2))$ *on* $D[0,1]^2$, *and* $\overline{B}_X(r) := B_X(r) - \int B_X(s)\,\mathrm{d}s$.

Remark A.1. In the pure drift model ($\theta \neq 0$ and $\lambda = 0$ in (3.2)), $\overline{\theta} = 0$ and $\overline{\Theta}(r) = \theta(r - \frac{1}{2})$.

Lemma A.2. *For ρ_t as in (3.8) above,*

$$\frac{1}{T^{2+\tau}}\sum_{t=1}^{T}\overline{\rho}_t\overline{z}_t + \frac{1}{T^{1+\tau}}\sum_{t=1}^{T}\overline{\rho}_t\overline{\Delta z}_t = o_p(1). \tag{A.1}$$

and $T^{-1}\sum_{t=1}^{n}\overline{\rho}_t^2 \xrightarrow{p} \sigma_\rho^2 = (1-\beta_1)^{-1}(\sigma_\epsilon^2 + 2\alpha^2\sigma_e^2)$.

Proofs of the preceding lemmas appear in Appendix A.2.

Proof. [Proof of Proposition 3.1] We shall derive the limiting distribution of the OLS estimators by supposing that the transformed regression model (3.4), reversing the (linear) transformations that lead from (3.1) to (3.4), allows us to recover the limiting distribution of $\widehat{\beta}$. Since

$$\mu_t = \beta_1\mu_{t-1} + (\beta_2 - \kappa)\Delta z_t + \epsilon_t,$$

it follows that $(\mu_{t-1}, \Delta z_t)$ is a linear process satisfying the requirements of Lemma A.1. Hence, by parts (i), (ii), (viii) and (x) of that result (and Remark A.1)

$$D_T^{-1}\left(\sum_{t=1}^{T}\begin{bmatrix}\overline{z}_{t-1}^2 & & \\ \overline{\mu}_{t-1}\overline{z}_{t-1} & \overline{\mu}_{t-1}^2 & \\ \overline{\Delta z}_t\overline{z}_{t-1} & \overline{\Delta z}_t\overline{\mu}_{t-1} & \overline{\Delta z}_t^2\end{bmatrix}\right)D_T^{-1} \xrightarrow{p} \begin{bmatrix}\theta^2(r-\frac{1}{2})^2 & & \\ 0 & \sigma_\mu^2 & \\ 0 & C_0 & \sigma_e^2\end{bmatrix}$$

where $D_T := \operatorname{diag}[T^{-3/2}, T^{-1/2}, T^{-1/2}]$, and

$$C_0 := \operatorname{cov}(\Delta z_t, \mu_{t-1}) = (\beta_2 - \kappa)\sum_{k=0}^{\infty}\beta_1^k\mathbb{E}e_te_{t-k-1}.$$

Thus $C_0 = 0$ here, but only owing to our rather special assumption that the stochastic component of z_t is a random walk; if this component were allowed to be a more general integrated process, then C_0 would be nonzero.

By the martingale CLT (Theorem 3.2 in Hall and Heyde (1980)),

$$\frac{1}{T^{1/2}}\sum_{t=1}^{T}\begin{bmatrix}\overline{\mu}_{t-1}\\ \overline{\Delta z}_t\end{bmatrix}\overline{\epsilon}_t = \frac{1}{T^{1/2}}\sum_{t=1}^{T}\begin{bmatrix}\mu_{t-1} - \mathbb{E}\mu_{t-1}\\ e_t\end{bmatrix}\epsilon_t + o_p(1) \rightsquigarrow \mathsf{N}\left[0, \begin{bmatrix}\sigma_\mu^2\sigma_\epsilon^2 & 0\\ 0 & \sigma_e^2\sigma_\epsilon^2\end{bmatrix}\right],$$

where the off-diagonals of the limiting variance matrix are zero, only in consequence of our assumption that e_t is i.i.d.; while by Lemma A.1(viii),

$$\frac{1}{T^{3/2}}\sum_{t=1}^{T}\overline{z}_{t-1}\overline{\epsilon}_t = O_p(1).$$

It thus follows that $\widehat{\beta}_1$ and $\widehat{\beta}_2$ have the claimed limiting distributions, while $\widehat{\kappa} - \kappa = O_p(T^{-3/2})$, where $\widehat{\kappa}$ denotes the OLS estimator of κ in (3.4). The second part of the result then follows from $\widehat{\beta}_3 = \widehat{\kappa} - (\kappa\widehat{\beta}_1 + \widehat{\beta}_2)$. □

Proof. [Proof of Proposition 3.2] Analogously to the preceding proof, we shall suppose that the parameters of the transformed (3.7) are estimated by OLS, and then reverse these transformations to recover the limiting distribution of $\widehat{\beta}$. By parts (i)–(iii) and (x) of Lemma A.1, and Lemma A.2, we have

$$D_T^{-1}\left(\sum_{t=1}^{T}\begin{bmatrix}\overline{z}_{t-1}^2 & & \\ \overline{\Delta z_t}\cdot\overline{z}_{t-1} & \overline{\Delta z}_t^2 & \\ \overline{\rho}_{t-1}\cdot\overline{z}_{t-1} & \overline{\rho}_{t-1}\cdot\overline{\Delta z_t} & \overline{\rho}_{t-1}^2\end{bmatrix}\right)D_T^{-1}\xrightarrow{p}\begin{bmatrix}\int\overline{\Theta}^2 & & \\ \int\overline{\Theta\theta} & \int\overline{\theta}^2 & \\ 0 & 0 & \sigma_\rho^2\end{bmatrix}, \tag{A.2}$$

where $D_T := \operatorname{diag}[T^{3/2+\tau}, T^{1/2+\tau}, T^{1/2}]$. By arguments similar to those used in the proof of Lemma A.2, and then the martingale CLT,

$$\frac{1}{T^{1/2}}\sum_{t=1}^{T}\overline{\rho}_{t-1}\overline{\epsilon}_t = \frac{1}{T^{1/2}}\sum_{t=1}^{T}\rho_{1,t-1}\epsilon_t + o_p(1) \rightsquigarrow \mathsf{N}[0,\ \sigma_\rho^2\sigma_\epsilon^2], \tag{A.3}$$

where ρ_{1t} is as defined in (A.10) below. Further, by Lemma A.1(ix) and the independence of e_t and ϵ_t,

$$\sum_{t=1}^{T}\overline{\Delta z}_t\overline{\epsilon}_t = \sum_{t=1}^{T}\overline{e}_t\overline{\epsilon}_t + \sum_{t=1}^{T}\overline{\theta}_t\epsilon_t = O_p(T^{1/2+\tau}) \tag{A.4}$$

while by Lemma A.1(viii)

$$\sum_{t=1}^{T}\overline{z}_{t-1}\overline{\epsilon}_t = O_p(T^{3/2+\tau}). \tag{A.5}$$

It follows from (A.2)–(A.5) that

$$T^{1/2}(\widehat{\beta}_1-\beta_1) = \frac{\frac{1}{T^{1/2}}\sum_{t=1}^{T}\overline{\rho}_{t-1}\overline{\epsilon}_t}{\frac{1}{T}\sum_{t=1}^{T}\overline{\rho}_{t-1}^2} + o_p(1),$$

which has the claimed limiting distribution; the asymptotic dependence of $(\widehat{\beta}_2, \widehat{\beta}_3)$ on $\widehat{\beta}_1$ in turn follows from

$$\begin{bmatrix}T^{3/2+\tau}(\widehat{\kappa}-\kappa)\\ T^{1/2+\tau}(\widehat{\gamma}-\gamma)\end{bmatrix} = O_p(1).$$

□

Proof. [Proof of Proposition 4.1] It will be convenient to rewrite the model as

$$y_t^* = \mu_{t-1}^* + \beta_2\Delta z_t^* + \kappa z_{t-1}^* + \zeta_t \tag{A.6}$$

where $\zeta_t := \epsilon_t + v_t - \beta_2\eta_t - (1-\beta_1)\mu_{t-1}$; note that a regression of $\overline{y}_t^*$ on $(\overline{z}_{t-1}^*, \overline{\mu}_{t-1}^*, \Delta\overline{z}_t^*)$ yields the OLS estimates of (β_1, β_2). We have

$$\mu_t^* = \mu_t + T^{-1/2}\pi_t \qquad\qquad z_t^* = z_t + u_t$$

where μ_t is a linear process, while π_t and u_t are I(1). Thus

$$\begin{aligned}\frac{1}{T^2}\sum_{t=1}^{T}\overline{\mu}_{t-1}^*\overline{z}_{t-1}^* &= \frac{1}{T^{5/2}}\sum_{t=1}^{T}\overline{\pi}_{t-1}(\overline{z}_{t-1}+\overline{u}_{t-1}) + \frac{1}{T^2}\sum_{t=1}^{T}\overline{\mu}_{t-1}(\overline{z}_{t-1}+\overline{u}_{t-1})\\ &= \frac{1}{T^{5/2}}\sum_{t=1}^{T}\overline{\pi}_{t-1}\overline{z}_{t-1} + o_p(1)\\ &\rightsquigarrow \theta\int\overline{B}_\pi(r)(r-\tfrac{1}{2})\,\mathrm{d}r\end{aligned}$$

successively by parts (vii), (viii), and (iv) of Lemma A.1. Further,

$$\begin{aligned}\frac{1}{T^2}\sum_{t=1}^{T}\overline{\Delta z_t^* z_{t-1}^*} &= \frac{1}{T^2}\sum_{t=1}^{T}(\overline{\Delta z_t} + \overline{\eta}_t)(\overline{z}_{t-1} + \overline{u}_{t-1}) \\ &= \frac{1}{T^2}\sum_{t=1}^{T}(\overline{\theta}_t + \overline{e}_t)\overline{u}_{t-1} + o_p(1) \\ &= o_p(1) \end{aligned} \tag{A.7}$$

successively by parts (ii), (vii), (viii), (ix) and (x) of Lemma A.1, and the fact that $\overline{\theta}_t = 0$; while

$$\frac{1}{T}\sum_{t=1}^{T}(\overline{\mu}_{t-1}^*)^2 = \frac{1}{T}\sum_{t=1}^{T}\overline{\mu}_{t-1}^2 + \frac{1}{T^2}\sum_{t=1}^{T}\overline{\pi}_{t-1}^2 + \frac{1}{T^{3/2}}\sum_{t=1}^{n}\overline{\mu}_{t-1}\overline{\pi}_{t-1} \rightsquigarrow \sigma_\mu^2 + \int \overline{B}_\pi^2,$$

by parts (vi), (vii) and (x) of Lemma A.1; and

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T}\overline{\Delta z_t^*\mu_{t-1}^*} &= \frac{1}{T}\sum_{t=1}^{T}(\overline{e}_t + \overline{\eta}_t)\overline{\mu}_{t-1} + \frac{1}{T^{3/2}}\sum_{t=1}^{T}(\overline{e}_t + \overline{\eta}_t)\overline{\pi}_{t-1} \\ &\xrightarrow{p} \operatorname{cov}(e_t, \mu_{t-1}) + \operatorname{cov}(\eta_t, \mu_{t-1}) \\ &= 0\end{aligned}$$

by $\overline{\theta}_t = 0$ and parts (vii) and (x) of Lemma A.1. Arguing similarly for the remaining terms in the design matrix, we thus obtain

$$\begin{aligned}&D_T^{-1}\left(\sum_{t=1}^{T}\begin{bmatrix}(\overline{z}_{t-1}^*)^2 & & \\ \overline{\mu_{t-1}^* z_{t-1}^*} & (\overline{\mu}_{t-1}^*)^2 & \\ \overline{\Delta z_t^* z_{t-1}^*} & \overline{\Delta z_t^*\mu_{t-1}^*} & (\overline{\Delta z_t^*})^2\end{bmatrix}\right)D_T^{-1} \\ &\qquad\rightsquigarrow \begin{bmatrix}\theta^2\int(r-\frac{1}{2})^2 & & \\ \theta\int \overline{B}_\pi(r)(r-\frac{1}{2})\,\mathrm{d}r & \sigma_\mu^2 + \int \overline{B}_\pi^2 & \\ 0 & 0 & \sigma_e^2 + \sigma_\eta^2\end{bmatrix},\end{aligned} \tag{A.8}$$

where $D_T \coloneqq \operatorname{diag}[T^{3/2}, T^{1/2}, T^{1/2}]$.

Since ζ_t is a linear process, we have

$$\frac{1}{T^2}\sum_{t=1}^{T}\overline{z}_{t-1}^*\overline{\zeta}_t = \frac{1}{T^2}\sum_{t=1}^{T}(\overline{z}_{t-1} + \overline{u}_t)\overline{\zeta}_t = o_p(1)$$

by parts (viii) and (x) of Lemma A.1; while

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T}\overline{\mu}_{t-1}^*\overline{\zeta}_t &= \frac{1}{T}\sum_{t=1}^{T}\overline{\mu}_{t-1}\overline{\zeta}_t + \frac{1}{T^{3/2}}\sum_{t=1}^{n}\overline{\pi}_{t-1}\overline{\zeta}_t \\ &\xrightarrow{p} \operatorname{cov}(\mu_{t-1}, \zeta_t) \\ &= -(1-\beta_1)\sigma_\mu^2\end{aligned}$$

by parts (vii) and (x) of Lemma A.1; and similarly

$$\frac{1}{T}\sum_{t=1}^{T}\overline{\Delta z_t^*\zeta_t} = \frac{1}{T}\sum_{t=1}^{T}(\overline{e}_t + \overline{\eta}_t)\overline{\zeta}_t \xrightarrow{p} \operatorname{cov}(e_t, \zeta_t) + \operatorname{cov}(\eta_t, \zeta_t) = -\beta_2\sigma_\eta^2.$$

Hence

$$T^{-1/2}D_T^{-1}\sum_{t=1}^{T}\begin{bmatrix}\overline{z}_{t-1}^{*}\\ \overline{\mu}_{t-1}^{*}\\ \overline{\Delta z_t^{*}}\end{bmatrix}\overline{\zeta}_t \xrightarrow{p} \begin{bmatrix}0\\ -(1-\beta_1)\sigma_\mu^2\\ -\beta_2\sigma_\eta^2\end{bmatrix}, \tag{A.9}$$

and so it follows from (A.8) and (A.9) that

$$\begin{bmatrix}n(\widehat{\kappa}-\kappa)\\ \widehat{\beta}_1-1\\ \widehat{\beta}_2-\beta_2\end{bmatrix} \rightsquigarrow \begin{bmatrix}\theta^2\int(r-\frac{1}{2})^2 & \theta\int\overline{B}_\pi(r)(r-\frac{1}{2})\,\mathrm{d}r & 0\\ \theta\int\overline{B}_\pi(r)(r-\frac{1}{2})\,\mathrm{d}r & \sigma_\mu^2+\int\overline{B}_\pi^2 & 0\\ 0 & 0 & \sigma_e^2+\sigma_\eta^2\end{bmatrix}^{-1}\begin{bmatrix}0\\ -(1-\beta_1)\sigma_\mu^2\\ -\beta_2\sigma_\eta^2\end{bmatrix}.$$

Direct calculation of the inverse on the right side of preceding yields

$$\widehat{\beta}_1 \rightsquigarrow 1-(1-\beta_1)\frac{\sigma_\mu^2}{\sigma_\mu^2+\int\widetilde{B}_\pi^2}=\beta_1+\frac{1-\beta_1}{1+R}$$

where $\widetilde{B}_\pi$ denotes the residual from projecting B_π onto a constant and a linear trend, and R is as in the text; it is also apparent that

$$\widehat{\beta}_2 \xrightarrow{p} \beta_2-\beta_2\frac{\sigma_\eta^2}{\sigma_e^2+\sigma_\eta^2}=\beta_2\frac{\sigma_e^2}{\sigma_e^2+\sigma_\eta^2}.$$

□

Proof. [Proof of Proposition 5.1] As in the proof of Proposition 3.2, we shall work with the estimates of the transformed model (3.7). Noting that $\rho_t^*=\rho_t+\delta_t-\alpha\eta_{t+1}$, we have,

$$\begin{aligned}\frac{1}{T^{2+\tau}}\sum_{t=1}^{T}\overline{\rho}_{t-1}^{*}\overline{z}_{t-1}^{*} &= \frac{1}{T^{2+\tau}}\sum_{t=1}^{T}(\overline{\rho}_{t-1}+\overline{\delta}_{t-1}-\alpha\overline{\eta}_t)(\overline{z}_{t-1}+\overline{u}_t)\\ &= \frac{1}{T^{2+\tau}}\sum_{t=1}^{T}\overline{\rho}_{t-1}\overline{z}_{t-1}+o_p(1)\\ &= o_p(1)\end{aligned}$$

by parts (viii) and (x) of Lemma A.1, and a straightfoward extension of Lemma A.2. Analogously,

$$\frac{1}{T^{1+\tau}}\sum_{t=1}^{T}\overline{\rho}_{t-1}^{*}\cdot\overline{\Delta z_t^{*}}=\frac{1}{T^{1+\tau}}\sum_{t=1}^{T}\overline{\rho}_{t-1}\overline{\Delta z}_t+o_p(1)=o_p(1),$$

whence

$$D_T^{-1}\left(\sum_{t=1}^{T}\begin{bmatrix}(\overline{z}_{t-1}^{*})^2 & & \\ \overline{\Delta z_t^{*}}\cdot\overline{z}_{t-1}^{*} & (\overline{\Delta z_t^{*}})^2 & \\ \overline{\rho}_{t-1}^{*}\cdot\overline{z}_{t-1}^{*} & \overline{\rho}_{t-1}^{*}\cdot\overline{\Delta z_t^{*}} & (\overline{\rho}_{t-1}^{*})^2\end{bmatrix}\right)D_T^{-1}\xrightarrow{p}\begin{bmatrix}\int\overline{\Theta}^2 & & \\ \int\overline{\Theta\theta} & \int\overline{\theta}^2 & \\ 0 & 0 & \sigma_\rho^2+\sigma_\delta^2+\alpha^2\sigma_\eta^2\end{bmatrix},$$

for $D_T\coloneqq\operatorname{diag}[T^{3/2+\tau},T^{1/2+\tau},T^{1/2}]$.

Since ξ_t is a linear process,

$$\frac{1}{T^{2+\tau}}\sum_{t=1}^{T}\overline{z}_t^{*}\overline{\xi}_t=\frac{1}{T^{2+\tau}}\sum_{t=1}^{T}\overline{z}_t\overline{\xi}_t+o_p(1)=o_p(1)$$

by parts (vii), (viii) and (x) of Lemma A.1; while

$$\frac{1}{T^{1+\tau}}\sum_{t=1}^{T}\overline{\Delta z}_t^*\overline{\xi}_t = \frac{1}{T^{1+\tau}}\sum_{t=1}^{T}(\overline{\theta}_t + \overline{e}_t + \overline{\eta}_t)\overline{\xi}_t + o_p(1) = o_p(1)$$

by parts (ix) and (x) of Lemma A.1; and

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T}\overline{\rho}_{t-1}^*\overline{\xi}_t &= \frac{1}{T}\sum_{t=1}^{T}(\overline{\rho}_{1,t-1} + \overline{\delta}_{t-1} - \alpha\overline{\eta}_t)\overline{\xi}_t + o_p(1)\\ &\xrightarrow{p} \operatorname{cov}(\rho_{1,t-1} + \delta_{t-1} - \alpha\eta_t,\ (1-\beta_1-\psi_1)\delta_{t-1} - \beta_2\eta_t + \epsilon_t)\\ &= (1-\beta_1-\psi_1)\sigma_\delta^2 + \alpha\beta_2\sigma_\eta^2.\end{aligned}$$

It follows that

$$T^{-1/2}D_T^{-1}\sum_{t=1}^{T}\begin{bmatrix}\overline{z}_{t-1}^*\\ \overline{\Delta z}_t^*\\ \overline{\rho}_{t-1}^*\end{bmatrix}\overline{\xi}_t \xrightarrow{p} \begin{bmatrix}0\\0\\(1-\beta_1-\psi_1)\sigma_\delta^2 + \alpha\beta_2\end{bmatrix},$$

whence $\widehat{\beta}_1$ exhibits the claimed asymptotic behaviour. In view of the preceding, $(\widehat{\kappa},\widehat{\alpha}) \xrightarrow{p} (\kappa,\alpha)$, whence the asymptotic bias in $(\widehat{\beta}_2,\widehat{\beta}_3)$ depends only on that in $\widehat{\beta}_1$. □

A.2 Proofs of auxiliary lemmas

Proof. [Proof of Lemma A.1] Since $T^{-1/2}V_{\lfloor Tr\rfloor} \rightsquigarrow B_V(r)$,

$$T^{-1-\tau}z_{\lfloor Tr\rfloor} = \frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\theta_t + \frac{1}{T^{1+\tau}}\sum_{s=1}^{\lfloor Tr\rfloor}e_s \xrightarrow{p} \int_0^r \theta(s)\,\mathrm{d}s =: \Theta(r)$$

and

$$T^{-\tau}\Delta z_{\lfloor Tr\rfloor} = \theta_{\lfloor Tr\rfloor} + T^{-\tau}e_{\lfloor Tr\rfloor} \xrightarrow{p} \theta(r)$$

on $D[0,1]$, parts (i)–(vi) follow by the continuous mapping theorem. (vii) and (viii) are a consequence of Lemma 2.3 in Watson (1994), while (ix) and (x) follow respectively from the CLT and LLN for linear processes. □

Proof. [Proof of Lemma A.2] Since K in (3.6) is fixed and finite, it suffices to consider the case where $K = 1$ and $r_1 \in (0,1)$. It follows from (3.8) that

$$\rho_t = \sum_{k=0}^{\infty}\beta_1^k\varphi_{t+1-k} - \alpha\sum_{k=0}^{\infty}\beta_1^k\Delta\theta_{t+1-k} =: \rho_{1t} - \alpha\rho_{2t}, \tag{A.10}$$

where we have defined $\varphi_t := \epsilon_t - \alpha(e_{t+1} - e_t)$. Since ρ_{1t} is a linear process, it follows by parts (viii) and (ix) of Lemma A.1 that (A.1) holds with $\overline{\rho}_{1t}$ in place of $\overline{\rho}_t$. Regarding ρ_{2t}, we note that since $K = 1$, $\Delta\theta_t = \lambda_1 T^\tau$ when $t = \lfloor Tr_1\rfloor$, and is zero otherwise. Thus

$$\rho_{2t} = \begin{cases}0 & \text{if } 0 \le t < \lfloor Tr_1\rfloor - 1,\\ \beta_1^{t+1-\lfloor Tr_1\rfloor}\lambda_1 T^\tau & \text{otherwise}\end{cases}$$

and so

$$\frac{1}{T^{1+\tau}}\sum_{t=1}^{T}\rho_{2t}\Delta z_t = \frac{\lambda_1}{T}\sum_{t=\lfloor Tr_1\rfloor}^{T}\beta_1^{t+1-\lfloor Tr_1\rfloor}\Delta z_t = o_p(1) \tag{A.11}$$

with the final order estimate following from

$$\frac{1}{T}\sum_{t=\lfloor Tr_1\rfloor}^{T}\beta_1^{t+1-\lfloor Tr_1\rfloor}\mathbb{E}|\Delta z_t| \leq \frac{1}{T^{1-\tau}}\sum_{k=0}^{\infty}\beta_1^k[\lambda_1+\mathbb{E}|e_t|] = o(1).$$

It follows easily that (A.11) holds when ρ_{2t} and Δz_t are replaced by their demeaned versions; and thus the second term on the right of (A.1) has the claimed order. An analogous argument establishes the result for the first term. The second part of the lemma follows from the fact that

$$\frac{1}{T}\sum_{t=1}^{n}\overline{\rho}_t^2 = \frac{1}{T}\sum_{t=1}^{n}\overline{\rho}_{1,t}^2 + o_p(1),$$

which may be proved via a further variation on the preceding arguments. □