Debt, Inflation and Growth

Robust Estimation of Long-Run Effects in Dynamic Panel Data Models

Alexander Chudik
Kamiar Mohaddes
M. Hashem Pesaran
Mehdi Raissi

CESIFO WORKING PAPER NO. 4508
CATEGORY 12: EMPIRICAL AND THEORETICAL METHODS
DECEMBER 2013

CESifo *Working Paper No. 4508*

Debt, Inflation and Growth

Robust Estimation of Long-Run Effects in Dynamic Panel Data Models

Abstract

This paper investigates the long-run effects of public debt and inflation on economic growth. Our contribution is both theoretical and empirical. On the theoretical side, we develop a cross-sectionally augmented distributed lag (CS-DL) approach to the estimation of long-run effects in dynamic heterogeneous panel data models with cross-sectionally dependent errors. The relative merits of the CS-DL approach and other existing approaches in the literature are discussed and illustrated with small sample evidence obtained by means of Monte Carlo simulations. On the empirical side, using data on a sample of 40 countries over the 1965-2010 period, we find significant negative long-run effects of public debt and inflation on growth. Our results indicate that, if the debt to GDP ratio is raised and this increase turns out to be permanent, then it will have negative effects on economic growth in the long run. But if the increase is temporary, then there are no long-run growth effects so long as debt to GDP is brought back to its normal level. We do not find a universally applicable threshold effect in the relationship between public debt and growth. We only find statistically significant threshold effects in the case of countries with rising debt to GDP ratios.

JEL-Code: C230, E620, F340, H600.

Keywords: long-run relationships, estimation and inference, large dynamic heterogeneous panels, cross-section dependence, debt, inflation and growth, debt overhang.

Alexander Chudik
Federal Reserve Bank of Dallas
Dallas / Texas / USA
alexander.chudik@dal.frb.org

*Kamiar Mohaddes**
Faculty of Economics & Girton College
University of Cambridge / UK
km418@cam.ac.uk

M. Hashem Pesaran
Department of Economics
University of Southern California / USA
pesaran@dornsife.usc.edu

Mehdi Raissi
International Monetary Fund
Washington DC / USA
mraissi@imf.org

*corresponding author

November 2013
We are grateful to Luis Catão, Markus Eberhardt, Thomas Moutos, Kenneth Rogoff, Alessandro Rebucci, Ron Smith, Martin Weale, Mark Wynne, and seminar participants at the International Monetary Fund and participants at the Conference on MENA Economies 2013 for constructive comments and suggestions. The views expressed in this paper are those of the authors and do not necessarily represent those of Federal Reserve Bank of Dallas, the Federal Reserve System, the International Monetary Fund or IMF policy. Hashem Pesaran acknowledges financial support under ESRC Grant ES/I031626/1.

1 Introduction

The debt-growth nexus has received renewed interest among academics and policy makers alike in the aftermath of the recent global financial crisis and the subsequent euro area sovereign debt crisis which has triggered trillions of dollars in fiscal stimulus across the globe. This paper investigates whether a build-up of public debt slows down the economy in the long run. The conventional view is that public debt (arising from deficit financing) can stimulate aggregate demand and output in the short run, but crowds out capital and reduces output in the long run. In addition, there are possible non-linear effects where the build-up of debt can harm economic growth especially when the level of debt exceeds a certain threshold, as estimated, for example, by Reinhart and Rogoff (2010) to be around 90% of the GDP. However, such results are obtained under strong homogeneity assumptions across countries, and without adequate attention to dynamics, feed-back effects from debt to GDP, and error cross-sectional dependencies that exist across countries, due to unobserved common factors or spill-over effects that tend to magnify at times of financial crises. Due to the intrinsic cross-country heterogeneities, the thresholds are most-likely country specific and estimation of a universal threshold based on pooling of observations across countries might not be informative to policy makers interested in a particular economy and their use could be even misleading. Relaxing the homogeneity assumption, whilst possible in a number of dimensions (as seen below), is difficult when it comes to the estimation of country-specific thresholds, because due to the non-linearity of the relationships involved, identification and estimation of country-specific thresholds require much larger time series data than are currently available.

In this paper we model the growth rates, as opposed to levels of (log) GDP and debt to GDP, which allows us to make inferences about the long-term effects of debt on growth, regardless of thresholds. Using recent developments in the literature on dynamic heterogeneous panels, we provide a fresh re-examination of debt-growth nexus while allowing for dynamic heterogeneities and cross-sectional error dependencies. Our focus will be on the long-run impacts of debt and inflation on GDP growth which will be shown to be robust to feedbacks from growth to debt and inflation. We use a relatively large panel of advanced and emerging market economies, and jointly model inflation, debt, and growth. We consider the role of inflation in our long-run analysis because, in some countries in the panel that do not have active government bond markets, deficit financing is often achieved through money creation with high inflation. Like excessively high levels of debt, high levels of inflation, when persistent, can also be detrimental for growth. By considering both inflation and debt we allow the regression analysis to accommodate both types of economies in the panel.

The paper also makes a theoretical contribution to the econometric analysis of the long run. A new approach to the estimation of the long-run coefficients in dynamic heterogeneous panels with cross-sectionally dependent errors is proposed. The approach is based on a distributed lag representation that does not feature lags of the dependent variable, and allows for a residual factor error structure and weak cross-section dependence of idiosyncratic errors. Similarly to Common Correlated Effects (CCE) estimators proposed by Pesaran (2006), we appropriately augment the individual regressions by cross-section averages to deal with the effects of common factors. We derive the asymptotic distribution of the proposed cross-section augmented distributed lag (or CS-DL in short) mean group and pooled estimators under the coefficient heterogeneity and large time (T) and cross section (N) dimensions. We also investigate consequences of various departures from our maintained assumptions by means of Monte Carlo experiments, including unit root in factors and/or in regressors, homogeneity of coefficients or breaks in error processes. The small sample evidence suggests that the CS-DL estimators often outperform the traditional approach based on estimating the full autoregressive distributed lag (ARDL) specification. However, the CS-DL approach should be seen as complementary and not as superior to the ARDL approach due to its two drawbacks: unlike the panel ARDL approach it does not allow for feedback effects from the dependent variable onto the regressors, and its small sample performance deteriorates when the roots of the AR polynomial in the ARDL representation are close to the unit circle. The relative merits of different approaches are carefully documented in the paper.

Our empirical contribution is in estimating long-run effects of debt and inflation on economic growth in a panel of 40 countries over the period 1965–2010. Cross-country experience shows that some economies have run into debt difficulties and experienced subdued growth at relatively low debt levels, while others have been able to sustain high levels of indebtedness for prolonged periods and grow strongly without experiencing debt distress. This suggests that the effects of public debt on growth varies across countries, depending critically on country-specific factors and institutions.[1] It is therefore important that we take account of cross-country heterogeneity. The dynamics should also be modelled properly, otherwise the estimates of the long-run effects might be inconsistent. Last but not least, it is now widely agreed that conditioning on observed variables specific to countries alone need not ensure error cross-section independence that underlies much of the panel data literature. It is, therefore, also important that we allow for the possibility of cross-sectional error correlations, which could arise due to omitted common effects, possibly correlated with the regressors. Neglecting such dependencies can lead to biased estimates and spurious inference.

We adopt a cross-section augmented ARDL approach (CS-ARDL), advanced in Chudik and Pesaran (2013a), and a CS-DL approach developed in this paper. This estimation strategy takes into account all three key features of the panel (i.e. dynamics, heterogeneity and cross-sectional dependence) jointly, in contrast with the earlier literature surveyed in Section 5. We study whether there is a common threshold for government debt ratios above

[1] These might include prospects for primary fiscal surpluses and growth; cost of borrowing including both the interest cost of debt already contracted and market perceptions of a country's ability to service future borrowings; regulatory requirements; nature of the investor base and the track record of meeting its debt obligations (whether it had debt distress/lost market access); and vulnerability to shocks (confidence effects).

which long-term growth rates are adversely affected (especially if the country is on an upward debt trajectory). We particularly look into debt *trajectory* beyond certain debt threshold levels as to our knowledge no such systematic analysis has been carried out in the past. We do not find a universally applicable threshold effect in the relationship between debt and growth. We only find a statistically significant threshold effect in the case of countries with rising debt to GDP ratios. The debt trajectory seems much more important than the level of debt itself. Provided that debt is on a downward path, a country with a high level of debt can grow just as fast as its peers. This "no-simple-debt-threshold-level" finding can be driven, among other possible factors, by cross-country differences in (i) overall net wealth (international investment position) and the depth of financial system; (ii) investor behavior (home bias); (iii) ability to generate primary surpluses and interest costs–growth considerations; and (iv) confidence factors. Our results also show that, regardless of the threshold, there are significant and robust negative long-run effects of debt on economic growth. By comparison, the evidence of a negative effect of inflation on growth is less strong, although it is statistically significant in the case of most specifications considered.

Our results suggest that if the debt level is raised and this increase is permanent, then it will have negative effects on growth in the long run. On the other hand, if the debt rises (for instance to help smooth out business cycle fluctuations) and this increase is temporary, then there are no long-run negative effects on output growth. The key in debt financing is the reassurance, backed by commitment and action, that the increase in government debt is temporary and will not be a permanent departure from the prevailing norms.

The remainder of the paper is organized as follows. We begin with the definition of long-run coefficients and discuss their estimation in Section 2. The next section introduces the CS-DL approach to the estimation of long-run relationships. Section 4 investigates the small sample performance of the CS-DL approach and compares it with the performance of the CS-ARDL approach by means of Monte Carlo experiments. Section 5 reviews the literature on long-run effects of inflation and debt on economic growth. Section 6 presents empirical findings on the long-run effects of debt and inflation on economic growth in our panel of countries. The last section concludes. Mathematical derivations and other supporting material are relegated to the Appendix.

A brief word on notation: All vectors are column vectors represented by bold lower case letters and matrices are represented by bold capital letters. $\|\mathbf{A}\| = \sqrt{\varrho(\mathbf{A}'\mathbf{A})}$ is the spectral norm of $\mathbf{A}$, $\varrho(\mathbf{A})$ is the spectral radius of $\mathbf{A}$.[2] $a_n = O(b_n)$ denotes the deterministic sequence $\{a_n\}$ is at most of order b_n. Convergence in probability and convergence in distribution are denoted by $\xrightarrow{p}$ and $\xrightarrow{d}$, respectively. $(N,T) \xrightarrow{j} \infty$ denotes joint asymptotic in N and T, with N and $T \to \infty$, in no particular order. We use K to denote a positive fixed constant that does not vary with N or T.

[2] Note that if $\mathbf{x}$ is a vector, then $\|\mathbf{x}\| = \sqrt{\varrho(\mathbf{x}'\mathbf{x})} = \sqrt{\mathbf{x}'\mathbf{x}}$ corresponds to the Euclidean length of vector $\mathbf{x}$.

2 Estimation of long-run or level relationships in economics

Estimating long-run or level relationships is of great importance in economics. The concept of the long-run in economics is associated with the steady-state solution of a structural model. Often the same long-run relations can also be obtained from arbitrage conditions within and across markets. As a result many long-run relationships in economics are free of particular model assumptions; examples being purchasing power parity, uncovered interest parity and the Fisher inflation parity. Other long-run relations, such as those between macroeconomic aggregates like consumption and income, output and investment, technological progress and real wages, are less grounded in arbitrage and hence are more controversial, but still form a major part of what is generally agreed in empirical macro modelling. This is in contrast to the analysis of short-run effects which are model specific and subject to identification problems.

The estimation of long-run relations can be carried out with or without constraining the short-run dynamics (possibly from a particular theory). In this section we focus on the estimation of long-run relations without restricting the short-run dynamics. In view of the empirical application that we have in mind, we shall assume that there exists a single long-run relationship between the dependent variable, y_t, and a set of regressors.[3] For illustrative purposes, suppose that there is one regressor x_t and suppose that $\mathbf{z}_t = (y_t, x_t)'$ is jointly determined by the following vector autoregressive model of order 1, VAR(1),

$$\mathbf{z}_t = \mathbf{\Phi}\mathbf{z}_{t-1} + \mathbf{e}_t, \tag{1}$$

where $\mathbf{\Phi} = (\phi_{ij})$ is a 2×2 matrix of unknown parameters, and $\mathbf{e}_t = (e_{yt}, e_{xt})'$ is 2-dimensional vector of reduced form errors. Denoting the covariance of e_{yt} and e_{xt} by $\omega Var(e_{xt})$, we can write

$$e_{yt} = E\left(e_{yt}\left|e_{xt}\right.\right) + u_t = \omega e_{xt} + u_t, \tag{2}$$

where by construction u_t is uncorrelated with e_{xt}, namely $E\left(u_t\left|e_{xt}\right.\right) = 0$. Substituting (2) for e_{yt}, the equation for the dependent variable y_t in (1) is

$$y_t = \phi_{11}y_{t-1} + \phi_{12}x_{t-1} + \omega e_{xt} + u_t. \tag{3}$$

Using the equation for the regressor x_t in (1), we obtain the following expression for e_{xt}

$$e_{xt} = x_t - \phi_{21}y_{t-1} - \phi_{22}x_{t-1},$$

[3] The problem of estimation and inference in the case of multiple long-run relations is further complicated by the identification problem and simultaneous determination of variables. The case of multiple long-run relations is discussed for example in Pesaran (1997).

and substituting this expression for e_{xt} back in (3) yields the following *conditional* model for y_t,

$$y_t = \varphi y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + u_t, \tag{4}$$

where

$$\varphi = \phi_{11} - \omega\phi_{21},\ \beta_0 = \omega,\ \beta_1 = \phi_{12} - \omega\phi_{22}. \tag{5}$$

Note that u_t is uncorrelated with the regressor x_t and its lag by construction. (4) is ARDL(1,1) representation of y_t conditional on x_t, and the short-run coefficients φ, β_0, and β_1 can be directly estimated from (4) by least squares. Model (4) can also be written as the following error-correction model,

$$\Delta y_t = -\left(1-\varphi\right)\left(y_{t-1} - \theta x_{t-1}\right) + \beta_0 \Delta x_t + u_t,$$

or as the following level relationship

$$y_t = \theta x_t + \alpha\left(L\right)\Delta x_t + \tilde{u}_t,$$

where the level coefficient is defined by the ratio

$$\theta = \frac{\beta_0 + \beta_1}{1-\varphi},$$

$\tilde{u}_t = \left(1-\varphi L\right)^{-1} u_t$ is uncorrelated with regressor x_t and its lags, and $\alpha\left(L\right) = \sum_{\ell=0}^{\infty}\alpha_\ell L^\ell$, with $\alpha_\ell = \sum_{s=\ell+1}^{\infty}\delta_s$, for $\ell = 0,1,2,...$, and $\delta\left(L\right) = \sum_{\ell=0}^{\infty}\delta_\ell L^\ell = \left(1-\varphi L\right)^{-1}\left(\beta_0 + \beta_1 L\right)$. Note that if $\mathbf{z}_t$ is $I\left(1\right)$ then $\left(1,-\theta\right)'$ is the cointegrating vector and the level relation is also cointegrating.

The level coefficient θ can still be motivated as the long-run outcome of a counterfactual exercise even if $\mathbf{z}_t$ is stationary . One possible counterfactual is to consider the effects of a permanent shock to the x_t process on y_t in the long run. Let

$$g_{yt} = \lim_{s\to\infty} E\left(y_{t+s} - \mu_{y,t+s}\middle| \mathcal{I}_{t-1}, e_{x,t+h} = \sigma_x, \text{ for } h = 0,1,2,...\right),$$

and similarly

$$g_{xt} = \lim_{s\to\infty} E\left(x_{t+s} - \mu_{x,t+s}\middle| \mathcal{I}_{t-1}, e_{x,t+h} = \sigma_x, \text{ for } h = 0,1,2,...\right),$$

where μ_{yt} and μ_{xt}, respectively, are the deterministic components of y_t and x_t (in the current illustrative example deterministic components are zero) and $\mathcal{I}_t$ is the information set containing all information up to the period t. Using (1) and noting that $E\left(e_{yt}\left|e_{xt}\right.\right) = \omega e_{xt}$, we obtain $g_{yt} = g_y$, $g_{xt} = g_x$,[4]

$$\mathbf{g} = \begin{pmatrix} g_y \\ g_x \end{pmatrix} = (\mathbf{I}_2 - \mathbf{\Phi})^{-1} \begin{pmatrix} \omega \\ 1 \end{pmatrix} \sigma_x = \begin{pmatrix} -\frac{\omega+\phi_{12}-\omega\phi_{22}}{\phi_{11}+\phi_{22}-\phi_{11}\phi_{22}+\phi_{12}\phi_{21}-1} \\ -\frac{\omega\phi_{21}-\phi_{11}+1}{\phi_{11}+\phi_{22}-\phi_{11}\phi_{22}+\phi_{12}\phi_{21}-1} \end{pmatrix} \sigma_x,$$

and

$$\frac{g_y}{g_x} = \frac{\omega + \phi_{12} - \omega\phi_{22}}{1 - (\phi_{11} - \omega\phi_{21})},$$

which upon using (5), yields, $g_y = \theta g_x$, namely the long-run impact of a permanent change in the mean of x on y is given by θ. Note that only in the special case when the reduced form errors are uncorrelated ($\omega = 0$) then the short-run coefficient β_0 in the ARDL model (4) is equal to 0 and the long-run coefficient θ reduces to $\phi_{12}/(1 - \phi_{11})$. But in general, when $\omega \neq 0$, the short-run coefficient β_0 is non-zero and contemporaneous values of the regressor should not be excluded from (4). In the stationary case with regressors not strictly exogenous, θ depends also on the parameters of the x_t process and the estimation of θ should therefore be based on (4).

An alternative way to show that θ is equal to the ratio g_y/g_x is to consider the ARDL representation (4) for the future period $t + s$, given the information at time $t - 1$. We first note that

$$y_{t+s} = \varphi y_{t+s-1} + \beta_0 x_{t+s} + \beta_1 x_{t+s-1} + u_{t+s},$$

and after taking the conditional expectation with respect to $\{\mathcal{I}_{t-1}, e_{x,t+h} = \sigma_x, \text{ for } h = 0, 1, 2, ...\}$, taking limits as $s \to \infty$, and noting that in the stationary case $g_{yt} = g_y$ and $g_{xt} = g_x$, we obtain

$$g_y = \varphi g_y + \beta_0 g_x + \beta_1 g_x,$$

and hence

$$\frac{g_y}{g_x} = \frac{\beta_0 + \beta_1}{1 - \varphi} = \theta,$$

as desired.

Regardless of whether the variables are $I(0)$ or $I(1)$, or whether the regressors are exogenous or not, the level coefficient θ is well defined and can be consistently estimated. The rates of convergence and the asymptotic distributions of the ARDL estimates of θ are established in Pesaran and Shin (1999). See in particular their Theorem 3.3.

2.1 Two approaches to the estimation of long-run effects

Let y_{it} be the dependent variable in country i, $\mathbf{x}_{it}$ be the $k \times 1$ vector of country-specific regressors, and suppose that the object of interest is the long-run coefficient vector of country i, denoted as $\boldsymbol{\theta}_i$, or, in a multicounty context, the average long-run coefficients vector, $\bar{\boldsymbol{\theta}} = N^{-1}\Sigma_{i=1}^{N}\boldsymbol{\theta}_i$. In modelling the relationship between the dependent variable and the

[4] Note that in the stationary case $\sum_{\ell=0}^{\infty} \mathbf{\Phi}^{\ell} = (\mathbf{I} - \mathbf{\Phi})^{-1}$.

regressors in a panel context, we need to allow for slope heterogeneity, dynamics and cross-sectional dependence. This is accomplished by assuming that the dependent variable is given by the following $ARDL(p_{yi}, p_{xi})$ specification,

$$y_{it} = \sum_{\ell=1}^{p_{yi}} \varphi_{i\ell} y_{i,t-\ell} + \sum_{\ell=0}^{p_{xi}} \boldsymbol{\beta}'_{i\ell} \mathbf{x}_{i,t-\ell} + u_{it}, \tag{6}$$

$$u_{it} = \boldsymbol{\gamma}'_i \mathbf{f}_t + \varepsilon_{it}, \tag{7}$$

for $i = 1, 2, ..., N$ and $t = 1, 2, ..., T$, where $\mathbf{f}_t$ is an $m \times 1$ vector of unobserved common factors, and p_{yi} and p_{xi} are the lag orders chosen to be sufficiently long so that u_{it} is a serially uncorrelated process across all i. The vector of long-run coefficients is then given by

$$\boldsymbol{\theta}_i = \frac{\sum_{\ell=0}^{p_{xi}} \boldsymbol{\beta}_{i\ell}}{1 - \sum_{\ell=1}^{p_{yi}} \varphi_{i\ell}}. \tag{8}$$

There are two approaches to estimating the long-run coefficients. One approach, considered in the literature, is to estimate the individual short-run coefficients $\{\varphi_{i\ell}\}$ and $\{\boldsymbol{\beta}_{i\ell}\}$ in the ARDL relation, (6), and then compute the estimates of long-run effects using formula (8) with the short-run coefficients replaced by their estimates $\{\hat{\varphi}_{i\ell}\}$ and $\left\{\hat{\boldsymbol{\beta}}_{i\ell}\right\}$. We shall refer to this approach as the "ARDL approach to the estimation of long-run effects". The advantage of this approach is that the estimates of short-run coefficients are also obtained. But when the focus is on the long-run then, under certain conditions to be clarified below, an alternative approach proposed in this paper can be undertaken to estimate $\boldsymbol{\theta}_i$ *directly*. This is possible by observing that the ARDL model, (6), can be written as

$$y_{it} = \boldsymbol{\theta}_i \mathbf{x}_{it} + \boldsymbol{\alpha}'_i (L) \Delta \mathbf{x}_{it} + \tilde{u}_{it}, \tag{9}$$

where $\tilde{u}_{it} = \varphi(L)^{-1} u_{it}$, $\varphi_i(L) = 1 - \sum_{\ell=1}^{p_{yi}} \varphi_{i\ell} L^\ell$, $\boldsymbol{\theta}_i = \boldsymbol{\delta}_i(1)$, $\boldsymbol{\delta}_i(L) = \varphi_i^{-1}(L) \boldsymbol{\beta}_i(L) = \sum_{\ell=0}^{\infty} \boldsymbol{\delta}_{i\ell} L^\ell$, $\boldsymbol{\beta}_i(L) = \sum_{\ell=0}^{p_{xi}} \boldsymbol{\beta}_{i\ell} L^\ell$, and $\boldsymbol{\alpha}_i(L) = \sum_{\ell=0}^{\infty} \sum_{s=\ell+1}^{\infty} \boldsymbol{\delta}_s L^\ell$. We shall refer to the estimation of $\boldsymbol{\theta}_i$ based on the distributed lag representation (9) as the "distributed lag (DL) approach to the estimation of long-run effects". Under the usual assumptions on the roots of $\varphi_i(L)$ falling strictly outside the unit circle, then the coefficients of $\boldsymbol{\alpha}_i(L)$ are exponentially decaying; and it is possible to show that, in the absence of feedback effects from lagged values of y_{it} onto the regressors $\mathbf{x}_{it}$, a consistent estimate of $\boldsymbol{\theta}_i$ can be obtained directly based on the least squares regression of y_{it} on $\mathbf{x}_{it}$ and $\{\Delta \mathbf{x}_{it-\ell}\}_{\ell=0}^{p}$, where the truncation lag order p is chosen appropriately as an increasing function of the sample size. But, when the feedback effects from the lagged values of the dependent variable to the regressors are present, $\tilde{u}_{it}$ will be correlated with $\mathbf{x}_{it}$ and the DL approach would no longer be consistent. Note that strict exogeneity is, however, not necessarily required for the consistency of the DL approach, since arbitrary correlations amongst the individual reduced form innovations in $\mathbf{e}_t$ are still allowed. After the individual estimates $\hat{\boldsymbol{\theta}}_i$ are obtained, either using ARDL or DL approach, they can then be averaged across i to obtain a consistent estimate of the average long-run effects, given by $\bar{\hat{\boldsymbol{\theta}}} = N^{-1}\Sigma_i^N \hat{\boldsymbol{\theta}}_i$.

2.2 Pros and cons of the two approaches to the estimation of long-run effects

Consider first the ARDL approach, where the estimates of long-run effects are computed based on the estimates of the short-run coefficients in (6). In the case where the unobserved common factors are serially uncorrelated and are also uncorrelated with the regressors, the long-run coefficients can be estimated consistently from the Ordinary Least Squares (OLS) estimates of the short-run coefficients, irrespective of whether the regressors are strictly exogenous or jointly determined with y_{it}, in the sense that $\mathbf{z}_{it} = (y_{it}, \mathbf{x}'_{it})'$ follows a VAR model. The long-run estimates are also consistent irrespective of whether the underlying variables are integrated of order one, $I(1)$ for short, or integrated of order zero, $I(0)$. These robustness properties are clearly important in empirical research. However, the ARDL approach has also a number of drawbacks. The sampling uncertainty could be large especially when the speed of convergence towards the long-run relation is rather slow and the time dimension is not sufficiently long. This is readily apparent from (8) since even a small change to $1 - \sum_{\ell=1}^{p_{yi}} \hat{\varphi}_{i\ell}$ could have large impact on the estimates of $\boldsymbol{\theta}_i$ when $\sum_{\ell=1}^{p_{yi}} \hat{\varphi}_{i\ell}$ is close to unity. In this respect, a correct specification of lag orders could be quite important for the performance of the ARDL estimates of $\boldsymbol{\theta}_i$. Underestimating the lag orders leads to inconsistent estimates, whilst overestimating the lag orders could result in loss of efficiency and low power when the ARDL long-run estimates are used for inference.

In the more general case when the unobserved common factors are correlated with the regressors then LS estimation of ARDL model is no longer consistent and the effects of unobserved common factors need to be taken into account. There are so far two possible estimators developed in the literature for this case:[5] a principal-components based approach by Song (2013) who extends the interactive effects estimator originally proposed Bai (2009) to dynamic heterogeneous panels, and the dynamic common correlated effects mean group estimator suggested by Chudik and Pesaran (2013a). A recent overview of these methods is provided in Chudik and Pesaran (2013b). These estimators have (so far) been proposed only for stationary panels, and are subject to the small T bias of the ARDL approach discussed above. Bias correction techniques can also be used, but overall they do not seem to be effective when the speed of adjustment to the steady state is slow.[6]

The main merits of the DL approach that we develop below is that, once (9) is appro-

[5] Related is also the quasi maximum likelihood estimator for dynamic panels by Moon and Weidner (2010), but this estimators has been developed only for homogeneous panels.

[6] Chudik and Pesaran (2013a) consider the application of two bias correction procedures to dynamic CCE type estimators, but find that they do not fully eliminate the bias.

priately augmented by cross-section averages, it is robust along a number of dimensions that are important in practice and it tends to show better small sample performance when the time dimension T is not very large. This includes robustness to the possibility of unit roots in regressors and/or factors, heterogeneity or homogeneity of short and/or long-run coefficients, arbitrary serial correlation in ε_{it} and $\mathbf{f}_t$ (note that $\boldsymbol{\theta}_i$ is identified even when ε_{it} is serially correlated), number of unobserved common factors (subject to certain conditions), and weak cross-sectional dependence in the idiosyncratic errors, ε_{it}. These are very important considerations in applied work. In addition, the CS-DL approach does not require specifying the individual lag orders, p_{yi} and p_{xi}, and is robust to possible breaks in ε_{it}. The main drawback of the CS-DL approach, however, is that $\tilde{u}_{it} = \varphi(L)^{-1} u_{it}$ is correlated with $\mathbf{x}_{it}$ when there are feedback effects from lagged values of y_{it} onto the regressors, $\mathbf{x}_{it}$. This correlation in turn introduces a bias that will not vanish as the sample size increase and therefore the CS-DL estimation of the long-run effects is consistent only in the case when the feedback effects (or reverse causality) are not present. The second drawback is that the small sample performance is very good only when the eigenvalues of $\varphi(L)$ are not close to the unit circle. We will provide small sample evidence on the two approaches by means of Monte Carlo experiments in Section 4.

3 Cross section augmented distributed lag (CS-DL) approach to estimation of mean long-run coefficients

3.1 The ARDL panel data model

Suppose y_{it} is generated according to the panel ARDL data model (6) with $p_{yi} = 1$ and $p_{xi} = 0$,

$$y_{it} = \varphi_i y_{i,t-1} + \boldsymbol{\beta}_i' \mathbf{x}_{it} + \boldsymbol{\gamma}_i' \mathbf{f}_t + \varepsilon_{it}, \tag{10}$$

for $i = 1, 2, ..., N$ and $t = 1, 2, ..., T$. To allow for correlation between the m unobserved factors, $\mathbf{f}_t$, and the k observed regressors, $\mathbf{x}_{it}$, suppose that the latter is generated according to the following canonical factor model

$$\mathbf{x}_{it} = \boldsymbol{\Gamma}_i' \mathbf{f}_t + \mathbf{v}_{it}, \tag{11}$$

for $i = 1, 2, ..., N$ and $t = 1, 2, ..., T$, where $\boldsymbol{\Gamma}_i$ is $m \times k$ matrix of factor loadings, and $\mathbf{v}_{it}$ are the idiosyncratic components of $\mathbf{x}_{it}$ which are assumed to be distributed independently of the idiosyncratic errors, ε_{it}. The panel data model (10) and (11) is identical to the model considered by Pesaran (2006) with the exception that the lagged dependent variable is included in (10). We have also omitted observed common effects and deterministics (such as intercepts and time trends) from (10) to simplify the exposition. Introducing these terms and additional lags of the dependent variable and regressors is relatively straightforward.

We are interested in the estimation of the mean long-run coefficients $\boldsymbol{\theta} = E(\boldsymbol{\theta}_i)$, where $\boldsymbol{\theta}_i$, $i = 1, 2, ..., N$ are the cross section specific long-run coefficients defined by (8), which for $p_{yi} = 1$ and $p_{xi} = 0$ reduces to

$$\boldsymbol{\theta}_i = \frac{\boldsymbol{\beta}_i}{1 - \varphi_i}. \tag{12}$$

We postulate the following assumptions.

Assumption 1 *(Individual Specific Errors) Individual specific errors* ε_{it} *and* $\mathbf{v}_{jt'}$ *are independently distributed for all* i, j, t *and* t'. ε_{it} *follows a linear stationary process with absolute summable autocovariances (uniformly in* i*),*

$$\varepsilon_{it} = \sum_{\ell=0}^{\infty} \alpha_{\varepsilon i\ell} \zeta_{i,t-\ell}, \tag{13}$$

for $i = 1, 2, ..., N$*, where the vector of innovations* $\boldsymbol{\zeta}_t = (\zeta_{1t}, \zeta_{2t}, ..., \zeta_{Nt})'$ *is spatially correlated according to*

$$\boldsymbol{\zeta}_t = \mathbf{R} \boldsymbol{\varsigma}_t,$$

in which the elements of $\boldsymbol{\varsigma}_t$ *are independently and identically distributed (IID) with mean zero, unit variance and finite fourth-order cumulants and the matrix* $\mathbf{R}$ *has bounded row and column matrix norms, namely* $\|\mathbf{R}\|_\infty < K$ *and* $\|\mathbf{R}\|_1 < K$*. In particular,*

$$Var(\varepsilon_{it}) = \sum_{\ell=0}^{\infty} \alpha_{\varepsilon i\ell}^2 \sigma_{\zeta i}^2 = \sigma_i^2 \leq K < \infty, \tag{14}$$

for $i = 1, 2, ..., N$*, where* $\sigma_{\zeta i}^2 = Var(\zeta_{it})$. $\mathbf{v}_{it}$ *follows a linear stationary process with absolute summable autocovariances uniformly in* i*,*

$$\mathbf{v}_{it} = \sum_{\ell=0}^{\infty} \mathbf{S}_{i\ell} \boldsymbol{\nu}_{i,t-\ell}, \tag{15}$$

for $i = 1, 2, ..., N$*, where* $\boldsymbol{\nu}_{it}$ *is* $k \times 1$ *vector of IID random variables, with mean zero, variance matrix* $\mathbf{I}_k$ *and finite fourth-order cumulants. In particular,*

$$\|Var(\mathbf{v}_{it})\| = \left\| \sum_{\ell=0}^{\infty} \mathbf{S}_{i\ell} \mathbf{S}_{i\ell}' \right\| = \|\boldsymbol{\Sigma}_i\| \leq K < \infty, \tag{16}$$

for $i = 1, 2, ..., N$*, where* $\|\mathbf{A}\|$ *is the spectral norm of the matrix* $\mathbf{A}$.

Assumption 2 *(Common Effects) The* $m \times 1$ *vector of unobserved common factors,* $\mathbf{f}_t = (f_{1t}, f_{2t}, ..., f_{mt})$*, is covariance stationary with absolute summable autocovariances, distributed independently of* $\varsigma_{it'}$ *and* $\mathbf{v}_{it'}$ *for all* i, t *and* t'*. Fourth moments of* $f_{\ell t}$*, for* $\ell = 1, 2, ..., m$*, are bounded.*

Assumption 3 *(Factor Loadings) Factor loadings $\boldsymbol{\gamma}_i$, and $\boldsymbol{\Gamma}_i$ are independently and identically distributed across i, and of the common factors $\mathbf{f}_t$, for all i and t, with fixed mean $\boldsymbol{\gamma}$ and $\boldsymbol{\Gamma}$, respectively, and bounded second moments. In particular,*

$$\boldsymbol{\gamma}_i = \boldsymbol{\gamma} + \boldsymbol{\eta}_{\gamma i},\ \boldsymbol{\eta}_{\gamma i} \sim IID\left(\underset{m\times 1}{\mathbf{0}}, \boldsymbol{\Omega}_\gamma\right),\ \text{for } i = 1, 2, ..., N,$$

and

$$vec\left(\boldsymbol{\Gamma}_i\right) = vec\left(\boldsymbol{\Gamma}\right) + \boldsymbol{\eta}_{\Gamma i},\ \boldsymbol{\eta}_{\Gamma i} \sim IID\left(\underset{km\times 1}{\mathbf{0}}, \boldsymbol{\Omega}_\Gamma\right),\ \text{for } i = 1, 2, ..., N,$$

where $\boldsymbol{\Omega}_\gamma$ and $\boldsymbol{\Omega}_\Gamma$ are $m\times m$ and $k\,m \times k\,m$ symmetric nonnegative definite matrices, $\|\boldsymbol{\gamma}\| < K$, $\|\boldsymbol{\Omega}_\gamma\| < K$, $\|\boldsymbol{\Gamma}\| < K$, and $\|\boldsymbol{\Omega}_\Gamma\| < K$.

Assumption 4 *(Coefficients) The level coefficients $\boldsymbol{\theta}_i$, defined in (12), follow the random coefficient model*

$$\boldsymbol{\theta}_i = \boldsymbol{\theta} + \boldsymbol{\upsilon}_i,\ \boldsymbol{\upsilon}_i \sim IID\left(\underset{k\times 1}{\mathbf{0}}, \boldsymbol{\Omega}_\theta\right),\ \text{for } i = 1, 2, ..., N, \tag{17}$$

where $\|\boldsymbol{\theta}\| < K$, $\|\boldsymbol{\Omega}_\theta\| < K$, $\boldsymbol{\Omega}_\theta$ is $k\times k$ symmetric nonnegative definite matrix, and the random deviations $\boldsymbol{\upsilon}_i$ are independently distributed of $\boldsymbol{\gamma}_j$, $\boldsymbol{\Gamma}_j$, ς_{jt}, $\mathbf{v}_{jt}$, and $\mathbf{f}_t$ for all i,j, and t. The coefficients φ_i are distributed with a support strictly inside the unit circle.

The polynomial $1 - \varphi_i L$ is invertible under Assumption 4, and multiplying (10) by $\left(1-\varphi_i L\right)^{-1}$ we obtain

$$\begin{aligned} y_{it} &= \left(1-\varphi_i L\right)^{-1} \boldsymbol{\beta}_i' \mathbf{x}_{it} + \left(1-\varphi_i L\right)^{-1} \boldsymbol{\gamma}_i' \mathbf{f}_t + \left(1-\varphi_i L\right)^{-1} \varepsilon_{it} \\ &= \boldsymbol{\theta}_i \mathbf{x}_{it} - \boldsymbol{\alpha}_i'\left(L\right) \Delta \mathbf{x}_{it} + \boldsymbol{\gamma}_i' \tilde{\mathbf{f}}_{it} + \tilde{\varepsilon}_{it}, \text{ for } i = 1, 2, ..., N, \end{aligned} \tag{18}$$

where $\Delta \mathbf{x}_{it} = \mathbf{x}_{it} - \mathbf{x}_{i,t-1}$, $\boldsymbol{\alpha}_i\left(L\right) = \sum_{\ell=0}^{\infty} \varphi_i^{\ell+1} \left(1-\varphi_i\right)^{-1} \boldsymbol{\beta}_i L^\ell$, $\tilde{\mathbf{f}}_{it} = \left(1-\varphi_i L\right)^{-1} \mathbf{f}_t$ and $\tilde{\varepsilon}_{it} = \left(1-\varphi_i L\right)^{-1} \varepsilon_{it}$. The distributed lag specification in (18) does not include lagged values of the dependent variable, and as a result the CCE estimation procedure can be applied to (18) directly. The level regression of y_{it} on $\mathbf{x}_{it}$ is estimated by augmenting the individual regressions by differences of unit specific regressors $\mathbf{x}_{it}$ and their lags, in addition to the augmentation by the cross section averages that take care of the effects of unobserved common factors.

Let $\mathbf{w} = \left(w_1, w_2, ..., w_N\right)'$ be an $N\times 1$ vector of weights that satisfies the following 'granularity' conditions

$$\|\mathbf{w}\| = O\left(N^{-\frac{1}{2}}\right), \tag{19}$$

$$\frac{w_i}{\|\mathbf{w}\|} = O\left(N^{-\frac{1}{2}}\right) \text{ uniformly in } i, \tag{20}$$

and the normalization condition

$$\sum_{i=1}^{N} w_i = 1. \tag{21}$$

Define the cross section averages $\bar{\mathbf{z}}_{wt} = (y_{wt}, \bar{\mathbf{x}}'_{wt})' = \sum_{i=1}^{N} w_i \mathbf{z}_{it}$, and consider augmenting the regressions of y_{it} on $\mathbf{x}_{it}$ and the current and lagged values of $\Delta \mathbf{x}_{it}$, with the following set of cross section averages, $\mathcal{S}_{Npt} = \bar{\mathbf{z}}_{wt} \cup \{\Delta \bar{\mathbf{x}}_{w,t-\ell}\}_{\ell=0}^{p}$. Cross section averages approximate the unobserved common factors arbitrarily well if

$$\vartheta_{fNp} = \mathbf{f}_t - E\left(\mathbf{f}_t | \mathcal{S}_{Npt}\right) \xrightarrow{p} \mathbf{0}, \tag{22}$$

uniformly in t, as N and $p \xrightarrow{j} \infty$. Sufficient conditions for result (22) to hold are given by Assumptions 1-4 and if the rank condition $rank\left(\mathbf{\Gamma}\right) = m$ holds. Different sets of cross section averages could also be considered. For example, if the set of cross section averages is defined as $\mathcal{S}_{Np_{\bar{z}}t} = \{\bar{\mathbf{z}}_{wt-\ell}\}_{\ell=0}^{p_{\bar{z}}}$, then the sufficient condition for (22) to hold under Assumption 1-4 would be the usual rank condition $rank\left(\mathbf{C}\right) = m$, where $\mathbf{C} = (\boldsymbol{\gamma}, \mathbf{\Gamma})$. Using covariates to enlarge the set of cross section averages could also be considered, as in Chudik and Pesaran (2013a). Theses rank conditions can be relaxed in the case $\boldsymbol{\gamma}_i$ and $\mathbf{\Gamma}_i$ are independently distributed.[7] In this case the asymptotic variance of the CCE estimators does depend on the rank condition, nevertheless the CS-DL estimators are consistent and the proposed non-parametric estimators of the covariance matrix of the CS-DL estimators given below are also valid regardless of whether the rank condition holds.

Let us also introduce the following notations, which will prove useful for setting up of the proposed estimators. Let $\mathbf{y}_i = (y_{i,p+1}, y_{i,p+2}, ..., y_{i,T})'$, $\mathbf{X}_i = \left(\mathbf{x}_{i,p+1}, \mathbf{x}_{i,p+2}, ..., \mathbf{x}_{i,T}\right)'$, $\bar{\mathbf{Z}}_w = (\bar{\mathbf{z}}_{w,p+1}, \bar{\mathbf{z}}_{w,p+2}, ..., \bar{\mathbf{z}}_{w,T})'$,

$$\underset{(T-p)\times pk}{\Delta \mathbf{X}_{ip}} = \begin{pmatrix} \Delta \mathbf{x}'_{i,p+1} & \Delta \mathbf{x}'_{i,p} & \cdots & \Delta \mathbf{x}'_{i2} \\ \Delta \mathbf{x}'_{i,p+2} & \Delta \mathbf{x}'_{i,p+1} & \cdots & \Delta \mathbf{x}'_{i3} \\ \vdots & \vdots & & \vdots \\ \Delta \mathbf{x}'_{iT} & \Delta \mathbf{x}'_{i,T-1} & \cdots & \Delta \mathbf{x}'_{i,T-p+1} \end{pmatrix},$$

$\Delta \bar{\mathbf{X}}_{wp} = \sum_{i=1}^{N} w_i \Delta \mathbf{X}_{ip}$, $\mathbf{Q}_{wi} = \left(\bar{\mathbf{Z}}_w, \Delta \bar{\mathbf{X}}_{wp}, \Delta \mathbf{X}_{ip}\right)$, and the define the projection matrix

$$\mathbf{M}_{qi} = \mathbf{I}_{T-p} - \mathbf{Q}_{wi}\left(\mathbf{Q}'_{wi}\mathbf{Q}_{wi}\right)^{+}\mathbf{Q}'_{wi}, \tag{23}$$

for $i = 1, 2, ..., N$, where $p = p\left(T\right)$ is a chosen non-decreasing truncation lag function such that $0 \leq p < T$, and $\mathbf{A}^{+}$ is the Moore-Penrose pseudoinverse of the matrix $\mathbf{A}$. We use the Moore-Penrose pseudoinverse as opposed to standard inverse in (23) because the column vectors of $\mathbf{Q}_{wi}$ could be asymptotically (as $N \to \infty$) linearly dependent.

[7] Correlation of $\boldsymbol{\gamma}_i$ and $\mathbf{\Gamma}_i$ could introduce a bias in the rank deficient case, as noted by Sarafidis and Wansbeek (2012).

The CS-DL mean group estimator of the mean long-run coefficients is given by

$$\widehat{\boldsymbol{\theta}}_{MG}=\frac{1}{N}\sum_{i=1}^{N}\widehat{\boldsymbol{\theta}}_{i}, \tag{24}$$

where

$$\widehat{\boldsymbol{\theta}}_{i}=\left(\mathbf{X}_{i}^{\prime}\mathbf{M}_{qi}\mathbf{X}_{i}^{\prime}\right)^{-1}\mathbf{X}_{i}^{\prime}\mathbf{M}_{qi}\mathbf{y}_{i}. \tag{25}$$

The CS-DL pooled estimator of the mean long-run coefficients is

$$\widehat{\boldsymbol{\theta}}_{P}=\left(\sum_{i=1}^{N}w_{i}\mathbf{X}_{i}^{\prime}\mathbf{M}_{qi}\mathbf{X}_{i}\right)^{-1}\sum_{i=1}^{N}w_{i}\mathbf{X}_{i}^{\prime}\mathbf{M}_{qi}\mathbf{y}_{i}. \tag{26}$$

Estimators $\widehat{\boldsymbol{\theta}}_{MG}$ and $\widehat{\boldsymbol{\theta}}_{P}$ differ from the mean group and pooled CCE estimator developed in Pesaran (2006), which only allows for the inclusion of a fixed number of regressors, whilst the CS-DL type estimators include p_T lags of $\Delta\mathbf{x}_{it}$ and their cross section averages, where p_T increases with T, albeit at a slower rate.

In addition to Assumptions 1-4 above, we shall also require the following assumption to hold. Assumption 5 below ensures that $\widehat{\boldsymbol{\theta}}_{MG}$ and $\widehat{\boldsymbol{\theta}}_{P}$ and their asymptotic distributions are well defined.

Assumption 5 *(a) The matrix* $\lim_{N,T,p\xrightarrow{j}\infty}\sum_{i=1}^{N}w_{i}\boldsymbol{\Sigma}_{i}=\boldsymbol{\Psi}^{*}$ *exists and is nonsingular, and* $\sup_{i,p}\left\Vert\boldsymbol{\Sigma}_{i}^{-1}\right\Vert<K$*, where* $\boldsymbol{\Sigma}_{i}=p\lim T^{-1}\mathbf{X}_{i}^{\prime}\mathbf{M}_{hi}\mathbf{X}_{i}$*, and* $\mathbf{M}_{hi}$ *is defined in (A.3).*

(b) Denote the t*-th row of matrix* $\widetilde{\mathbf{X}}_{i}=\mathbf{M}_{hi}\mathbf{X}_{i}$ *by* $\widetilde{\mathbf{x}}_{it}^{\prime}=\left(\widetilde{x}_{i1t},\widetilde{x}_{i2t},....,\widetilde{x}_{ikt}\right)$*. The individual elements of* $\widetilde{\mathbf{x}}_{it}$ *have uniformly bounded fourth moments, namely there exists a positive constant* $K<\infty$ *such that* $E\left(\widetilde{x}_{ist}^{4}\right)<K$*, for any* $t=1,2,...,T$*,* $i=1,2,...,N$ *and* $s=1,2,...,k$*.*

(c) There exists T_0 *such that for all* $T\geq T_{0}$*,* $\left(\sum_{i=1}^{N}w_{i}\mathbf{X}_{i}^{\prime}\mathbf{M}_{qi}\mathbf{X}_{i}/T\right)^{-1}$ *exists.*

(d) There exists N_0*,*T_0 *and* $p_{0}=p(T_{0})$ *such that for all* $N\geq N_{0}$*,* $T\geq T_{0}$ *and* $p(T)\geq p(T_{0})$*, the* $k\times k$ *matrices* $\left(\mathbf{X}_{i}^{\prime}\mathbf{M}_{qi}\mathbf{X}_{i}/T\right)^{-1}$ *exist for all* i*, uniformly.*

Our main findings are summarized in the following theorems.

Theorem 1 (Asymptotic distribution of $\widehat{\boldsymbol{\theta}}_{MG}$) *Suppose* y_{it}*, for* $i=1,2,...,N$ *and* $t=1,2,...,T$ *is given by the panel data model (10)-(11), Assumptions 1-5 hold, and* $(N,T,p(T))\xrightarrow{j}\infty$ *such that* $\sqrt{N}p(T)\rho^{p}\rightarrow 0$*, for any constant* $0<\rho<1$ *and* $p(T)^{3}/T\rightarrow\varkappa$*,* $0<\varkappa<\infty$*. Then, if rank* $(\boldsymbol{\Gamma})=m$ *we have*

$$\sqrt{N}\left(\widehat{\boldsymbol{\theta}}_{MG}-\boldsymbol{\theta}\right)\xrightarrow{d}N\left(0,\boldsymbol{\Omega}_{\theta}\right), \tag{27}$$

where $\mathbf{\Omega}_\theta = Var(\boldsymbol{\theta}_i)$ and $\widehat{\boldsymbol{\theta}}_{MG}$ is given by (24). If rank $(\mathbf{\Gamma}) \neq m$ and $\boldsymbol{\gamma}_i$ is independently distributed of $\mathbf{\Gamma}_i$, we have

$$\sqrt{N}\left(\widehat{\boldsymbol{\theta}}_{MG} - \boldsymbol{\theta}\right) \xrightarrow{d} N\left(0, \mathbf{\Sigma}_{MG}\right), \tag{28}$$

where

$$\mathbf{\Sigma}_{MG} = \mathbf{\Omega}_\theta + \lim_{p,N\to\infty}\left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{\Sigma}_i^{-1}\mathbf{Q}_{if}\mathbf{\Omega}_\gamma\mathbf{Q}_{if}'\mathbf{\Sigma}_i^{-1}\right], \tag{29}$$

in which $\mathbf{\Omega}_\gamma = Var(\boldsymbol{\gamma}_i)$, $\mathbf{\Sigma}_i = p\lim_{T\to\infty} T^{-1}\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{X}_i$ and $\mathbf{Q}_{if} = p\lim_{T\to\infty} T^{-1}\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{F}$. In both cases, the asymptotic variance of $\widehat{\boldsymbol{\theta}}_{MG}$ can be consistently estimated nonparametrically by

$$\widehat{\mathbf{\Sigma}}_{MG} = \frac{1}{N-1}\sum_{i=1}^{N}\left(\widehat{\boldsymbol{\theta}}_i - \widehat{\boldsymbol{\theta}}_{MG}\right)\left(\widehat{\boldsymbol{\theta}}_i - \widehat{\boldsymbol{\theta}}_{MG}\right)'. \tag{30}$$

Theorem 2 (Asymptotic distribution of $\widehat{\boldsymbol{\theta}}_P$) *Suppose y_{it}, for $i = 1, 2, ..., N$ and $t = 1, 2, ..., T$ are generated by the panel data model (10)-(11), Assumptions 1-5 hold, and $(N, T, p(T)) \xrightarrow{j} \infty$, such that $\sqrt{N}p(T)\rho^p \to 0$, for any constant $0 < \rho < 1$ and $p(T)^3/T \to \varkappa$, $0 < \varkappa < \infty$. Then, if $\boldsymbol{\gamma}_i$ is independently distributed of $\mathbf{\Gamma}_i$, we have*

$$\left(\sum_{i=1}^{N} w_i^2\right)^{-1/2}\left(\widehat{\boldsymbol{\theta}}_P - \boldsymbol{\theta}\right) \xrightarrow{d} N\left(\mathbf{0}, \mathbf{\Sigma}_P\right), \tag{31}$$

where $\widehat{\boldsymbol{\theta}}_P$ is given by (26),

$$\mathbf{\Sigma}_P = \mathbf{\Psi}^{*-1}\mathbf{R}^*\mathbf{\Psi}^{*-1},\ \mathbf{\Psi}^* = \lim_{N\to\infty}\sum_{i=1}^{N} w_i\mathbf{\Sigma}_i, \tag{32}$$

$$\mathbf{R}^* = \mathbf{R}_\theta^* + \mathbf{R}_\gamma^*,\ \mathbf{R}_\theta^* = \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\widetilde{w}_i^2\mathbf{\Sigma}_i\mathbf{\Omega}_\theta\mathbf{\Sigma}_i,\ \mathbf{R}_\gamma^* = \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\widetilde{w}_i^2\mathbf{Q}_{if}\mathbf{\Omega}_\gamma\mathbf{Q}_{if}',$$

$\mathbf{\Omega}_\theta = Var(\boldsymbol{\theta}_i)$, $\mathbf{\Omega}_\gamma = Var(\boldsymbol{\gamma}_i)$, $\mathbf{\Sigma}_i = p\lim T^{-1}\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{X}_i$, $\mathbf{Q}_{if} = p\lim T^{-1}\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{F}$, and $\widetilde{w}_i = \sqrt{N}w_i\left(\sum_{i=1}^{N} w_i^2\right)^{-1/2}$. If rank $(\mathbf{\Gamma}) = m$, then $\boldsymbol{\gamma}_i$ is no longer required to be independently distributed of $\mathbf{\Gamma}_i$ and (31) continues to hold with $\mathbf{\Sigma}_P = \mathbf{\Psi}^{*-1}\mathbf{R}_\theta^*\mathbf{\Psi}^{*-1}$. In both cases, $\mathbf{\Sigma}_P$ can be consistently estimated by $\hat{\mathbf{\Sigma}}_P$ defined by equation (A.25) in the Appendix.

Theorems 1-2 establish asymptotic distribution of $\widehat{\boldsymbol{\theta}}_{MG}$ and $\widehat{\boldsymbol{\theta}}_P$ under slope heterogeneity. These theorems distinguish between cases where the rank condition that ensures (22) is satisfied or not. In the former case, unobserved common factors can be approximated by cross section averages when N is large and regardless of whether $\boldsymbol{\gamma}_i$ is correlated with $\mathbf{\Gamma}_i$, $\widehat{\boldsymbol{\theta}}_{MG}$ and $\widehat{\boldsymbol{\theta}}_P$ are consistent and asymptotically normal. In the latter case, where the unobserved common factors cannot be approximated by cross section averages when N is large, then so long as $\boldsymbol{\gamma}_i$ and $\boldsymbol{\Gamma}_i$ are independently distributed, both $\widehat{\boldsymbol{\theta}}_{MG}$ and $\widehat{\boldsymbol{\theta}}_P$ continue to be consistent and asymptotically normal, but the asymptotic variance depends also on unobserved common factors and their loadings. In both (full rank or rank deficient) cases, the asymptotic variance of the CS-DL estimators can be estimated consistently using the same non-parametric formulae as in the full rank case.

There are several departures from the assumptions of these theorems that might be of interest in applied work, such as the consequences of breaks in the error processes, ε_{it}, possibility of unit roots in factors and/or regressor specific components, and situations where some or all coefficients are homogeneous over the cross-section units. These theoretical extensions are outside the scope of the present paper but we investigate the robustness of the proposed CS-DL estimator to such departures by means of Monte Carlo simulations in the next section.

4 Monte Carlo experiments

This section investigates small sample properties of the CS-DL estimators and compare them with the estimates obtained from the panel ARDL approach using the dynamic CCEMG estimator of the short-run coefficients advanced in Chudik and Pesaran (2013a), which we denote by CS-ARDL. First, we present results from the baseline experiments with heterogeneous slopes (long- and short-run coefficients), and then we document small sample performance of the alternative estimators under various deviations from the baseline experiments, including robustness of the estimators to the introduction of unit roots in the regressors or factors, possible breaks in the idiosyncratic error processes, and the consequences of feedback effects from lagged values of y_{it} onto $\mathbf{x}_{it}$. Second, we investigate whether it is possible to improve on the estimation of short-run coefficients, provided the model is correctly specified, by imposing CS-DL estimates of the long-run coefficients.

We start with a brief summary of the estimation methods and a description of the data generating processes. Then we present findings on the estimation of mean long-run coefficient and on the extent to which estimates of the short-run coefficients can be improved by using the CS-DL estimators of the long-run effects.

4.1 Estimation methods

The CS-DL estimators are based on the following auxiliary regressions:

$$y_{it} = c_{yi} + \boldsymbol{\theta}_i' \mathbf{x}_{it} + \sum_{\ell=0}^{p-1} \delta_{i\ell} \Delta x_{i,t-\ell} + \sum_{\ell=0}^{p_{\bar{y}}} \omega_{y,i\ell} \bar{y}_{t-\ell} + \sum_{\ell=0}^{p_{\bar{x}}} \boldsymbol{\omega}_{x,i\ell}' \bar{\mathbf{x}}_{t-\ell} + e_{it}, \tag{33}$$

where $\bar{\mathbf{x}}_t = N^{-1} \sum_{i=1}^{N} \mathbf{x}_{it}$, $\bar{y}_t = N^{-1} \sum_{i=1}^{N} y_{it}$, $p_{\bar{x}}$ is set equal to the integer part of $T^{1/3}$, denoted as $\left[T^{1/3}\right]$, $p = p_{\bar{x}}$ and $p_{\bar{y}}$ is set to 0. We consider both CS-DL mean group and pooled estimators based on (33).

The CS-ARDL estimator is based on the following regressions:

$$y_{it} = c_{yi}^{*} + \sum_{\ell=1}^{p_y} \varphi_{i\ell} y_{i,t-\ell} + \sum_{\ell=0}^{p_x} \boldsymbol{\beta}_{i\ell}' \mathbf{x}_{i,t-\ell} + \sum_{\ell=0}^{p_{\bar{z}}} \boldsymbol{\psi}_{i\ell}' \bar{\mathbf{z}}_{t-\ell} + e_{it}^{*}, \tag{34}$$

where $\bar{\mathbf{z}}_t = (\bar{y}_t, \bar{\mathbf{x}}_t')'$, $p_{\bar{z}} = \left[T^{1/3}\right]$ and two options for the remaining lag orders are considered: ARDL(2,1) specification, $p_y = 2$ and $p_x = 1$, and ARDL(1,0) specification, $p_y = 1$ and $p_x = 0$. The CS-ARDL estimates of individual mean level coefficient are then given by

$$\hat{\boldsymbol{\theta}}_{CS-ARDL,i} = \frac{\sum_{\ell=0}^{p_x} \hat{\boldsymbol{\beta}}_{i\ell}}{1 - \sum_{\ell=1}^{p_y} \hat{\varphi}_{i\ell}}, \tag{35}$$

where the estimates of short run coefficients ($\hat{\varphi}_{i\ell}$,$\hat{\boldsymbol{\beta}}_{i\ell}$) are based on (34). The mean long-run effects are estimated as $N^{-1} \sum_{i=1}^{N} \hat{\boldsymbol{\theta}}_{CS-ARDL,i}$ and the inference is based on the usual non-parametric estimator of asymptotic variance of the mean group estimator.

4.2 Data generating process

The dependent variable and regressors are generated from the following ARDL(2,1) panel data model with factor error structure,

$$y_{it} = c_{yi} + \varphi_{i1} y_{i,t-1} + \varphi_{i2} y_{i,t-2} + \beta_{i0} x_{it} + \beta_{i1} x_{i,t-1} + u_{it},\ u_{it} = \boldsymbol{\gamma}_i' \mathbf{f}_t + \varepsilon_{it}, \tag{36}$$

and

$$x_{it} = c_{xi} + \kappa_{yi} y_{i,t-1} + \boldsymbol{\gamma}_{xi}' \mathbf{f}_t + v_{it}. \tag{37}$$

We generate $y_{it}, \mathbf{x}_{it}$ for $i = 1, 2, ..., N$, and $t = -99, ..., 0, 1, 2, ..., T$ with the starting values $y_{i,-101} = y_{i,-100} = 0$, and the first 100 time observations ($t = -99, -48, ..., 0$) are discarded to reduce the effects of the initial values on the outcomes. The fixed effects are generated as $c_{iy} \sim IIDN\,(1, 1)$, and $c_{xi} = c_{yi} + \varsigma_{c_x i}$, where $\varsigma_{c_x i} \sim IIDN\,(0, 1)$, thus allowing for dependence between x_{it} and c_{yi}.

We consider three cases depending on the heterogeneity/homogeneity of the slopes:

- (heterogeneous slopes - baseline) $\varphi_{i1} = (1 + \varkappa_{\varphi i})\, \eta_{\varphi i}$, $\varphi_{i2} = -\varkappa_{\varphi i} \eta_{\varphi i}$, $\varkappa_{\varphi i} \sim IIDU\,(0.2, 0.3)$, $\eta_{\varphi i} \sim IIDU\,(0, \varphi_{\max})$. The long-run coefficients are generated as $\theta_i \sim IIDN\,(1, 0.2^2)$ and the regression coefficient are generated as $\beta_{i0} = \varkappa_{\beta i} \eta_{\beta i}$, $\beta_{i1} = (1 - \varkappa_{\beta i})\, \eta_{\beta i}$, where $\eta_{\beta i} = \theta_i / \left(1 - \varphi_{i1} - \varphi_{i2}\right)$ and $\varkappa_{\beta i} \sim IIDU\,(0, 1)$.

- (homogeneous long-run, heterogenous short-run slopes) $\theta_i = 1$ for all i and the remaining coefficients $(\varphi_{i1}, \varphi_{i2}, \beta_{i0}, \beta_{i1})$ are generated as in the previous fully heterogeneous case.

- (homogeneous long- and short-run slopes) $\varphi_{i1} = 1.15\varphi_{\max}/2$, $\varphi_{i2} = -0.15\varphi_{\max}/2$, $\theta_i = 1$, and $\beta_{i0} = \beta_{i1} = 0.5/\left(1 - \varphi_{\max}/2\right)$.

We also consider the case of ARDL(1,0) panel model by setting $\varkappa_{\varphi i} = 0$ and $\varkappa_{\beta i} = 1$ for all i, which gives $\varphi_{i2} = \beta_{i1} = 0$ for all i. We consider three values for $\varphi_{\max} = 0.6$, 0.8 or 0.9.

The unobserved common factors in $\mathbf{f}_t$ and the unit-specific components, v_{it}, are generated as independent AR(1) processes:

$$f_{t\ell} = \rho_{f\ell} f_{t-1,\ell} + \varsigma_{ft\ell},\ \varsigma_{ft\ell} \sim IIDN\left(0, \sigma^2_{\varsigma f\ell}\right), \tag{38}$$

$$v_{it} = \rho_{xi} v_{i,t-1} + \nu_{it},\ \varsigma_{xit} \sim IIDN\left(0, \sigma^2_{\nu i}\right), \tag{39}$$

for $i = 1, 2, ..., N$, $\ell = 1, 2, .., m$, and for $t = -99, ..., 0, 1, 2, ..., T$ with the starting values $f_{\ell,-100} = 0$, and $v_{i,-100} = 0$. The first 100 time observations ($t = -99, -48, ..., 0$) are discarded. We consider three possibilities for the AR(1) coefficients $\rho_{f\ell}$ and ρ_{xi}:

- (stationary baseline) $\rho_{xi} \sim IIDU\left[0.0.95\right]$, $\sigma^2_{\nu i} = 1 - \rho^2_{xi}$ for all i; $\rho_{f\ell} = 0.6$, and $\sigma^2_{\varsigma f\ell} = 1 - \rho^2_{f\ell}$ for $\ell = 1, 2, ..., m$.

- (nonstationary factors) $\rho_{xi} \sim IIDU\left[0.0.95\right]$, $\sigma^2_{\nu i} = 1 - \rho^2_{xi}$ for all i; and $\rho_{f\ell} = 1$, $\sigma^2_{\varsigma f\ell} = 0.1^2$ for $\ell = 1, 2, ..., m$.

- (nonstationary regressors and stationary factors) $\rho_{xi} = 1$, $\sigma^2_{\nu i} = 0.1^2$ for all i; and $\rho_{f\ell} = 0.6$, $\sigma^2_{\varsigma f\ell} = 1 - \rho^2_{f\ell}$, for $\ell = 1, 2, ..., m$.

We consider also two options for the feedback coefficients κ_{yi}: no feedback effects, $\kappa_{yi} = 0$ for all i, and with feedback effects, $\kappa_{yi} \sim IIDU\left(0, 0.2\right)$.

Factor loadings are generated as

$$\gamma_{i\ell} \sim IIDN\left(\gamma_\ell, 0.2^2\right) \text{ and } \gamma_{xi\ell} \sim IIDN\left(\gamma_{x\ell}, 0.2^2\right),$$

for $\ell = 1, 2, .., m$, and $i = 1, 2, ..., N$. Also, without loss of generality, the means of factor loadings are calibrated so that $Var(\boldsymbol{\gamma}'_i \mathbf{f}_t) = Var\left(\boldsymbol{\gamma}'_{xi}\mathbf{f}_t\right) = 1$ in the stationary case. We set $\gamma_\ell = \sqrt{b_\gamma}$, and $\gamma_{x\ell} = \sqrt{\ell b_x}$, for $\ell = 1, 2, ..., m$, where $b_\gamma = 1/m - 0.2^2$, and $b_x = 2/\left[m\left(m+1\right)\right] - 2/\left(m+1\right)0.2^2$. This ensures that the contribution of the unobserved factors to the variance of y_{it} does not rise with m in the stationary case. We consider $m = 2$ or 3 unobserved common factors.

Finally, the idiosyncratic errors, ε_{it}, are generated to be heteroskedastic, weakly cross-sectionally dependent and serially correlated. Specifically,

$$\varepsilon_{it} = \rho_{\varepsilon i}\varepsilon_{i,t-1} + \zeta_{it}, \tag{40}$$

where $\boldsymbol{\zeta}_t = (\zeta_{1t}, \zeta_{2t}, ..., \zeta_{Nt})'$ are generated using the following spatial autoregressive model (SAR),

$$\boldsymbol{\zeta}_t = a_\epsilon \mathbf{S}_\epsilon \boldsymbol{\zeta}_t + \boldsymbol{\varsigma}_t, \tag{41}$$

in which the elements of $\boldsymbol{\varsigma}_t$ are drawn as $IIDN\left[0, \frac{1}{2}\sigma_i^2\left(1-\rho_{\varepsilon i}^2\right)\right]$, with σ_i^2 obtained as independent draws from $\chi^2(2)$ distribution,

$$\mathbf{S}_\epsilon = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & & 0 \\ 0 & \frac{1}{2} & 0 & \ddots & & \vdots \\ 0 & 0 & \ddots & \ddots & \frac{1}{2} & 0 \\ \vdots & & & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & \cdots & 0 & 1 & 0 \end{pmatrix},$$

and the spatial autoregressive parameter is set to $a_\epsilon = 0.6$. Note that $\{\epsilon_{it}\}$ is cross-sectionally weakly dependent for $|a_\epsilon| < 1$. We consider $\rho_{\varepsilon i} = 0$ for all i or $\rho_{\varepsilon i} \sim IIDU\left(0, 0.8\right)$. We also consider the possibility of breaks in ε_{it} by generating for each i random break points $b_i \in \{1, 2, ..T\}$ and

$$\begin{aligned} \varepsilon_{it} &= \rho_{\varepsilon i}^a \varepsilon_{i,t-1} + \zeta_{it}, \text{ for } t = 1, 2, ..., b_i \\ \varepsilon_{it} &= \rho_{\varepsilon i}^b \varepsilon_{i,t-1} + \zeta_{it}, \text{ for } t = b_i + 1, b_i + 2, ..., T, \end{aligned}$$

where $\rho_{\varepsilon i}^a, \rho_{\varepsilon i}^b \sim IIDU\left(0, 0.8\right)$, and $\boldsymbol{\zeta}_t = (\zeta_{1t}, \zeta_{2t}, ..., \zeta_{Nt})'$ is generated using SAR model (41) with $\varsigma_{it} \sim IIDN\left[0, \frac{1}{2}\sigma_i^2\left(1-\rho_{\varepsilon i}^{a2}\right)\right]$.

The above DGP is more general than the other DGPs used in MC experiments in the literature and allows the factors and regressors to be correlated and persistent. The above DGPs also include models with unit roots, breaks in the error processes, and allows for correlated fixed effects. To summarize, we consider the following cases:

1. (3 options for heterogeneity of coefficients) heterogeneous baseline, homogeneous long-run with heterogeneous short-run, and both long-and short-run homogeneous,

2. (2 options for lags) ARDL(2,1) baseline, and ARDL(1,0) model where $\varkappa_{\varphi i} = 0$ and $\varkappa_{\beta i} = 1$ for all i, which gives $\varphi_{i2} = \beta_{i1} = 0$ for all i.

3. (3 options for $\varphi_{\max}$) $\varphi_{\max} = 0.6$ (baseline), 0.8, or 0.9

4. (3 options for the persistence of factors and regressors) stationary baseline, I(1) factors, or I(1) regressor specific components v_{it},

5. (2 options for the number of factors) full rank case baseline $m = 2$, or rank deficient case $m = 3$,

6. (3 options for the persistence of idiosyncratic errors) serially uncorrelated baseline $\rho_{\varepsilon i}=0$, $\rho_{\varepsilon i}\sim IIDU\left(0,0.8\right)$, or breaks in the error process.

7. (2 options for feedback effects) $\kappa_{yi}=0$ for all i (baseline), or $\kappa_{yi}\sim IIDU\left(0,0.2\right)$.

Due to the large number of possible cases (648 in total), we only consider baseline experiments and various departures from the baseline. We consider the following combinations of sample sizes: $N,T\in\{30,50,100,150,200\}$, and set the number of replications to $R=2,000$, in the case of all experiments.

4.3 Monte Carlo findings on the estimation of mean long-run coefficients

The results for the baseline DGP are summarized in Table 1. This table shows good performance of the CS-DL estimators in the baseline experiments. This table also shows problems with the CS-ARDL approach when T is not large (<100) due to the small sample problems arising when $\sum_{\ell=1}^{p_y}\hat{\varphi}_{i\ell}$ is close to unity. Also, CS-ARDL estimates based on misspecified lags orders are inconsistent, as to be expected.

Next, we investigate robustness of the results to different assumptions regarding slope heterogeneity. Table 2 presents findings for the experiment that depart from the baseline DGP by assuming homogeneous long-run slopes, while allowing the short-run slopes to be heterogeneous. Table 3 gives the results when both long- and short-run slopes are homogeneous. These results show that the CS-DL estimators continue to have good size and power properties in all cases.

Experiments based on the ARDL(1,0) specification (as the DGP) are summarized in Table 4. CS-DL estimators continue to perform well, showing their robustness to the underlying ARDL specification.

The effects of increasing the value of $\varphi_{\max}$ on the properties of the various estimators are summarized in Tables 5 (for $\varphi_{\max}=0.8$) and 6 (for $\varphi_{\max}=0.9$). Small sample performance of the CS-DL estimators deteriorates as $\varphi_{\max}$ moves closer to unity, as to be expected. Tables 5-6 show that the performance deteriorates substantially for values of $\varphi_{\max}$ close to unity, due to the bias that results from the truncation of lags for the first differences of regressors. It can take a large lag order for the truncation bias to be negligible when the largest eigenvalue of the dynamic specification (given by the lags of the dependent variable) is close to one. We see quite a substantial bias when $\varphi_{\max}=0.9$. Therefore, it is important that the CS-DL approach is used when the speed of convergence towards equilibrium is not too slow and/or T is sufficiently large so that biases arising from the approximation of dynamics by distributed lag functions can be controlled.

The robustness of the results to the number of unobserved factors (m) is investigated in Table 7. This table provides a summary in the case of $m=3$ factors, which represents the rank deficient case. It is interesting to note that despite the failure of the rank condition, the CS-DL estimators continue to perform well (the results are almost unchanged as compared with those in Table 1), while the CS-ARDL estimates are affected by two types of biases (the time series bias and the bias due to rank deficiency) that operate in opposite directions.

Consider now the robustness of the results to the presence of unit roots in the unobserved factors (Table 8) or in the regressors (Table 9). As can be seen the CS-DL estimators continue to perform well when factors contain unit roots. Table 9, on the other hand, shows large RMSE and low power for $T = 30$ and 50, when the idiosyncratic errors have unit roots. But, interestingly enough, the reported size is correct and biases are very small for all sample sizes.

The results in Table 10 consider the robustness of the CS-DL estimators to the problem of serial correlation in the errors, whilst those in Table 11 consider the robustness of these estimators to the breaks in the error processes. As can be seen, and as predicted by the theory, the CS-DL estimators are robust to both of these departures from the baseline scenario, whereas the CS-ARDL approach is not. Recall, that CS-ARDL approach requires that the lag orders are correctly specified, and does not allow for residual serial correlation and/or breaks in the error processes, whilst CS-DL does.

Last but not least, the consequences of feedback effects from y_{it} to the regressors, x_{it}, is documented in Table 12. This table shows that the CS-ARDL approach is consistent regardless of the feedback effects, provided that the lag orders are correctly specified, again as predicted by the theory. But a satisfactory performance (in terms of bias and size of the test) for the CS-ARDL approach requires T to be sufficiently large. On the other hand, in the presence of feedbacks, the CS-DL estimators are inconsistent and show positive bias even for T sufficiently large. But the bias due to feedback effects seem to be quite small; between -0.02 and 0.06, and the CS-DL estimators tend to outperform the CS-ARDL estimators when $T < 100$.

Given the above MC results, and considering that output growth is only moderately persistent[8], and given that the time dimension is 45 years, the CS-DL estimates are likely to provide a valuable complement to the ARDL estimates in our empirical investigation below.

4.4 Monte Carlo findings on the improvement in estimation of short-run coefficients

As a final exercise, we consider if it is possible to improve on the estimation of short-run coefficients by imposing the CS-DL estimates of the long-run, before estimating the short-run coefficients. We consider the experiment that departs from the baseline model by assuming a homogeneous long-run coefficient, whilst all the short-run slopes are heterogeneous,

[8] In our empirical application the first order autoregressive coefficient of output growth ranges from -0.53 (Morocco) to 0.65 (Japan), with mean and median of 0.274 and 0.273, respectively.

and use the ARDL(1,0) as the data generating process. More specifically, we impose the CS-DL pooled estimator of the long-run coefficient, $\hat{\theta}_P$, when estimating the short-run coefficients using the CS-ARDL approach. In particular, we estimate the following unit-specific regressions,

$$\Delta y_{it} = c_{yi}^* + \lambda_i \left(y_{i,t-1} - \hat{\theta}_P x_{it} \right) + \sum_{\ell=0}^{p_{\bar{z}}} \boldsymbol{\delta}_{i\ell}' \bar{\mathbf{z}}_{t-\ell} + \varepsilon_{it}^*, \tag{42}$$

for $i = 1, 2, ..., N$, and the resulting mean group estimator of $E(\varphi_{i1}) = 1 + E(\lambda_i)$ is denoted by

$$\tilde{\varphi}_{1,MG} = \frac{1}{N} \sum_{i=1}^{N} \tilde{\varphi}_{i1}, \;\; \tilde{\varphi}_{i1} = 1 - \tilde{\lambda}_i,$$

where $\tilde{\lambda}_i$ is the least square estimate of λ_i based on (42). The results of these experiments are summarized in Table 13. Imposing the CS-DL pooled estimator of the long-run coefficient improves the small sample properties of the short-run estimates substantially, about 80-90% reduction of the difference between the RMSE of the infeasible CS-ARDL estimator and the RMSE of the unconstrained estimator when $T = 30$.

We are now in a position to apply the various estimation techniques discussed in this paper to our central empirical question of interest, namely the relationship between inflation, debt to GDP and output growth across a panel of developed and emerging economies. But first we provide an overview of the literature so that our empirical results can be placed within the extant literature.

5 Effects of inflation and debt on economic growth: a literature review

5.1 Debt and growth

Economic theory provides mixed results on the relationship between public debt and growth. Elmendorf and Mankiw (1999) argue that profligate debt-generating fiscal policy (and high public debt) can have a negative impact on long-term growth by crowding out private investment, although it is argued that this effect is quantitatively small. The negative growth effect of public debt could be larger in the presence of policy uncertainty or expectations of future confiscation (possibly through inflation and financial repression). See, for example, Cochrane (2011a) and Cochrane (2011b). Contrary to this view, DeLong and Summers (2012) argue that hysteresis arising from recessions can lead to a situation in which expansionary fiscal policies may have positive effect on long-run growth. Krugman (1988) argues that nonlinearities and threshold effects can arise from the presence of external debt overhang, but it is not clear whether such an argument is applicable to advanced economies where the majority of debt-holders are residents. Nonlinearities may also arise if there is a turning point above which public debt suddenly becomes unsustainable - see Ghosh et al. (2013).

Overall, the predictions of the theoretical literature on the long-run effects of public debt on output growth are ambiguous, predicting negative as well as a positive effect under certain conditions. Even if we rely on theoretical models that predict a negative relationship between output growth and debt, we still need to estimate the magnitude of such effects empirically. The empirical evidence on the relationship between debt and growth until recently focussed on the role of *external debt* in *developing countries*, and so far there has been only a few studies that include evidence on the developed economies. One such study is by Reinhart and Rogoff (2010) who argue for a non-linear relationship between debt and growth. Using a sample of 20 advanced economies over the period 1946-2009, they split these countries into four groups: (i) country-years for which public debt to GDP levels were below 30 percent (low debt); (ii) country-years for which public debt to GDP levels were between 30 and 60 percent (medium debt); (iii) country-years for which public debt to GDP levels were between 60 and 90 percent (high debt); and (iv) country-years for which public debt to GDP levels were above 90 percent (very high). They calculate the median and average GDP growth rates for each group and show that there is generally a weak relationship between government debt and economic growth for countries with public debt levels below 90% of GDP. However, for countries with debt-to-GDP ratio over this threshold, they find that debt can have adverse effects on growth. They show that in the high-debt group, median growth is approximately one percentage point lower and average growth is nearly four percentage points lower as compared to the other groups. They also perform a similar exercise for 24 emerging economies over the periods 1946-2009 and 1900-2009.

The analysis of Reinhart and Rogoff (RR) has generated a considerable degree of debate in the literature. See, for example, Kumar and Woo (2010), Checherita-Westphal and Rother (2012), Eberhardt and Presbitero (2013), and Reinhart et al. (2012); who discuss the choice of debt brackets used, changes in country coverage, data frequency; econometric specification, and reverse causality going from output to debt. See also Panizza and Presbitero (2013) for a survey and additional references to the literature.

Kumar and Woo (2010) study the impact of high public debt on subsequent growth of real per capita GDP for a panel of 38 advanced and emerging market economies over the period 1970–2007. They apply a variety of *homogeneous* estimation methodologies, such as pooled OLS, fixed effects (FE) panel regression, and system GMM approach (to account for endogeneity of growth regressors), and consider a variety of possible covariates of debt and growth. They complement their analysis by a growth accounting framework which allows for an exploration of the channels (factor accumulation versus factor productivity) through which public debt may influence growth. Checherita-Westphal and Rother (2012) employ an alternative strategy to deal with simultaneous determination of public debt and growth (by using external instruments). They restrict their sample to 12 euro area countries over the period 1970-2008 and instrument the debt-to-GDP ratio of a typical country at each point in time with the average debt-to-GDP ratio of the other 11 countries in the sample during the same time period. With this strategy, the authors find a non-linear relationship between debt and growth with a threshold ranging between 90 and 100 percent of debt to GDP levels. They use fixed-effects, 2SLS and GMM techniques for estimation and employ a quadratic functional form for the growth-debt regression equation. They also analyze the channels through which public debt is likely to affect economic growth.

The above studies address a number of important modelling issues not considered by Reinhart and Rogoff, but they nevertheless employ panel data models that impose slope homogeneity and do not adequately allow for cross-sectional dependence across individual country errors. It is implicitly assumed that different countries converge to their equilibrium at the same rate, and there are no spillover effects of debt overhang from one country to another. These assumptions do not seem plausible given the diverse historical and institutional differences that exist across countries, and the increasing degree of interdependence of the economies in the global economy.

The paper which deals with some of these issues and is closest in approach to ours is by Eberhardt and Presbitero (2013), which studies the debt-growth relationship in the context of a heterogeneous panel data model covering 105 countries over the period from 1972 to 2009. However, their analysis is subject to three main problems. First, they include the capital stock along with the level of debt as the two main variables determining the level of aggregate output. Given the endogeneity of these variables, the analysis of the effects of debt on output becomes complicated since changes in debt are likely to influence interest rates and hence investment, and such indirect effects of changes in debt on the capital stock must also be taken into account (see Pesaran and Smith (2013) for a related discussion). Second, they assume the existence of long-run relations between output, capital stock and debt across all countries in their sample, without providing any empirical evidence to support it. Third, their analysis could be subject to the reverse causality problem since they only include one lagged values of the dependent variable and the regressors, and this might not be sufficient for the ARDL specification to capture the feedback effects running from output growth to debt/GDP ratio.

5.2 Inflation and growth

Economic theory provides mixed predictions on the effects of inflation on economic growth. Depending on how money is introduced into the model and the assumptions about its functions, inflation can have either positive or negative effects on real variables such as output and investment. Within a money-in-the-utility-function model, Sidrauski (1967) presents a superneutrality result where changes in the rate of money growth and inflation have no effects on steady-state capital and output. The same effect is obtained by Ireland (1994) within a cash-in-advance model where money is needed in advance to finance investment expenditures and at the same time capital accumulation affects money's role in the payments system. Tobin (1965) regards money as a substitute for capital and shows that higher inflation enhances investment and causes a higher level of output. Bayoumi and Gagnon (1996) show that a positive relationship between inflation and investment can also arise if there are distortions in the tax system. Stockman (1981) examines the implications of a cash-in-advance constraint applying to investment and argues that higher inflation decreases steady-state real-money balances and capital stock, and hence produces a reversed Tobin effect. Dornbusch and Frenkel (1973) show that the effects of inflation on real variables are ambiguous if money is introduced into the model through a transaction cost function. However, this ambiguity disappears when money is introduced as a transaction device through a shopping-time technology, Saving (1971) and Kimbrough (1986).

Gillman and Kejak (2005) surveys the *theoretical literature* on inflation and endogenous growth, and show that a broad range of models can generate a negative association between inflation and growth; see Gomme (1993) and De Gregorio (1993) among others. They also analyze whether the inflation–growth relationship is non-linear (becomes weaker as the inflation rate rises). In such models, the inflation rate affects growth because it changes the marginal product of capital, either that of physical capital (AK models), or that of human capital (AH models), or that of both in combined capital models. Considering AK and AH models, inflation acts as a tax on physical or human capital which decreases the marginal product of capital and lowers growth. The non-linearity property of the inflation-growth relationship can be explained through models that explicitly account for unemployment; see Akerlof et al. (2000). According to these models, low inflation favors both employment and productivity, resulting in higher capacity utilization, a lower output gap and, as a consequence, higher growth. Therefore, the relationship between inflation and output growth may be positive for low levels of the inflation rate.

There also exists a large empirical literature on the relationship between inflation and growth. A brief summary of these empirical findings is as follows. First, inflation could reduce growth by lowering investment and productivity. Barro (2001) provides evidence for a strongly significant negative effect of inflation on growth. Bruno and Easterly (1998) show that the inflation-growth correlation is present only when they base their cross-section regressions on annual observations, with the correlation weakening as longer term time averages are used. There is also a strong inflation-growth relation with pooled annual data. Third, the relationship between inflation and growth is highly non-linear. Khan and Senhadji (2001) find a 'threshold' rate of inflation, above which the effect is strongly significant and negative, but below which the effect is insignificant and positive. Gylfason and Herbertsson (2001) list some 17 studies for which all but one find a significant decrease in the growth rate from increasing the inflation rate from 5 to 50%; while Chari et al. (1996) review the empirical results from increasing the inflation rate from 10 to 20%, and report a significant fall in the growth rate within the interval, 0.2% to 0.7%. Roubini and Sala-i-Martin (1992) study the relationship between inflation and growth in a panel of 98 countries over 1960-1985 and find that an increase in the annual rate of inflation from 5 to 50 percent reduces per capita growth, *ceteris paribus,* by 2.2 percent per annum. Rousseau and Wachtel (2001) report a smaller but still significant negative effect of inflation on growth in their panel study of 84 countries during 1960-1995. The negative and highly non-linear inflation–growth effect is also supported in Judson and Orphanides (1999), Ghosh and Phillips (1998), and López-Villavicencio and Mignon (2011). Forth, *inflation volatility* is found to negatively affect production decisions, and hence growth; see Judson and Orphanides (1999).

The inflation-growth relationship is not robust though, due to the sample selection bias, temporal aggregation, and omission of consequential variables in levels. Trying to address these misspecifications, Ericsson et al. (2001), using 40 years of data (1953-1992), show that *output and inflation* are positively related. They find that, for most G-7 countries, annual time series of inflation and the log-level of output are cointegrated, thus rejecting the existence of a long-run relation between output growth and inflation. Following a different econometric approach, Bullard and Keating (1995), using a large sample of postwar countries, find that a permanent shock to inflation is not associated with a long-run change in real output for high inflation economies. Using instrumental variables to account for inflation–growth endogeneity bias, Gillman and Nakov (2004) show that the negative non-linear effect is reinstated at all positive inflation levels for both developed and developing countries.

6 Empirical results

In this section, we examine the long-term effects of debt and inflation on economic growth using both ARDL and DL specifications. We also look at the effects of debt thresholds and its trajectory on long-run growth. But first we begin with a description of the data used.

6.1 Data sources

The inflation and output growth are calculated based on consumer price index (CPI) and real gross domestic product (GDP) data series obtained from the International Monetary Fund *International Financial Statistics* database, except for the CPI data for Brazil, China and Tunisia which are obtained from the International Monetary Fund, *World Economic Outlook* database, and the CPI data for the UK, which is obtained from the Reinhart and Rogoff (2010) *Growth in a Time of Debt* database.

The gross government debt/GDP data series are from Reinhart and Rogoff (2011) which are updated and made available online (http://www.carmenreinhart.com/data/browse-by-topic/topics/9/), except for Iran, Morocco, Nigeria, and Syria for which the International Monetary Fund FAD *Historical Public Debt* database was used instead. We focus on gross debt data due to difficulty of collecting net debt data on a consistent basis over time and across countries. Moreover, we use public debt at the general government level for as many countries as possible (Austria, Belgium, Germany, Italy, Netherlands, New Zealand, Singapore, Spain, Sweden, and Tunisia), but given the lack of general public debt data for many countries, central government debt data is used as an alternative.[9]

Since our analysis allows for slope heterogeneity across countries, we need a sufficient number of time periods to estimate country-specific coefficients. To this end, we include only countries in our sample for which we have at least 30 consecutive annual observations on debt, inflation and GDP. Subject to this requirement we ended up with 40 countries listed in Table 14. These countries cover most regions in the world and include advanced, emerging and developing countries. To account for error cross-sectional dependence, we need to form cross-section averages based on a sufficient number of units, and hence set the minimum cross-section dimension to 20. Overall, we ended up with an unbalanced panel covering the sample period 1965-2010, with $T_{\min} = 30$, and $N_{\min} = 20$ across all countries and time periods.[10]

6.2 Estimates based on the ARDL approach not augmented by CS averages

We first consider the long-run effects of debt and inflation on output growth using the traditional panel ARDL approach, in which the long-run effects are calculated from OLS estimates of the short-run coefficients in the following equation:

$$\Delta y_{it} = c_i + \sum_{\ell=1}^{p} \varphi_{i\ell} \Delta y_{i,t-\ell} + \sum_{\ell=0}^{p} \boldsymbol{\beta}'_{i\ell} \mathbf{x}_{i,t-\ell} + u_{it}, \tag{43}$$

where y_{it} is the log of real GDP, $\mathbf{x}_{it} = (\Delta d_{it}, \pi_{it})'$, d_{it} is the log of debt to GDP ratio, and π_{it} is the inflation rate. In a series of papers, Pesaran and Smith (1995), Pesaran (1997), and Pesaran and Shin (1999) show that the traditional ARDL approach can be used for long-run analysis, and that the ARDL methodology is valid regardless of whether the regressors are exogenous, or endogenous, and irrespective of whether the underlying variables are $I(0)$ or $I(1)$. These features of the panel ARDL approach are appealing as reverse causality could be very important in our empirical application. It is well recognized that while high debt burden may have an adverse impact on economic growth, low GDP growth (by reducing tax revenues and increasing public expenditures) could also lead to high debt to GDP ratios. We are indeed interested in looking at the relationship between public debt, inflation and output growth after accounting for these possible feedback effects. Our panel ARDL specification also allows for a significant degree of cross-county heterogeneity and accounts for the fact

[9] The complete dataset, Matlab codes, and Stata do files needed to generate the empirical results in this paper are available from people.ds.cam.ac.uk/km418.

[10] See Section 7 in Chudik and Pesaran (2013b) for further details on the application of the CCE estimators to unbalanced panels.

that the effect of public debt and inflation on growth could vary across countries (particularly in the short run), depending on country-specific factors such as institutions, geographical location, or cultural heritage.

As mentioned in Section 2 and illustrated by MC simulations in Section 4, sufficiently long lags are necessary for the consistency of the ARDL approach, whereas specifying longer lags than necessary can lead to estimates with poor small sample properties. We use the same lag order, p, for all variables/countries, but consider different values of p in the range of 1 to 3. Given that we are working with growth rates which are only moderately persistent, a lag order of 3 should be sufficient to fully account for the short-run dynamics. Also, using the same lag order across all variables and countries help reduce the possible adverse effects of data mining that could accompany the use of country and variable specific lag order selection procedures such as Akaike or Schwarz criteria. Note that our primary focus here is on the long-run estimates rather than the specific dynamics that might be relevant for a particular country.

The Least Squares (LS) estimates obtained from the panel ARDL specifications are reported for three cases, (a), (b) and (c), in Tables 15 and 16.[11] Panel (a) depicts the results when only the debt/GDP variable is included in the ARDL model, panel (b) when only inflation is included, and panel (c) when both variables are included. Each panel gives the average estimates of the long-run effects of debt/GDP growth and inflation on GDP growth (denoted by $\theta_{\Delta d}$ and θ_{π}), and the mean estimate of the coefficients of the error correction term, denoted by λ. For each lag order $p = 1$, 2 and 3, we provide fixed effects (FE) estimates in Table 15 (assuming slope homogeneity), and Mean Group (MG) estimates in Table 16 that allow for slope coefficients to vary across countries. As shown in Pesaran and Smith (1995), the FE estimators will be inconsistent in the presence of slope heterogeneity even if T is sufficiently large. In contrast the MG estimates are consistent under fairly general conditions so long as the errors are cross-sectionally independent.

The results across all specifications suggest an inverse relationship between debt/GDP growth (inflation) and economic growth. Specifically, for case (a) Tables 15 and 16 show that the coefficients of debt/GDP growth are negative and always statistically significant at the 1 percent level, with their values ranging from -0.055 to -0.075 across various estimation techniques and lag orders.[12] For case (b) and when considering the FE estimates, we note that the negative effects of inflation on output growth is -0.025 at various lag orders, while the MG estimates are much larger (falling between -0.054 and -0.104). These estimates are statistically significant at the 1 percent level, with one exception.

Focusing on case (c), where we jointly model debt/GDP growth, inflation, and output growth, we note that a one percentage point increase in debt-to-GDP growth is associated

[11] Individual country estimates are available on request, but it should be noted that they are likely to be individually unreliable given the fact that the time dimension of the panel is relatively small.

[12] The reported standard errors are robust to cross-sectional heteroskedasticity and residual serial correlation as in Arellano (1987).

with a slowdown in GDP growth of between 0.044 and 0.083 percentage points (statistically significant at the 1% level), depending on the selected lag order and estimator, with the MG estimates being generally larger than those of the FE. On the other hand, while the long-run growth effects of inflation are negative (between -0.024 and -0.026) and significant at 1 percent level based on the FE estimates, the MG coefficients are only significant in the case of $p = 1$, suggesting that once we control for debt/GDP and allow for longer lags ($p = 2$ and 3) the long-run impact of inflation on output growth is no longer evident. Overall, the results presented in Tables 15 and 16 are suggestive of negative relationships between debt, inflation, and growth. However, the estimated coefficients vary considerably with different lag augmentation and with/without pooling. It is also worth noting that in all cases, (a)–(c) in Tables 15 and 16, the speed of adjustment to long-run equilibrium is very quick and is in line with the relatively low persistence of output growth in the case of most countries. However, this does not mean that the effects of changes to debt/GDP ratio will also be very quick on the *level* of real output.

6.3 Estimates based on the CS-ARDL approach

The above panel ARDL methodology assumes that the errors in the debt-inflation-growth relationships are cross-sectionally independent, which is likely to be problematic as there are a number of factors such as trade and financial integration, external-debt financing of budget deficits, and exposures to common shocks (i.e. oil price disturbances), that could lead to cross-sectional error dependencies. These global factors are mostly unobserved and can simultaneously affect both domestic growth and public debt, and can lead to badly biased estimates if the unobserved common factors are indeed correlated with the regressors.

Tables 15 and 16 report the CD (Cross-section Dependence) test of Pesaran (2004, 2013), which is based on the average of the pair-wise correlations of the OLS residuals from the individual-country regressions (a-c), and which under the null of cross-section independence is distributed as standard normal.[13] For each $p = 1, 2$, and 3, we observe that the error terms across countries in our model exhibit a considerable degree of cross-sectional dependence as the reported CD statistics are highly significant with very large test statistics. The presence of the cross-sectional dependence implies that estimates obtained using standard panel ARDL models might be misleading. To overcome this problem, we employ the CS-ARDL approach, based on Chudik and Pesaran (2013a), which augments the ARDL regressions with cross-sectional averages of the regressors, the dependant variable and a sufficient number of their lags, which in our case is set to 3 regardless of p, the lag order chosen for the underlying ARDL specification. More specifically, the cross-sectionally augmented ARDL regressions

[13] Theoretical properties of the CD test have been established in the case of strictly exogenous regressors and pure autoregressive models. The properties of the CD test for dynamic panels that include lagged dependent variables and other (weakly or strictly exogenous) regressors have not yet been investigated. However, the Monte Carlo findings reported in Chudik and Pesaran (2013b) suggest that the CD test continues to be valid even when the panel data model contains lagged dependent variable and other regressors.

are given by

$$\Delta y_{it} = c_i + \sum_{\ell=1}^{p} \varphi_{i\ell} \Delta y_{i,t-\ell} + \sum_{\ell=0}^{p} \boldsymbol{\beta}'_{i\ell} \mathbf{x}_{i,t-\ell} + \sum_{\ell=0}^{3} \boldsymbol{\psi}'_{i\ell} \bar{\mathbf{z}}_{t-\ell} + e_{it}, \tag{44}$$

where $\bar{\mathbf{z}}_t = \left(\overline{\Delta y}_t, \bar{\mathbf{x}}'_t\right)'$, and all the other variables are as defined in equation (43).

The estimation results are summarized in Table 17, where we provide MG estimates for the three specifications, (a), (b), and (c), discussed above. For specification (a), we note that the long-run estimates of the debt/GDP growth variable are somewhat larger (ranging between -0.072 and -0.096) than those in Table 16, but still statistically significant at the 1 percent level. The long-run effects of inflation on output growth are similar in most cases to those of the ARDL estimates, except for the CCEMG estimate with $p = 3$ which is not statistically significant. Turning to specification (c), there is now more evidence for negative growth effects of inflation in the long run as the estimates are significant (at the 1% level) in all cases but one. The long-run effects of inflation on growth lies in the range of -0.080 and -0.164. These estimates are much larger than those obtained in Table 16, as the latter does not take into account the possibility that the unobserved common factors are correlated with the regressors. The CD test statistics in Table 17, confirm a substantial decline in the average pair-wise correlation of residuals after the cross-section augmentation of the ARDL models. The coefficients of debt/GDP growth under specification (c) are also larger (between -0.079 to -0.120) using the CS-ARDL regressions, and all of the estimates are statistically significant at the 1 percent level. Finally, the speed of convergence to equilibrium is very fast (and in some instances faster than in the case without augmentation, see Tables 15–17). But as noted earlier and due to the small sample bias in the estimates of the short-run dynamics, the adjustment speeds reported in these tables should be viewed as indicative.

6.4 Estimates based on the CS-DL approach

The results in Tables 15-17 provide evidence of long-run negative effects of both debt and inflation on GDP growth. However, as discussed earlier in the paper, the ARDL and CS-ARDL approaches have their own drawbacks. The sampling uncertainty could be large when the time dimension is moderate and the performance of the estimators also depends on a correct specification of the lag orders of the underlying ARDL specifications. The direct approach to estimating the long-run relationships proposed in this paper (the CS-DL method), is more generally applicable and only requires that a truncation lag order is selected. Also, as can be seen from Section 4, this method has better small sample performance for moderate values of T, which is often the case in applied work. Furthermore, it is robust to a number of departures from the baseline specification such as residual serial correlation, and possible breaks in the error processes.

We estimate the CS-DL versions of the three specifications (a)-(c) and obtain the MG estimates for different truncation lag orders, $p = 1, 2, 3$. We always include three lags of the cross-sectional averages of the regressors in all specifications; namely, we run the following regressions

$$\Delta y_{it} = c_i + \boldsymbol{\theta}_i^{'} \mathbf{x}_{it} + \sum_{\ell=0}^{p-1} \boldsymbol{\delta}_{i\ell}^{'} \Delta \mathbf{x}_{i,t-\ell} + \omega_{iy} \overline{\Delta y}_t + \sum_{\ell=0}^{3} \boldsymbol{\omega}_{i,x\ell}^{'} \overline{\mathbf{x}}_{t-\ell} + e_{it}, \tag{45}$$

where the regressors are defined as in equation (43), with $p = 1, 2, 3$.

The MG estimates based on the above CS-DL regressions are summarized in Table 18. Overall, the estimates are similar to those obtained based on panel ARDL and CS-ARDL regressions given in Tables 15–17. Specifically, the mean group estimates, $\widehat{\boldsymbol{\theta}}_{MG}$, of the effects of debt/GDP and inflation on economic growth are negative and statistically significant (in most cases at the 1% level). The estimated coefficients for the debt/GDP growth variable range from -0.068 to -0.087, and those of inflation fall between -0.066 and -0.089. These estimates fall in a narrow range and tend to be robust to the choice of the truncation lag order. The estimates indicate that, if the debt to GDP ratio is raised permanently, then it will negatively affect economic growth in the long run. But if the increase is temporary and the debt to GDP ratio is actually brought back to its normal level, then there are no long-run adverse effects on economic growth.

However, one drawback of the CS-DL approach is that the estimated long-run effects are only consistent when the feedback effects from the lagged values of the dependent variable to the regressors are absent, although as we have seen in the MC section that, even with this bias, the performance of CS-DL in terms of RMSE is much better than that of the CS-ARDL approach when T is moderate (which is the case in our empirical application). Having said that, it should be noted that no one estimator is perfect and each technique involves a trade-off. Estimators that effectively address a specific econometric problem may lead to a different type of bias. For instance, while the CS-DL estimator is capable of dealing with many modeling issues (cross sectional dependences, robustness to different lag-orders, serial correlations in errors, and breaks in country-specific error processes), it leaves the feedback effects problem unresolved. To deal with different types of econometric issues, and to ensure more robust results, we conducted the debt-inflation-growth exercise based on a range of estimation methods (ARDL, CS-ARDL, and CS-DL). We note that the direction/sign of the long-run relationship between debt and growth is always negative and statistically significant (across different specification and lag orders). This is also the case for the relationship between inflation and growth in most of the models estimated (20 out of 24 coefficients). This gives one more assurance that debt and inflation have a dampening effect on long-run output growth, but given the different biases associated with the direct and indirect approaches to estimating the long-run relationship between debt, inflation and growth, we expect the exact magnitude of the effects to be somewhere in between the two estimates (CS-ARDL and CS-DL).

Given that the CS-DL approach is robust to the possibility of unit roots in variables, we also investigate the long-run effects of the log *level* of debt to GDP ratio and inflation on the log *level* of output. The results are reported in Table 19 from which we observe that a one percent increase in the level of debt/GDP, if sustained, reduces real output by -0.048 to -0.068 percent. These estimates continue to be statistically highly significant in all cases, and suggest, for example, that if a country's debt-to-GDP rises from its normal level of say 70% to 90% and if this increase is maintained, then eventually the country's output might decline by as much as 1.7%.

Finally, we also run regressions where inflation is replaced with the log of CPI in the regressions of log GDP levels and obtained very similar results for the effects of debt/GDP on real output. (Table 19). However, in contrast, the long-run effects of inflation (or log of CPI) on output growth in the level regressions turn out not to be statistically significant.

6.5 Debt/GDP threshold effects on growth

The above results clearly suggest that maintaining high levels of debt-to-GDP are likely to be unsustainable, and if persistent can lead to long-run growth stagnation. However, the estimates obtained so far do not provide any information regarding the normal or acceptable levels of debt-to-GDP. This issue has been addressed by Reinhart and Rogoff (2010) and Checherita-Westphal and Rother (2012) who argue for the presence of a threshold effect in the relationship between debt/GDP and economic growth. RR's analysis is informal and, as noted in our literature review, involves in comparisons of average growth rate differentials across economies classified by their average debt/GDP ratios. They find that these differentials peak when debt/GDP ratio is around 90-100%. Krugman (1988) and Ghosh et al. (2013) also consider possible threshold effects in the relationship between external debt and output growth, which is known as the debt overhang. However, these results are based on strong homogeneity restrictions, in particular the assumption that there exists a *universal* debt/GDP threshold, applicable to all countries equally. It is further assumed (albeit implicitly) that all countries are similarly affected by the threshold effect.

The debt overhang phenomenon in itself seems plausible. What is difficult to accept is the assumption that the level of debt/GDP threshold and its effects on output growth are the same across all countries irrespective of their degree of external debt exposure, historical performance in servicing their public debt, and market perceptions of their economic potential in meeting their debt obligations in future. Due to such intrinsic cross-country heterogeneities, debt thresholds are most-likely country specific and must be estimated as such. However, identification and estimation of country-specific debt thresholds are not feasible due to short time-series data that are currently available.

To explore the importance of heterogeneity and potential nonlinearity in the debt-growth relationship, initially we begin with the following baseline homogeneous panel data model

$$\Delta y_{it} = c_{\tau} + \gamma_{\tau} I_{it}(\tau) + e_{it}, \tag{46}$$

where $I_{it}(\tau)$ is a "threshold dummy", defined by the indicator variable $I(d_{it} \geq \log \tau)$ which takes the value of 1 if debt/GDP is above the given threshold value of τ, and zero otherwise. As before y_{it} is the log of real GDP, and d_{it} is the log of debt/GDP. In addition to assuming a universal threshold, τ, this model also assumes that the coefficients of the "threshold dummy", γ_{τ}, is the same across all countries whose debt/GDP ratio is above the same threshold. c_{τ} is the average GDP growth of countries with debt/GDP below τ.

The estimates of c_{τ} and γ_{τ} for values of $\tau = 30\%, 40\%, ..., 90\%$, are given on the top panel of Table 20.[14] The results show estimates of c_{τ} that are quite stable across different values of τ, which is in line with the rather small estimates obtained for γ_{τ}. The differences between average GDP growth for countries above a certain debt/GDP ratio and countries below the same threshold level are relatively flat over a range of values for τ. The estimates of γ_{τ} also show that while average GDP growth declines when the public debt/GDP ratio increases, one cannot find a tipping point beyond which long-term growth is reduced substantially.

We now consider a less restrictive model which uses a universal threshold, but allows the effects of the threshold dummy to differ across countries. This is a more plausible specification since it allows the threshold dummy, for example, to have a zero loading for a country like Japan, and possibly a large negative estimate for a country like Greece or Spain. Specifically, we consider

$$\Delta y_{it} = c_{i\tau} + \gamma_{i\tau} I_{it}(\tau) + e_{it}, \tag{47}$$

and report MG estimates of c_{τ} and γ_{τ}, defined as averages of the estimates of $c_{i\tau}$ and $\gamma_{i\tau}$ across countries with a given threshold, in Table 20. The results are qualitatively similar to those obtained for the homogenous case, but with larger estimates for γ_{τ}. If anything, the heterogenous specification is more supportive of the Reinhart and Rogoff position, partly due to the fact that it does not treat all the countries similarly.

Although specification (47) deals with heterogeneity, it does not allow for cross-country dependencies, dynamics, and non-threshold effects of debt/GDP growth and inflation variables on output growth. To address these problems, we consider the following specification which is a generalization of our earlier set up:

$$\Delta y_{it} = c_{i\tau} + \gamma_{i\tau} I_{it}(\tau) + \boldsymbol{\theta}'_{i\tau}\mathbf{x}_{it} + \sum_{\ell=0}^{2} \boldsymbol{\delta}'_{i\ell,\tau}\Delta\mathbf{x}_{i,t-\ell} + \omega_{iy,\tau}\overline{\Delta y}_t + \sum_{\ell=0}^{3} \boldsymbol{\omega}'_{i,x\ell,\tau}\overline{\mathbf{x}}_{t-\ell} + e_{it}, \tag{48}$$

where $\mathbf{x}_{it} = (\Delta d_{it}, \pi_{it})'$. The MG estimates of the parameters of interest, γ_{τ} and $\boldsymbol{\theta}_{\tau}$, are summarized in Table 20. In sharp contrast to the estimates based on (46) and (47), none of

[14] We report heteroscedasticity-robust standard errors.

the estimates of γ_τ are statistically significant. We note that, as before, the long-run effects of debt on growth are always statistically significant and negative in the range of -0.063 and -0.109 depending on τ. Therefore, our results show that there is no simple common threshold for the *level* of government debt above which growth is more adversely affected.

As our results have consistently shown that higher and sustained debt/GDP growth tend to adversely affect output growth, and having shown that the presence of simple threshold effects is not supported by the data, we turned to other non-linear threshold effects which became binding only in the case of countries with rising debt/GDP rates. Accordingly, we estimated the following specification,

$$\begin{aligned}\Delta y_{it} &= c_{i\tau} + \gamma_{i\tau} I_{it}(\tau) + \gamma_{i\tau}^{+} \left[I_{it}(\tau) \times \max\left(0, \Delta d_{it}\right)\right] + \boldsymbol{\theta}_{i\tau}^{\prime} \mathbf{x}_{it} + \sum_{\ell=0}^{2} \boldsymbol{\delta}_{i\ell,\tau}^{\prime} \Delta \mathbf{x}_{i,t-\ell} \\ &\quad + \omega_{iy,\tau} \overline{\Delta y}_{t} + \sum_{\ell=0}^{3} \boldsymbol{\omega}_{i,x\ell,\tau}^{\prime} \overline{\mathbf{x}}_{t-\ell} + e_{it}, \end{aligned} \tag{49}$$

which is the same as (48), except for the interactive term, $I_{it}(\tau) \times \max\left(0, \Delta d_{it}\right)$, which is non-zero only if $\Delta d_{it} > 0$, and $d_{it} > \log(\tau)$. The MG estimates for this model are summarized in Table 21. The results show that when samples of country episodes with an upward debt trajectory above certain thresholds are chosen, the coefficients of the interactive threshold dummy variable (i.e. $\widehat{\gamma}_\tau^{+}$) becomes negative and statistically significant if debt/GDP ratio is above 60%. However, as before the coefficient of the threshold dummy ($\widehat{\gamma}_\tau$) is not statistically significant. We therefore remove $I_{it}(\tau)$ and instead estimate

$$\Delta y_{it} = c_{i\tau} + \gamma_{i\tau}^{+} \left[I_{it}(\tau) \times \max\left(0, \Delta d_{it}\right)\right] + \boldsymbol{\theta}_{i\tau}^{\prime} \mathbf{x}_{it} + \sum_{\ell=0}^{2} \boldsymbol{\delta}_{i\ell,\tau}^{\prime} \Delta \mathbf{x}_{i,t-\ell} + \omega_{iy,\tau} \overline{\Delta y}_{t} + \sum_{\ell=0}^{3} \boldsymbol{\omega}_{i,x\ell,\tau}^{\prime} \overline{\mathbf{x}}_{t-\ell} + e_{it}. \tag{50}$$

Again we observe that the coefficients of the interactive threshold dummy variable are negative and statistically significant beyond 60 percent debt/GDP ratio while at the same time the coefficient of debt growth ($\widehat{\theta}_{\Delta d,\tau}$) is significant and falls between -0.056 and -0.100, which is in line with the results obtained in Tables 17-18. The results in Table 21 indicate that *debt trajectory* is probably more important than the *level* of debt itself.

7 Concluding remarks

Estimation of the long-run effects of public debt on economic growth has received renewed interest among economists and policy makers in the aftermath of the global financial crisis and the European sovereign debt crisis. Due to a significant worsening of public finances in many advanced economics and more limited fiscal space in these countries (compared with 2008), the interaction between public debt and economic growth is attracting greater attention. Recent sovereign debt problems in Greece and other European economies and negative feedback loops between sovereigns and the banking system have also contributed to this renewed interest in the interplay between public debt and economic growth, and in general on the design of policies that balance short-run gains from fiscal expansion with possible adverse effects on growth in the long run. This paper revisited the question of the long-run effects of debt on growth empirically in a dynamic heterogeneous and cross-sectionally correlated unbalanced panel of countries. Our findings suggest that there is a significant negative long-run relationship between *rising* debt and economic growth, and that the trajectory of the debt can have more important consequences for economic growth than the level of the debt itself, particularly beyond certain debt level thresholds.

In particular, our results show that following episodes of increasing public debt, governments need to adopt fiscal measures that credibly reduce the overall debt/GDP ratio to normal levels in order to prevent the negative long-run growth effects of debt. This policy is compatible with Keynesian fiscal deficit spending, so long as it is coupled with credible fiscal policy announcements that aim at reducing the debt burden to levels considered as normal for the country in question. Our analysis does not provide any guidelines as to what might be considered normal levels of debt/GDP ratio, except in cases where debt/GDP ratio is high and rising, and there is no credible expectations of a reversal in the debt/GDP trajectory.

Estimation of long-run effects is an important applied problem in many fields of economics. We have discussed how to estimate long-run effects in a typical macroeconomic panel, where errors are cross-sectionally dependent, slopes are heterogeneous, and dynamic effects include lagged values of the dependent variable. We have provided new Monte Carlo results showing the robustness of the estimates of the long-run effects based on panel ARDL models to the endogeneity problem. We have also contributed to the econometric analysis of long-run effects by proposing a new cross-section augmented distributed lag (CS-DL) approach which is robust to residual serial correlation, breaks in error processes and dynamic misspecifications. But unlike the ARDL approach, the CS-DL procedure is not robust to the endogeneity problem, and could be subject to simultaneity bias. Nevertheless, the extensive Monte Carlo experiments reported in the paper suggest that the endogeneity bias of the CS-DL approach is more than compensated for its better small sample performance as compared to the ARDL procedure when the time dimension is not very large. ARDL seems to dominate CS-DL only if the time dimension is sufficiently large, which is often lacking in empirical applications.

Table 1: Monte Carlo Estimates of Bias, RMSE, Size and Power for Estimation of LR Coefficient (θ) in Baseline Experiment

DGP is ARDL(2,1) model with heterogeneous coefficients, $\varphi_{\max} = 0.6$, stationary regressors, $m = 2$ factors, no feedback effects and $\rho_{\varepsilon i} = 0$.

	Bias (×100)					Root Mean Square Errors (×100)					Size (5% level, $H_0 : \theta = 1$)					Power (5% level, $H_1 : \theta = 1.2$)				
(N,T)	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**
	CS-DL mean group ($p_{\bar{y}} = 0$, $p_{\bar{x}} = \lfloor T^{1/3} \rfloor$ and $p = p_{\bar{x}} - 1$)																			
30	-0.65	-0.49	0.04	-0.11	-0.15	16.88	11.24	8.55	7.34	6.44	6.35	6.15	7.75	5.80	6.40	28.30	50.20	70.30	80.10	87.15
50	-1.12	-1.00	-0.34	-0.12	-0.03	13.19	8.83	6.33	5.82	4.92	5.90	6.15	5.25	6.45	5.20	39.45	70.20	89.10	93.45	97.80
100	-1.32	-0.92	-0.09	-0.11	0.15	9.66	6.25	4.49	4.03	3.56	5.95	6.30	5.55	4.50	5.45	62.95	92.45	99.50	99.75	100.00
150	-1.19	-0.96	-0.11	0.16	-0.05	7.91	5.24	3.78	3.38	2.94	5.90	6.50	5.90	6.25	5.75	79.45	98.20	99.85	100.00	100.00
200	-1.06	-0.75	-0.24	-0.07	0.03	6.70	4.38	3.17	2.86	2.47	5.60	6.00	4.85	5.10	4.50	88.65	99.80	100.00	100.00	100.00
	CS-DL pooled ($p_{\bar{y}} = 0$, $p_{\bar{x}} = \lfloor T^{1/3} \rfloor$ and $p = p_{\bar{x}} - 1$)																			
30	-0.40	-0.19	0.16	-0.08	-0.04	15.31	10.44	8.12	7.10	6.41	6.95	6.60	7.20	6.95	6.50	33.75	53.65	73.15	82.45	88.25
50	-0.93	-1.00	-0.29	-0.16	-0.06	11.47	8.31	6.08	5.55	4.80	6.00	6.70	5.70	6.00	4.60	47.05	75.50	91.40	94.80	97.80
100	-1.09	-0.83	-0.13	-0.10	0.14	8.17	5.88	4.26	3.90	3.53	5.85	6.35	5.20	4.60	5.80	74.20	95.10	99.75	99.90	99.95
150	-1.02	-0.72	-0.09	0.11	-0.02	6.83	4.82	3.55	3.28	2.87	5.90	6.10	5.80	5.95	5.55	87.60	98.95	99.95	100.00	100.00
200	-0.81	-0.68	-0.22	-0.06	0.03	5.89	4.11	3.04	2.74	2.43	5.25	5.70	5.20	5.15	5.05	94.55	99.95	100.00	100.00	100.00
	Based on CS-ARDL estimates of short-run coefficients (ARDL(2,1) specifications with $p_{\bar{z}} = \lfloor T^{1/3} \rfloor$).																			
30	-10.40	-3.87	-1.73	-1.38	-0.88	361.68	20.62	7.59	6.24	5.54	10.75	9.80	9.45	8.25	7.50	39.95	66.05	86.05	94.10	96.00
50	3.72	-3.96	-2.12	-1.30	-1.00	182.36	9.56	5.83	4.89	4.32	10.45	12.75	8.10	7.65	6.55	45.40	81.30	98.00	98.90	99.70
100	17.57	-4.03	-2.02	-1.39	-0.83	966.97	7.31	4.34	3.58	3.16	12.90	13.90	10.05	8.45	8.05	61.30	96.40	100.00	100.00	100.00
150	-9.46	-3.93	-2.03	-1.20	-1.07	159.90	6.46	3.84	3.03	2.68	13.65	18.45	12.05	9.45	8.95	67.60	99.50	100.00	100.00	100.00
200	11.29	-3.97	-2.15	-1.42	-1.01	678.37	6.66	3.52	2.79	2.31	13.90	20.10	13.10	11.05	7.70	71.00	99.40	100.00	100.00	100.00
	Based on CS-ARDL estimates of short-run coefficients (ARDL(1,0) specifications with $p_{\bar{z}} = \lfloor T^{1/3} \rfloor$).																			
30	-23.29	-26.91	-23.58	-22.66	-22.01	274.46	28.91	24.83	23.77	22.99	58.05	75.70	85.90	90.40	91.95	86.10	98.00	99.90	100.00	100.00
50	-27.78	-27.49	-23.95	-22.72	-22.28	109.97	28.68	24.70	23.36	22.83	73.20	90.75	97.25	98.60	99.25	92.00	99.80	100.00	100.00	100.00
100	-31.85	-27.64	-24.18	-22.94	-22.21	62.82	28.23	24.54	23.26	22.50	87.75	99.45	100.00	100.00	100.00	96.10	100.00	100.00	100.00	100.00
150	-30.11	-27.60	-24.01	-22.82	-22.34	81.52	28.02	24.26	23.05	22.53	93.35	99.90	100.00	100.00	100.00	97.20	100.00	100.00	100.00	100.00
200	-31.20	-27.73	-24.20	-22.96	-22.40	50.42	28.04	24.39	23.12	22.55	94.25	100.00	100.00	100.00	100.00	97.30	100.00	100.00	100.00	100.00

Notes: The dependent variable and regressors are generated according to (36)-(37) with correlated fixed effects, and with cross-sectionally weakly dependent and serially correlated heteroskedastic idiosyncratic innovations generated according to (40)-(41) with $a_{\varepsilon} = 0.6$. The knowledge of lag orders is not used in the estimation stage and the integer part of $T^{1/3}$ gives $3, 3, 4, 5$ and 5 for $T = 30, 50, 100, 150$ and 200, respectively.

Table 2: Monte Carlo Estimates of Bias, RMSE, Size and Power for Estimation of LR Coefficient (θ) in the Case of Homogeneous Long-Run

DGP is ARDL(2,1) model with homogeneous long-run, heterogeneous short-run, $\varphi_{\max} = 0.6$, stationary regressors, $m = 2$ factors, no feedback effects and $\rho_{\varepsilon i} = 0$.

	Bias (×100)					Root Mean Square Errors (×100)					Size (5% level, $H_0 : \theta = 1$)					Power (5% level, $H_1 : \theta = 1.2$)				
(N,T)	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**
	CS-DL mean group ($p_{\bar{y}} = 0$, $p_{\bar{x}} = \lceil T^{1/3} \rceil$ and $p = p_{\bar{x}} - 1$)																			
30	-1.59	-0.63	0.00	-0.05	0.01	16.55	10.27	7.40	6.21	5.31	6.00	5.30	6.40	5.25	5.85	30.85	55.85	78.45	88.30	95.05
50	-1.49	-0.95	-0.34	-0.02	-0.11	12.92	8.04	5.89	4.97	4.22	5.25	5.85	6.55	5.60	6.30	41.75	74.65	93.15	97.30	99.55
100	-0.96	-0.55	-0.20	0.02	0.02	9.16	5.76	4.09	3.53	3.02	5.10	5.70	5.40	5.10	5.70	64.90	94.40	99.75	100.00	100.00
150	-1.06	-0.83	-0.18	0.00	-0.10	7.63	4.94	3.35	2.88	2.41	5.45	6.50	5.45	5.35	5.35	79.55	98.70	100.00	100.00	100.00
200	-1.10	-0.77	0.02	-0.02	-0.08	6.47	4.22	2.87	2.52	2.10	4.60	5.95	5.20	5.30	5.00	89.70	99.90	100.00	100.00	100.00
	CS-DL pooled ($p_{\bar{y}} = 0$, $p_{\bar{x}} = \lceil T^{1/3} \rceil$ and $p = p_{\bar{x}} - 1$)																			
30	-0.97	-0.42	-0.04	-0.08	0.04	14.33	9.31	6.81	5.83	4.92	7.00	5.55	6.75	5.50	5.95	36.45	62.00	84.30	91.90	97.00
50	-1.01	-0.83	-0.29	-0.02	-0.09	11.07	7.38	5.40	4.59	3.91	5.90	5.20	6.20	5.30	6.85	49.75	80.45	95.50	98.15	99.75
100	-0.72	-0.51	-0.16	0.00	0.05	7.87	5.32	3.76	3.21	2.76	5.30	5.70	5.45	5.60	5.80	76.25	96.60	99.85	100.00	100.00
150	-0.80	-0.61	-0.14	0.05	-0.08	6.55	4.36	3.06	2.59	2.24	6.00	6.05	5.40	5.05	5.15	89.55	99.80	100.00	100.00	100.00
200	-0.95	-0.67	0.04	-0.03	-0.07	5.49	3.83	2.59	2.29	1.91	5.00	6.35	5.10	5.60	5.25	96.15	100.00	100.00	100.00	100.00
	Based on CS-ARDL estimates of short-run coefficients (ARDL(2,1) specifications with $p_{\bar{z}} = \lceil T^{1/3} \rceil$).																			
30	-0.25	-3.59	-1.80	-1.24	-0.80	284.59	10.67	6.61	5.00	4.10	10.90	9.90	10.00	8.30	7.90	41.75	72.45	92.20	98.40	99.65
50	-8.02	-3.87	-2.03	-1.39	-1.03	85.61	8.89	5.21	4.08	3.37	10.80	11.10	9.65	9.75	8.55	50.70	85.65	99.35	99.90	100.00
100	-3.69	-3.87	-2.17	-1.38	-0.95	154.40	6.94	4.03	3.01	2.45	12.25	14.85	12.45	10.05	9.10	62.05	97.65	100.00	100.00	100.00
150	-0.25	-3.94	-2.11	-1.37	-1.06	90.79	6.27	3.51	2.60	2.12	14.05	18.10	14.80	12.30	11.75	67.95	99.50	100.00	100.00	100.00
200	-1.53	-3.98	-2.06	-1.37	-1.05	320.07	5.75	3.16	2.39	1.92	15.25	22.55	16.25	14.25	12.15	71.90	99.70	100.00	100.00	100.00
	Based on CS-ARDL estimates of short-run coefficients (ARDL(1,0) specifications with $p_{\bar{z}} = \lceil T^{1/3} \rceil$).																			
30	-30.79	-27.20	-23.78	-22.80	-22.12	55.20	28.90	24.89	23.63	22.85	62.55	81.15	91.10	96.25	98.30	87.00	98.90	100.00	100.00	100.00
50	-32.12	-27.63	-24.23	-23.14	-22.57	47.11	28.66	24.88	23.66	23.00	76.85	94.25	99.05	99.75	100.00	94.25	99.90	100.00	100.00	100.00
100	-34.77	-27.78	-24.41	-23.21	-22.57	252.97	28.33	24.72	23.46	22.79	89.85	99.60	99.95	100.00	100.00	96.65	100.00	100.00	100.00	100.00
150	-32.53	-27.88	-24.35	-23.16	-22.62	36.50	28.27	24.56	23.33	22.76	92.55	99.90	100.00	100.00	100.00	96.90	100.00	100.00	100.00	100.00
200	-33.96	-27.90	-24.38	-23.16	-22.63	53.98	28.17	24.54	23.28	22.75	94.35	99.85	100.00	100.00	100.00	97.05	99.95	100.00	100.00	100.00

Notes: The dependent variable and regressors are generated according to (36)-(37) with correlated fixed effects, and with cross-sectionally weakly dependent and serially correlated heteroskedastic idiosyncratic innovations generated according to (40)-(41) with $a_\varepsilon = 0.6$. The knowledge of lag orders is not used in the estimation stage and the integer part of $T^{1/3}$ gives $3, 3, 4, 5$ and 5 for $T = 30, 50, 100, 150$ and 200, respectively.

Table 3: Monte Carlo Estimates of Bias, RMSE, Size and Power for Estimation of LR Coefficient (θ) in the Case of Homogeneous Short-Run

DGP is ARDL(2,1) models with homogeneous short-run, $\varphi_{\max} = 0.6$, stationary regressors, $m = 2$ factors, no feedback effects and $\rho_{\varepsilon i} = 0$.

	Bias (×100)					Root Mean Square Errors (×100)					Size (5% level, $H_0 : \theta = 1$)					Power (5% level, $H_1 : \theta = 1.2$)				
(N,T)	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**
	CS-DL mean group ($p_{\bar{y}} = 0$, $p_{\bar{x}} = \lceil T^{1/3} \rceil$ and $p = p_{\bar{x}} - 1$)																			
30	-2.20	-1.52	-0.09	-0.34	-0.35	19.01	12.19	8.55	7.37	6.14	7.25	6.00	6.40	5.75	5.70	27.80	49.20	69.05	80.30	90.00
50	-2.23	-1.69	-0.51	0.17	0.05	14.93	9.35	6.85	5.79	4.84	6.10	5.45	6.30	6.55	5.60	35.95	67.20	86.15	92.40	97.25
100	-2.24	-1.84	-0.36	-0.21	0.06	10.55	6.93	4.68	4.01	3.33	6.10	7.15	5.45	5.10	4.90	59.10	89.95	98.75	99.90	100.00
150	-1.98	-1.99	-0.47	-0.11	-0.03	8.79	5.82	3.91	3.35	2.66	6.30	7.30	6.50	5.35	4.20	75.35	98.00	99.95	100.00	100.00
200	-2.22	-1.86	-0.35	-0.20	-0.01	7.94	4.96	3.38	2.85	2.38	6.70	6.90	5.25	4.75	4.80	84.40	99.65	100.00	100.00	100.00
	CS-DL pooled ($p_{\bar{y}} = 0$, $p_{\bar{x}} = \lceil T^{1/3} \rceil$ and $p = p_{\bar{x}} - 1$)																			
30	-1.94	-1.39	-0.03	-0.35	-0.44	16.68	10.97	7.95	6.83	5.81	7.05	6.00	6.55	6.20	5.85	32.40	54.20	74.35	84.85	93.15
50	-1.96	-1.45	-0.40	0.16	0.02	12.88	8.70	6.29	5.36	4.43	6.80	6.65	6.25	6.60	5.45	44.55	72.80	89.75	95.50	98.75
100	-2.00	-1.66	-0.31	-0.16	0.04	9.10	6.34	4.37	3.70	3.07	6.25	6.10	6.20	5.35	5.25	70.55	93.50	99.30	99.95	100.00
150	-1.68	-1.62	-0.43	-0.08	-0.04	7.61	5.22	3.57	3.05	2.48	6.40	7.10	6.00	4.95	4.05	84.15	99.25	100.00	100.00	100.00
200	-1.94	-1.61	-0.31	-0.19	-0.04	6.76	4.50	3.13	2.59	2.20	6.95	6.55	5.05	3.95	4.45	92.70	99.75	100.00	100.00	100.00
	Based on CS-ARDL estimates of short-run coefficients (ARDL(2,1) specifications with $p_{\bar{z}} = \lceil T^{1/3} \rceil$).																			
30	-14.08	-3.81	-1.92	-1.59	-1.26	310.27	12.81	7.66	6.02	5.04	10.50	10.45	9.20	9.40	9.35	32.50	61.25	84.10	95.30	98.25
50	-4.56	-4.15	-2.14	-1.35	-1.00	242.69	10.59	6.20	4.72	3.98	10.15	11.85	9.80	8.75	8.95	37.90	76.00	96.15	99.50	99.75
100	2.62	-4.32	-2.28	-1.51	-1.11	203.52	8.24	4.61	3.48	2.88	10.70	14.15	11.60	9.35	9.25	47.05	93.10	99.85	100.00	100.00
150	-3.39	-4.50	-2.35	-1.56	-1.12	163.77	7.29	4.09	3.03	2.43	9.90	18.25	14.30	12.10	9.25	51.30	98.35	100.00	100.00	100.00
200	-13.55	-4.32	-2.31	-1.64	-1.18	298.99	6.58	3.71	2.77	2.22	11.40	21.30	15.50	14.05	11.80	56.00	99.45	100.00	100.00	100.00
	Based on CS-ARDL estimates of short-run coefficients (ARDL(1,0) specifications with $p_{\bar{z}} = \lceil T^{1/3} \rceil$).																			
30	-14.02	-10.86	-6.61	-5.78	-5.30	53.22	15.60	9.81	8.21	7.30	24.60	25.10	22.30	22.70	23.35	58.60	82.10	94.65	98.45	99.90
50	-15.83	-11.12	-7.07	-5.88	-5.26	67.84	14.14	9.05	7.41	6.59	30.70	34.25	30.60	31.40	33.65	71.95	93.60	99.90	99.95	100.00
100	-8.04	-11.36	-7.41	-6.14	-5.54	407.90	12.88	8.38	6.88	6.20	41.55	53.15	51.55	52.70	56.60	83.55	99.55	100.00	100.00	100.00
150	-15.24	-11.63	-7.51	-6.24	-5.54	43.70	12.68	8.17	6.77	5.99	50.75	68.95	68.25	70.15	72.20	88.85	100.00	100.00	100.00	100.00
200	-14.75	-11.53	-7.54	-6.33	-5.62	37.89	12.32	8.03	6.70	5.94	56.55	79.15	79.75	82.75	84.60	90.60	99.95	100.00	100.00	100.00

Notes: The dependent variable and regressors are generated according to (36)-(37) with correlated fixed effects, and with cross-sectionally weakly dependent and serially correlated heteroskedastic idiosyncratic innovations generated according to (40)-(41) with $a_\varepsilon = 0.6$. The knowledge of lag orders is not used in the estimation stage and the integer part of $T^{1/3}$ gives $3, 3, 4, 5$ and 5 for $T = 30, 50, 100, 150$ and 200, respectively.

Table 4: Monte Carlo Estimates of Bias, RMSE, Size and Power for Estimation of LR Coefficient (θ) in the Case of ARDL(1,0) Model

DGP is ARDL(1,0) model with heterogeneous coefficients, $\varphi_{\max} = 0.6$, stationary regressors, $m = 2$ factors, no feedback effects and $\rho_{\varepsilon i} = 0$.

	Bias (×100)					Root Mean Square Errors (×100)					Size (5% level, $H_0 : \theta = 1$)					Power (5% level, $H_1 : \theta = 1.2$)				
(N,T)	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**
	CS-DL mean group ($p_{\bar{y}} = 0$, $p_{\bar{x}} = \lceil T^{1/3} \rceil$ and $p = p_{\bar{x}} - 1$)																			
30	-2.81	-2.34	-1.01	-0.45	-0.51	16.57	10.87	7.93	7.39	6.40	7.00	6.50	5.60	6.80	6.35	34.90	59.65	77.10	82.55	88.80
50	-2.77	-2.29	-0.96	-0.53	-0.57	13.13	8.64	6.25	5.48	5.04	7.30	6.50	6.25	5.45	6.40	47.70	77.20	92.30	95.45	98.20
100	-2.34	-2.30	-0.99	-0.44	-0.57	8.99	6.39	4.51	3.97	3.57	5.90	7.15	6.05	4.90	5.75	71.20	96.00	99.65	99.90	100.00
150	-2.50	-2.18	-1.04	-0.52	-0.52	7.94	5.21	3.71	3.21	2.95	7.10	7.35	6.10	4.90	5.80	85.10	99.45	100.00	100.00	100.00
200	-2.95	-2.36	-1.08	-0.51	-0.57	7.05	4.73	3.37	2.88	2.57	8.25	8.55	7.70	6.50	5.70	94.55	99.90	100.00	100.00	100.00
	CS-DL pooled ($p_{\bar{y}} = 0$, $p_{\bar{x}} = \lceil T^{1/3} \rceil$ and $p = p_{\bar{x}} - 1$)																			
30	-2.50	-2.03	-0.98	-0.38	-0.55	14.54	10.27	7.74	7.23	6.34	6.60	7.20	6.10	6.65	5.90	39.80	62.10	79.85	82.80	89.35
50	-2.58	-2.10	-0.81	-0.51	-0.52	11.58	8.08	6.09	5.31	4.95	6.85	6.50	6.15	5.75	5.60	55.75	81.10	93.30	96.45	98.40
100	-2.08	-2.14	-0.88	-0.37	-0.54	7.97	6.04	4.31	3.89	3.51	5.70	7.25	6.15	5.40	6.10	80.65	96.75	99.80	99.95	100.00
150	-2.40	-1.92	-0.92	-0.49	-0.46	6.89	4.88	3.54	3.10	2.90	7.25	7.10	5.80	5.10	5.55	93.30	99.70	100.00	100.00	100.00
200	-2.69	-2.14	-0.96	-0.44	-0.49	6.22	4.39	3.22	2.75	2.53	8.25	7.75	6.75	6.05	5.60	98.45	100.00	100.00	100.00	100.00
	Based on CS-ARDL estimates of short-run coefficients (ARDL(2,1) specifications with $p_{\bar{z}} = \lceil T^{1/3} \rceil$).																			
30	-20.46	-3.19	-1.79	-1.05	-0.78	1278.76	10.93	7.21	6.12	5.43	9.95	9.10	7.80	7.70	6.80	39.75	66.70	87.80	93.50	96.25
50	-9.43	-3.23	-1.59	-1.12	-0.88	356.48	8.81	5.55	4.70	4.34	8.95	9.90	7.20	7.25	6.60	48.55	81.60	97.10	99.15	99.70
100	-2.25	-3.30	-1.72	-1.13	-0.96	99.43	6.69	4.25	3.42	3.11	9.05	12.20	9.20	6.95	7.60	58.00	97.15	99.95	100.00	100.00
150	-34.29	-2.95	-1.83	-1.15	-0.87	819.84	5.85	3.63	2.94	2.65	11.15	11.85	10.50	8.60	8.60	65.35	98.90	100.00	100.00	100.00
200	-2.09	-3.32	-1.76	-1.18	-0.93	101.18	6.98	3.22	2.59	2.30	11.70	16.50	10.80	8.95	7.70	70.55	99.45	100.00	100.00	100.00
	Based on CS-ARDL estimates of short-run coefficients (ARDL(1,0) specifications with $p_{\bar{z}} = \lceil T^{1/3} \rceil$).																			
30	-6.54	-4.05	-1.86	-0.96	-0.64	19.35	9.58	6.56	5.50	4.98	13.80	12.15	9.75	8.60	6.90	62.55	82.50	93.95	96.85	98.10
50	-6.10	-4.52	-1.89	-1.25	-0.89	26.10	8.02	5.05	4.35	4.04	17.55	14.40	9.00	7.40	6.65	76.65	95.00	99.25	99.75	99.85
100	-18.34	-4.40	-2.12	-1.38	-1.04	537.53	6.50	4.06	3.21	2.91	23.30	19.50	12.10	8.55	7.80	90.55	99.85	100.00	100.00	100.00
150	-11.14	-4.26	-2.27	-1.43	-1.01	129.28	5.75	3.58	2.83	2.51	32.05	23.20	15.65	11.10	9.65	95.65	99.90	100.00	100.00	100.00
200	-7.45	-4.62	-2.21	-1.46	-1.11	11.91	5.73	3.25	2.59	2.24	36.55	31.35	17.45	12.40	9.95	97.30	100.00	100.00	100.00	100.00

Notes: The dependent variable and regressors are generated according to (36)-(37) with correlated fixed effects, and with cross-sectionally weakly dependent and serially correlated heteroskedastic idiosyncratic innovations generated according to (40)-(41) with $a_{\varepsilon} = 0.6$. The knowledge of lag orders is not used in the estimation stage and the integer part of $T^{1/3}$ gives $3, 3, 4, 5$ and 5 for $T = 30, 50, 100, 150$ and 200, respectively.

Table 5: Monte Carlo Estimates of Bias, RMSE, Size and Power for Estimation of LR Coefficient (θ) in the Case of $\phi_{\max} = 0.8$

DGP is ARDL(2,1) model with heterogeneous coefficients, $\varphi_{\max} = 0.8$, stationary regressors, $m = 2$ factors, no feedback effects and $\rho_{\varepsilon i} = 0$.

	Bias ($\times$100)					Root Mean Square Errors ($\times$100)					Size (5% level, $H_0: \theta = 1$)					Power (5% level, $H_1: \theta = 1.2$)				
(N,T)	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**
	CS-DL mean group ($p_{\bar{y}} = 0$, $p_{\bar{x}} = \lceil T^{1/3} \rceil$ and $p = p_{\bar{x}} - 1$)																			
30	-5.95	-5.68	-2.87	-1.19	-1.86	21.83	14.77	10.87	9.28	8.23	7.70	8.65	7.45	5.75	6.90	28.35	51.30	63.45	66.20	79.25
50	-6.32	-5.87	-2.92	-1.59	-1.70	17.52	12.12	8.42	7.36	6.68	7.90	9.45	6.90	6.60	6.50	40.50	70.75	82.50	86.35	92.55
100	-6.47	-5.47	-3.03	-1.81	-1.67	13.09	9.40	6.46	5.24	4.83	9.35	12.35	8.85	5.85	7.70	65.10	91.90	98.25	98.95	99.70
150	-6.24	-5.60	-2.95	-1.65	-1.70	11.21	8.34	5.55	4.40	4.07	10.55	13.85	9.40	6.25	8.50	80.95	98.45	99.95	99.85	99.95
200	-6.32	-5.68	-3.08	-1.66	-1.68	10.26	7.94	5.05	4.00	3.55	12.30	18.90	10.70	8.05	8.10	90.40	99.65	100.00	100.00	100.00
	CS-DL pooled ($p_{\bar{y}} = 0$, $p_{\bar{x}} = \lceil T^{1/3} \rceil$ and $p = p_{\bar{x}} - 1$)																			
30	-5.46	-5.46	-2.49	-1.12	-1.73	19.66	13.97	10.21	8.88	7.85	7.20	7.75	7.45	5.90	6.50	32.70	55.50	65.80	69.15	81.85
50	-5.54	-5.31	-2.80	-1.29	-1.52	15.73	11.23	8.01	7.04	6.45	8.00	8.75	6.45	6.35	6.70	44.60	73.95	84.45	87.80	93.65
100	-6.04	-5.03	-2.78	-1.64	-1.48	11.81	8.76	6.07	4.98	4.68	10.20	10.65	8.65	6.20	7.95	72.05	94.20	98.50	99.10	99.70
150	-5.61	-5.12	-2.63	-1.46	-1.54	10.06	7.77	5.20	4.19	3.86	9.95	13.25	10.20	6.90	7.70	86.00	98.80	99.95	99.85	99.95
200	-5.78	-5.05	-2.70	-1.47	-1.56	9.20	7.19	4.72	3.79	3.41	12.45	17.45	10.35	7.05	7.55	94.60	99.75	100.00	100.00	100.00
	Based on CS-ARDL estimates of short-run coefficients (ARDL(2,1) specifications with $p_{\bar{z}} = \lceil T^{1/3} \rceil$).																			
30	-12.89	-5.14	-2.82	-1.50	-1.28	786.15	19.34	9.43	7.65	6.69	10.85	11.55	8.90	7.35	7.60	35.00	55.00	74.90	82.50	90.00
50	-2.09	-8.46	-2.72	-1.69	-1.29	394.84	187.62	7.57	6.04	5.40	11.55	10.80	9.15	8.00	7.45	37.90	68.50	89.35	95.00	97.90
100	-30.77	-5.24	-2.87	-2.08	-1.45	768.23	19.56	5.70	4.55	4.04	11.50	14.75	10.10	9.15	9.00	44.10	85.50	98.85	99.75	100.00
150	-15.09	-4.79	-2.98	-1.96	-1.47	375.14	18.01	5.25	3.85	3.37	12.30	16.90	14.00	10.65	9.30	46.45	92.40	99.50	100.00	100.00
200	-1.15	-6.97	-3.07	-1.96	-1.47	229.52	69.18	4.68	3.59	2.98	12.30	22.20	15.75	12.30	9.85	49.15	95.30	100.00	100.00	100.00
	Based on CS-ARDL estimates of short-run coefficients (ARDL(1,0) specifications with $p_{\bar{z}} = \lceil T^{1/3} \rceil$).																			
30	-26.96	-31.57	-23.02	-21.56	-20.97	145.97	203.84	24.91	23.16	22.29	48.25	60.65	70.50	74.85	79.35	72.25	89.15	98.15	99.60	99.70
50	-25.05	-27.49	-23.46	-21.76	-21.08	424.65	29.74	24.99	22.68	21.90	57.20	76.25	87.55	91.00	94.00	77.80	96.45	99.70	100.00	100.00
100	-35.41	-27.96	-23.42	-22.14	-21.19	148.21	32.71	24.01	22.60	21.62	68.30	92.05	98.20	99.70	99.75	84.25	98.30	100.00	100.00	100.00
150	-96.09	-27.09	-23.65	-22.06	-21.17	2622.39	41.48	24.07	22.37	21.46	72.70	96.05	99.85	99.95	100.00	85.90	99.10	100.00	100.00	100.00
200	-30.44	-28.11	-23.68	-21.92	-21.27	248.48	30.67	24.00	22.18	21.48	73.70	97.70	100.00	100.00	100.00	85.25	98.95	100.00	100.00	100.00

Notes: The dependent variable and regressors are generated according to (36)-(37) with correlated fixed effects, and with cross-sectionally weakly dependent and serially correlated heteroskedastic idiosyncratic innovations generated according to (40)-(41) with $a_{\varepsilon} = 0.6$. The knowledge of lag orders is not used in the estimation stage and the integer part of $T^{1/3}$ gives $3, 3, 4, 5$ and 5 for $T = 30, 50, 100, 150$ and 200, respectively.

Table 6: Monte Carlo Estimates of Bias, RMSE, Size and Power for Estimation of LR Coefficient (θ) in the Case of $\phi_{\max} = 0.9$

DGP is ARDL(2,1) model with heterogeneous coefficients, $\varphi_{\max} = 0.9$, stationary regressors, $m = 2$ factors, no feedback effects and $\rho_{\varepsilon i} = 0$.

	Bias ($\times$100)					**Root Mean Square Errors ($\times$100)**					**Size (5% level, $H_0: \theta = 1$)**					**Power (5% level, $H_1: \theta = 1.2$)**				
(N,T)	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**
	CS-DL mean group ($p_{\bar{y}} = 0$, $p_{\bar{x}} = \lfloor T^{1/3} \rfloor$ and $p = p_{\bar{x}} - 1$)																			
30	-12.05	-11.37	-6.94	-4.56	-4.89	27.31	19.55	14.78	12.60	11.66	9.00	9.85	9.75	8.05	9.45	31.80	53.15	61.35	63.65	73.00
50	-11.64	-10.52	-6.96	-4.69	-4.71	22.01	16.59	12.23	10.16	9.21	9.15	13.45	11.15	7.85	10.40	41.40	69.35	80.35	79.80	89.90
100	-12.19	-10.74	-6.77	-4.77	-4.63	18.30	14.13	9.85	7.95	7.19	16.35	21.95	16.95	11.60	13.30	67.65	92.05	97.05	96.85	99.45
150	-11.60	-10.76	-6.67	-4.79	-4.63	15.71	13.28	8.88	7.07	6.48	19.90	31.35	22.20	15.05	17.95	84.40	98.20	99.70	99.45	100.00
200	-11.87	-10.66	-6.58	-4.88	-4.77	15.17	12.52	8.22	6.67	6.15	24.00	38.55	27.40	19.25	22.05	92.45	99.80	99.95	100.00	100.00
	CS-DL pooled ($p_{\bar{y}} = 0$, $p_{\bar{x}} = \lfloor T^{1/3} \rfloor$ and $p = p_{\bar{x}} - 1$)																			
30	-11.26	-10.51	-6.36	-4.21	-4.53	24.79	18.53	14.13	12.10	11.24	9.00	10.00	9.75	8.25	9.35	34.00	55.35	63.80	65.80	74.20
50	-10.81	-9.80	-6.44	-4.41	-4.45	20.53	15.76	11.48	9.76	8.83	9.60	13.35	10.95	8.35	9.35	47.60	71.95	81.60	82.50	91.20
100	-11.12	-9.97	-6.24	-4.39	-4.29	16.72	13.31	9.26	7.47	6.85	15.10	20.80	15.20	11.10	13.00	73.60	92.65	97.50	97.55	99.20
150	-10.72	-10.13	-6.22	-4.53	-4.33	14.51	12.54	8.31	6.71	6.13	19.60	30.10	20.00	14.40	17.05	88.05	98.60	99.65	99.70	99.85
200	-10.97	-9.94	-6.18	-4.53	-4.41	14.03	11.76	7.82	6.30	5.81	25.35	35.70	26.10	17.80	22.55	94.70	99.80	99.90	99.95	100.00
	Based on CS-ARDL estimates of short-run coefficients (ARDL(2,1) specifications with $p_{\bar{z}} = \lfloor T^{1/3} \rfloor$).																			
30	-21.58	-6.64	-3.38	-1.70	-1.36	438.72	121.05	12.25	9.96	8.51	11.25	11.05	8.35	8.00	7.25	29.00	46.80	61.95	69.25	76.05
50	4.94	-5.49	-3.31	-2.27	-1.61	922.24	34.12	9.95	7.80	6.81	10.25	10.70	8.70	7.55	7.60	32.30	57.05	75.60	85.50	90.60
100	8.35	-3.40	-3.87	-2.43	-1.82	770.17	112.28	7.72	5.92	5.01	10.95	13.90	11.85	9.55	8.00	35.55	71.20	94.40	97.10	99.25
150	-22.01	-7.03	-3.45	-2.61	-1.84	513.85	91.47	6.40	5.07	4.24	10.65	16.85	11.95	11.30	8.75	35.85	78.90	97.40	99.75	99.80
200	41.07	-4.61	-3.61	-2.70	-1.92	2063.94	127.98	5.90	4.57	3.80	12.45	17.90	14.55	12.30	10.25	38.60	82.55	99.35	100.00	100.00
	Based on CS-ARDL estimates of short-run coefficients (ARDL(1,0) specifications with $p_{\bar{z}} = \lfloor T^{1/3} \rfloor$).																			
30	21.99	-27.42	-22.99	-20.75	-20.04	1921.14	70.72	26.25	23.39	22.18	40.55	51.10	57.40	60.25	63.65	60.40	80.05	90.60	94.50	97.40
50	-40.53	-25.17	-23.32	-21.48	-20.71	238.17	219.25	25.35	22.96	21.99	46.45	61.40	71.85	78.45	82.15	65.90	87.55	96.65	98.90	99.65
100	-93.36	-34.65	-23.62	-21.97	-21.09	3030.40	242.83	27.18	22.76	21.72	54.20	75.05	92.25	93.50	97.30	72.30	91.90	99.55	99.80	100.00
150	-33.97	-33.48	-23.56	-21.88	-20.96	472.17	329.85	24.24	22.39	21.39	55.15	81.65	96.60	99.15	99.60	71.60	93.60	99.85	100.00	99.95
200	-37.78	-14.74	-23.75	-22.07	-21.10	247.78	701.23	24.30	22.42	21.41	59.35	86.65	98.80	99.90	100.00	73.45	94.25	100.00	100.00	100.00

Notes: The dependent variable and regressors are generated according to (36)-(37) with correlated fixed effects, and with cross-sectionally weakly dependent and serially correlated heteroskedastic idiosyncratic innovations generated according to (40)-(41) with $a_{\varepsilon} = 0.6$. The knowledge of lag orders is not used in the estimation stage and the integer part of $T^{1/3}$ gives $3, 3, 4, 5$ and 5 for $T = 30, 50, 100, 150$ and 200, respectively.

Table 7: Monte Carlo Estimates of Bias, RMSE, Size and Power for Estimation of LR Coefficient (θ) in the Case of $m = 3$ Factors

DGP is ARDL(2,1) model with heterogeneous coefficients, $\varphi_{\max} = 0.6$, stationary regressors, $m = 3$ factors, no feedback effects and $\rho_{\varepsilon i} = 0$.

	Bias ($\times$100)					Root Mean Square Errors ($\times$100)					Size (5% level, $H_0 : \theta = 1$)					Power (5% level, $H_1 : \theta = 1.2$)				
(N,T)	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**
	CS-DL mean group ($p_{\bar{y}} = 0$, $p_{\bar{x}} = \lfloor T^{1/3} \rfloor$ and $p = p_{\bar{x}} - 1$)																			
30	-1.85	-1.02	0.00	-0.01	-0.02	17.23	11.03	8.20	7.45	6.63	6.70	6.15	6.90	5.75	6.40	29.80	51.70	68.05	78.05	85.25
50	-0.79	-0.63	-0.24	0.03	0.10	12.90	8.79	6.68	5.77	5.04	4.90	6.55	6.35	6.20	5.05	38.65	67.25	85.85	92.70	96.60
100	-1.00	-0.94	-0.23	0.09	-0.15	9.57	6.23	4.73	4.12	3.72	5.70	5.40	5.55	5.20	5.35	61.05	92.80	99.10	99.60	99.95
150	-1.19	-0.88	-0.02	0.00	-0.06	7.77	5.03	3.73	3.34	2.92	6.05	4.85	5.15	5.40	4.60	78.45	98.30	99.90	100.00	100.00
200	-0.99	-0.78	-0.03	-0.03	0.09	6.62	4.50	3.23	2.88	2.61	5.00	5.60	5.20	4.85	4.90	89.15	99.60	100.00	100.00	100.00
	CS-DL pooled ($p_{\bar{y}} = 0$, $p_{\bar{x}} = \lfloor T^{1/3} \rfloor$ and $p = p_{\bar{x}} - 1$)																			
30	-1.59	-0.79	0.08	-0.04	0.01	15.07	10.40	7.96	7.15	6.57	7.05	6.40	7.05	5.95	6.45	34.10	54.25	72.20	80.00	86.40
50	-0.91	-0.57	-0.17	0.03	0.11	11.51	8.27	6.44	5.57	4.93	5.65	6.50	6.55	5.50	5.00	47.50	71.95	88.25	93.90	97.65
100	-1.01	-0.88	-0.21	0.10	-0.16	8.33	5.78	4.51	3.97	3.58	5.55	5.85	5.50	5.35	5.20	72.30	94.90	99.55	99.85	100.00
150	-0.90	-0.72	-0.04	0.00	-0.07	6.78	4.73	3.61	3.28	2.88	6.20	5.15	5.70	5.35	4.80	87.25	99.10	100.00	100.00	100.00
200	-0.96	-0.63	-0.05	0.01	0.08	5.79	4.19	3.08	2.81	2.55	5.25	6.30	5.10	5.40	5.60	94.80	99.85	100.00	100.00	100.00
	Based on CS-ARDL estimates of short-run coefficients (ARDL(2,1) specifications with $p_{\bar{z}} = \lfloor T^{1/3} \rfloor$).																			
30	-5.34	-2.54	-0.89	-0.10	0.34	144.86	11.33	7.34	6.29	5.65	9.10	8.80	7.80	6.70	7.15	37.10	61.60	83.70	88.95	93.35
50	-5.60	-2.52	-1.03	-0.20	0.23	157.20	9.36	5.91	4.86	4.28	9.20	9.80	8.10	6.20	5.85	44.65	76.50	94.60	98.05	99.40
100	-0.04	-3.29	-1.21	-0.23	-0.07	83.58	10.20	4.28	3.43	3.10	9.65	11.05	7.85	6.65	6.25	56.30	95.20	100.00	100.00	100.00
150	-4.99	-2.93	-0.94	-0.44	0.01	94.15	7.97	3.48	2.82	2.54	11.80	13.50	8.20	4.80	5.40	64.25	98.65	99.95	100.00	100.00
200	7.81	1.08	-1.01	-0.36	0.08	383.41	178.92	3.08	2.51	2.23	12.10	15.05	8.15	6.75	6.55	67.70	99.60	100.00	100.00	100.00
	Based on CS-ARDL estimates of short-run coefficients (ARDL(1,0) specifications with $p_{\bar{z}} = \lfloor T^{1/3} \rfloor$).																			
30	-24.61	-25.74	-22.19	-20.87	-20.48	271.79	27.83	23.55	22.09	21.54	57.05	70.00	81.45	84.80	86.65	84.95	97.50	99.85	99.95	100.00
50	-29.47	-25.95	-22.39	-21.20	-20.52	39.80	27.22	23.27	21.90	21.18	70.30	88.00	94.45	97.15	98.05	91.90	99.65	100.00	100.00	100.00
100	-30.85	-26.44	-22.73	-21.26	-20.94	37.04	27.17	23.16	21.63	21.28	85.75	98.60	100.00	99.95	100.00	96.20	99.95	100.00	100.00	100.00
150	-31.83	-26.42	-22.46	-21.61	-20.86	48.65	26.97	22.76	21.85	21.08	89.95	99.95	99.95	100.00	100.00	96.05	99.95	100.00	100.00	100.00
200	-29.13	-26.22	-22.55	-21.46	-20.80	108.14	26.59	22.77	21.66	20.97	91.85	99.85	100.00	100.00	100.00	96.20	100.00	100.00	100.00	100.00

Notes: The dependent variable and regressors are generated according to (36)-(37) with correlated fixed effects, and with cross-sectionally weakly dependent and serially correlated heteroskedastic idiosyncratic innovations generated according to (40)-(41) with $a_\varepsilon = 0.6$. The knowledge of lag orders is not used in the estimation stage and the integer part of $T^{1/3}$ gives $3, 3, 4, 5$ and 5 for $T = 30, 50, 100, 150$ and 200, respectively.

Table 8: Monte Carlo Estimates of Bias, RMSE, Size and Power for Estimation of LR Coefficient (θ) in the Case of Unit Roots in Factors

DGP is ARDL(2,1) model with heterogeneous coefficients, $\varphi_{\max} = 0.6$, unit roots in factors, $m = 2$ factors, no feedback effects and $\rho_{\varepsilon i} = 0$.

	Bias ($\times$100)					Root Mean Square Errors ($\times$100)					Size (5% level, $H_0: \theta = 1$)					Power (5% level, $H_1: \theta = 1.2$)				
(N,T)	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**
	CS-DL mean group ($p_{\bar{y}} = 0$, $p_{\bar{x}} = \lceil T^{1/3} \rceil$ and $p = p_{\bar{x}} - 1$)																			
30	-1.04	-1.04	-0.14	0.06	0.05	16.26	11.33	8.28	7.27	6.55	5.65	6.85	6.60	5.95	6.55	28.50	53.25	71.45	78.30	86.40
50	-0.84	-0.84	-0.30	0.20	-0.16	12.76	8.51	6.56	5.79	5.18	5.15	4.85	5.35	6.40	5.70	38.25	70.50	87.65	92.80	97.10
100	-1.42	-0.99	-0.04	0.03	-0.14	9.37	6.29	4.55	4.05	3.62	5.30	5.50	5.10	5.55	5.40	63.55	92.60	98.75	99.80	100.00
150	-1.15	-0.91	-0.14	-0.08	0.01	7.87	5.26	3.73	3.31	2.91	5.90	6.40	4.75	4.95	4.90	78.30	98.10	99.95	100.00	100.00
200	-1.14	-0.79	-0.21	-0.03	0.03	6.79	4.43	3.24	2.90	2.50	5.50	4.95	5.20	5.10	4.95	88.00	99.65	100.00	100.00	100.00
	CS-DL pooled ($p_{\bar{y}} = 0$, $p_{\bar{x}} = \lceil T^{1/3} \rceil$ and $p = p_{\bar{x}} - 1$)																			
30	-0.69	-0.78	-0.15	0.09	0.04	14.53	10.66	8.01	7.04	6.47	6.25	6.75	6.80	6.15	7.20	33.10	57.20	72.95	80.50	88.00
50	-0.69	-0.81	-0.27	0.19	-0.24	11.42	7.97	6.17	5.62	5.06	5.05	5.25	5.80	6.40	5.60	45.45	74.10	90.60	94.15	97.55
100	-1.17	-0.80	-0.05	0.02	-0.09	8.25	5.77	4.38	4.03	3.60	5.40	5.55	5.80	5.65	5.60	72.60	95.00	99.50	99.90	100.00
150	-0.94	-0.82	-0.14	-0.05	0.03	6.98	4.88	3.55	3.20	2.85	6.05	5.90	5.70	5.15	5.40	86.00	98.90	100.00	100.00	100.00
200	-0.98	-0.70	-0.20	-0.01	0.01	5.96	4.15	3.16	2.83	2.48	5.60	5.45	5.60	5.40	5.20	94.05	99.90	100.00	100.00	100.00
	Based on CS-ARDL estimates of short-run coefficients (ARDL(2,1) specifications with $p_{\bar{z}} = \lceil T^{1/3} \rceil$).																			
30	14.28	-3.42	-1.78	-0.89	-0.48	760.36	14.43	7.34	6.04	5.47	9.40	10.10	8.35	6.75	7.20	38.75	66.45	88.85	93.30	96.40
50	-3.16	-3.92	-2.23	-1.19	-0.94	106.12	9.21	6.14	4.87	4.52	11.35	10.70	10.25	7.65	8.40	49.10	83.25	96.90	99.25	99.55
100	-9.23	-5.09	-2.62	-1.63	-1.20	293.70	7.95	4.67	3.75	3.31	15.00	18.15	12.30	9.90	8.40	65.15	97.95	100.00	100.00	100.00
150	-12.59	-5.44	-2.95	-1.97	-1.39	395.28	7.37	4.33	3.35	2.78	18.40	25.50	18.25	13.50	8.80	73.35	99.55	100.00	100.00	100.00
200	-5.66	-5.90	-3.07	-2.04	-1.47	91.39	7.55	4.15	3.13	2.56	21.95	34.35	22.30	16.80	11.60	77.50	99.90	100.00	100.00	100.00
	Based on CS-ARDL estimates of short-run coefficients (ARDL(1,0) specifications with $p_{\bar{z}} = \lceil T^{1/3} \rceil$).																			
30	-28.65	-26.01	-22.78	-21.75	-21.38	43.69	28.01	24.09	22.81	22.34	54.90	72.60	84.75	89.45	91.55	84.60	97.80	99.90	100.00	100.00
50	-30.99	-26.05	-23.52	-21.99	-21.78	80.99	27.34	24.34	22.65	22.38	67.60	87.65	96.40	98.10	99.00	89.65	99.60	100.00	100.00	100.00
100	-32.82	-27.17	-23.82	-22.65	-21.99	69.50	27.81	24.20	22.98	22.28	86.30	99.45	100.00	100.00	100.00	95.80	99.95	100.00	100.00	100.00
150	-26.37	-27.61	-24.12	-22.92	-22.31	175.08	28.01	24.37	23.13	22.50	92.10	99.95	100.00	100.00	100.00	96.95	100.00	100.00	100.00	100.00
200	-32.57	-27.95	-24.27	-22.94	-22.32	41.83	28.25	24.46	23.11	22.46	93.85	100.00	100.00	100.00	100.00	96.45	100.00	100.00	100.00	100.00

Notes: The dependent variable and regressors are generated according to (36)-(37) with correlated fixed effects, and with cross-sectionally weakly dependent and serially correlated heteroskedastic idiosyncratic innovations generated according to (40)-(41) with $a_{\varepsilon} = 0.6$. The knowledge of lag orders is not used in the estimation stage and the integer part of $T^{1/3}$ gives $3, 3, 4, 5$ and 5 for $T = 30, 50, 100, 150$ and 200, respectively.

Table 9: Monte Carlo Estimates of Bias, RMSE, Size and Power for Estimation of LR Coefficient (θ) in the Case of Unit Roots in Regressor Specific Components

DGP is ARDL(2,1) model with heterogeneous coefficients, $\varphi_{\max} = 0.6$, unit roots in v_{it}, $m = 2$ factors, no feedback effects and $\rho_{\varepsilon i} = 0$.

	Bias ($\times$100)					**Root Mean Square Errors** ($\times$100)					**Size** (5% level, $H_0 : \theta = 1$)					**Power** (5% level, $H_1 : \theta = 1.2$)				
(N,T)	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**
	CS-DL mean group ($p_{\bar{y}} = 0$, $p_{\bar{x}} = \lceil T^{1/3} \rceil$ and $p = p_{\bar{x}} - 1$)																			
30	-2.99	0.37	0.25	0.33	-0.13	65.72	32.60	15.74	10.52	8.06	5.70	5.00	5.45	5.70	5.65	7.85	11.45	29.45	52.30	72.85
50	-0.20	0.66	-0.40	0.24	0.20	51.19	26.47	12.29	8.26	6.42	5.45	6.40	6.00	5.25	5.30	8.30	14.55	42.20	69.35	87.30
100	-0.76	-0.50	-0.09	0.11	0.09	37.01	18.87	9.00	5.88	4.44	4.70	5.65	6.00	5.00	4.45	9.40	22.00	64.90	91.60	99.05
150	-0.85	-0.17	-0.25	-0.06	0.08	30.36	15.18	7.17	4.76	3.72	4.55	5.35	5.80	4.90	5.45	10.75	27.15	81.85	98.60	99.90
200	0.15	-0.14	-0.18	-0.17	0.18	27.24	13.26	6.28	4.21	3.15	5.90	5.55	5.40	6.00	5.00	12.25	34.05	89.65	99.60	100.00
	CS-DL pooled ($p_{\bar{y}} = 0$, $p_{\bar{x}} = \lceil T^{1/3} \rceil$ and $p = p_{\bar{x}} - 1$)																			
30	-1.75	0.62	-0.04	-0.01	-0.06	47.00	24.71	12.38	9.06	7.13	5.20	5.60	6.50	7.40	6.70	8.95	14.60	40.85	64.45	78.95
50	-0.67	0.37	0.00	-0.01	0.02	36.30	19.19	9.39	6.88	5.64	4.95	5.40	5.70	6.45	6.30	10.40	20.25	57.30	82.95	93.15
100	-0.92	-0.95	-0.24	-0.01	-0.01	27.06	13.97	6.90	4.75	4.01	6.15	5.90	5.85	4.95	5.10	14.50	35.45	84.50	98.25	99.50
150	-1.09	-0.34	-0.08	-0.13	0.13	21.26	11.24	5.47	3.98	3.31	4.70	5.55	4.75	5.45	4.85	16.95	46.00	95.00	99.75	100.00
200	-0.05	-0.11	-0.03	-0.10	0.07	18.94	9.81	4.74	3.45	2.76	4.95	5.45	5.00	5.05	5.20	20.10	55.20	98.25	100.00	100.00
	Based on CS-ARDL estimates of short-run coefficients (ARDL(2,1) specifications with $p_{\bar{z}} = \lceil T^{1/3} \rceil$).																			
30	11.59	-2.76	-0.64	-0.09	-0.27	1487.96	52.46	15.63	10.51	8.10	5.00	5.80	5.60	5.70	6.00	7.30	12.85	31.35	55.35	73.00
50	-28.86	-0.82	-0.92	-0.12	-0.02	1128.77	27.42	12.19	8.26	6.34	4.60	6.10	6.70	5.80	5.60	7.95	14.95	44.25	70.55	88.30
100	-6.61	-1.86	-0.68	-0.24	-0.12	361.66	19.50	9.00	5.77	4.47	4.70	5.10	6.35	5.25	5.80	8.75	22.25	67.60	93.20	99.20
150	-61.57	-1.90	-0.82	-0.28	-0.11	2620.58	16.18	7.17	4.67	3.67	4.05	5.40	5.50	4.95	5.80	10.05	30.85	83.80	99.00	99.95
200	2.04	-1.54	-0.68	-0.41	0.00	336.21	14.05	6.24	4.21	3.12	5.05	5.25	5.45	5.55	4.90	11.10	36.55	91.45	99.70	100.00
	Based on CS-ARDL estimates of short-run coefficients (ARDL(1,0) specifications with $p_{\bar{z}} = \lceil T^{1/3} \rceil$).																			
30	-14.85	-9.27	-4.73	-2.76	-2.30	146.76	36.98	17.15	11.44	8.94	6.65	6.45	6.85	7.05	7.30	12.20	17.30	38.80	58.70	76.15
50	-18.57	-8.15	-4.38	-2.42	-1.79	136.51	28.29	13.16	8.92	6.79	6.80	7.15	7.35	6.70	6.15	12.90	22.80	53.80	76.50	91.50
100	-16.57	-8.00	-3.42	-2.03	-1.39	72.70	20.94	9.73	6.19	4.75	6.60	7.90	7.75	6.75	7.15	16.30	33.75	76.10	95.75	99.40
150	-12.87	-7.13	-3.19	-1.80	-1.24	85.37	17.40	7.87	5.12	3.91	8.00	7.75	7.90	6.95	7.00	19.85	43.15	89.60	99.30	100.00
200	-16.82	-6.73	-2.89	-1.82	-1.04	346.63	15.13	6.99	4.62	3.33	8.45	7.75	8.35	7.75	6.15	23.20	51.75	95.40	99.90	100.00

Notes: The dependent variable and regressors are generated according to (36)-(37) with correlated fixed effects, and with cross-sectionally weakly dependent and serially correlated heteroskedastic idiosyncratic innovations generated according to (40)-(41) with $a_\varepsilon = 0.6$. The knowledge of lag orders is not used in the estimation stage and the integer part of $T^{1/3}$ gives $3, 3, 4, 5$ and 5 for $T = 30, 50, 100, 150$ and 200, respectively.

Table 10: Monte Carlo Estimates of Bias, RMSE, Size and Power for Estimation of LR Coefficient (θ) in the Case of Serially Correlated Idiosyncratic Errors

DGP is ARDL(2,1) model with heterogeneous coefficients, $\varphi_{\max} = 0.6$, stationary regressors, $m = 2$ factors, no feedback effects and $\rho_{\varepsilon i} \sim IIDU\,(0, 0.8)$.

	Bias (×100)					Root Mean Square Errors (×100)					Size (5% level, $H_0: \theta = 1$)					Power (5% level, $H_1: \theta = 1.2$)				
(N,T)	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**
	CS-DL mean group ($p_{\bar{y}} = 0$, $p_{\bar{x}} = \lceil T^{1/3} \rceil$ and $p = p_{\bar{x}} - 1$)																			
30	-1.11	-0.45	-0.21	0.01	0.12	21.91	14.75	11.44	10.33	8.89	7.05	6.00	5.40	6.25	5.55	20.20	31.40	47.00	53.70	64.15
50	-0.29	-1.04	0.08	-0.15	-0.04	16.85	11.71	8.88	7.89	6.96	5.25	6.20	5.95	5.00	5.95	24.05	46.65	64.10	71.85	82.20
100	-1.22	-0.75	-0.11	0.03	0.08	12.21	8.14	6.04	5.55	4.97	4.80	5.05	4.85	4.45	5.90	43.90	72.35	90.05	94.50	97.70
150	-1.21	-1.04	-0.07	0.09	-0.03	10.05	6.90	5.22	4.65	4.16	5.05	5.40	5.55	5.35	6.65	58.00	86.40	96.80	99.15	99.70
200	-1.28	-0.80	0.05	0.01	-0.08	8.62	5.89	4.56	3.99	3.45	5.40	5.25	6.25	5.45	4.15	70.00	94.45	99.05	99.95	100.00
	CS-DL pooled ($p_{\bar{y}} = 0$, $p_{\bar{x}} = \lceil T^{1/3} \rceil$ and $p = p_{\bar{x}} - 1$)																			
30	-0.86	-0.35	-0.06	-0.12	0.22	20.29	14.11	11.03	10.01	8.71	6.85	6.30	6.25	6.20	6.30	24.05	33.80	49.75	58.05	67.25
50	-0.35	-0.97	0.10	-0.11	-0.05	15.46	11.22	8.51	7.57	6.73	5.70	5.95	5.10	5.60	5.55	28.65	50.00	67.10	75.40	83.80
100	-1.11	-0.57	-0.11	0.08	0.15	11.05	7.75	5.89	5.28	4.75	5.30	4.50	4.65	5.00	4.55	49.65	75.05	91.10	95.30	98.35
150	-1.00	-0.77	-0.07	0.02	-0.05	9.06	6.50	5.02	4.43	4.01	5.20	5.45	5.55	4.90	6.00	66.10	90.00	97.60	99.45	99.75
200	-1.01	-0.52	0.00	0.04	-0.05	7.93	5.59	4.35	3.82	3.36	5.95	5.30	5.70	4.60	4.75	76.15	95.40	99.60	100.00	100.00
	Based on CS-ARDL estimates of short-run coefficients (ARDL(2,1) specifications with $p_{\bar{z}} = \lceil T^{1/3} \rceil$).																			
30	-15.73	22.36	16.11	16.50	17.10	1331.93	276.49	20.13	19.44	19.45	6.85	14.05	31.85	40.85	50.60	9.65	10.65	11.80	11.75	11.80
50	-1.80	13.61	16.51	16.65	17.00	568.06	81.68	19.19	18.46	18.44	7.50	20.60	46.55	61.45	71.70	8.95	11.60	11.20	12.80	11.60
100	43.49	16.66	15.91	16.56	17.04	1184.81	34.38	17.18	17.47	17.81	10.55	36.15	72.45	88.25	94.65	7.80	10.90	12.10	12.80	14.00
150	21.78	16.67	16.16	16.54	16.87	382.97	38.39	17.12	17.17	17.40	11.05	47.10	88.60	96.60	99.05	7.10	14.00	16.20	15.60	15.90
200	15.42	-93.00	16.23	16.65	16.86	323.97	4968.11	16.94	17.14	17.26	13.40	57.50	95.35	99.35	99.90	5.85	13.25	17.45	17.80	17.90
	Based on CS-ARDL estimates of short-run coefficients (ARDL(1,0) specifications with $p_{\bar{z}} = \lceil T^{1/3} \rceil$).																			
30	16.40	-6.69	16.36	17.54	18.91	697.94	880.16	23.81	22.65	23.23	6.75	7.70	17.15	24.65	30.80	17.05	15.95	12.15	10.70	9.60
50	-20.75	12.88	15.81	17.78	18.31	2302.74	61.28	23.24	20.87	21.06	4.45	9.05	23.80	38.50	44.45	15.95	16.20	11.95	8.50	8.80
100	48.11	12.90	15.75	18.12	18.87	1567.80	164.94	41.04	22.77	20.28	5.10	11.70	46.05	66.15	75.90	17.95	17.55	11.35	8.55	6.65
150	26.90	12.79	16.79	17.82	18.81	1050.80	491.93	18.79	18.97	19.79	3.85	16.40	63.75	83.05	91.10	15.30	18.55	11.85	8.90	8.05
200	0.27	-25.36	16.89	16.07	18.58	354.88	1004.96	19.27	84.64	19.27	3.65	20.55	79.10	92.90	96.70	16.75	17.85	11.20	9.20	7.35

Notes: The dependent variable and regressors are generated according to (36)-(37) with correlated fixed effects, and with cross-sectionally weakly dependent and serially correlated heteroskedastic idiosyncratic innovations generated according to (40)-(41) with $a_{\varepsilon} = 0.6$. The knowledge of lag orders is not used in the estimation stage and the integer part of $T^{1/3}$ gives $3, 3, 4, 5$ and 5 for $T = 30, 50, 100, 150$ and 200, respectively.

Table 11: Monte Carlo Estimates of Bias, RMSE, Size and Power for Estimation of LR Coefficient (θ) in the Case of Breaks in Errors

DGP is ARDL(2,1) model with heterogeneous coefficients, $\varphi_{\max} = 0.6$, stationary regressors, $m = 2$ factors, no feedback effects and breaks in errors.

	Bias (×100)					Root Mean Square Errors (×100)					Size (5% level, $H_0: \theta = 1$)					Power (5% level, $H_1: \theta = 1.2$)				
(N,T)	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**
	CS-DL mean group ($p_{\bar{y}} = 0$, $p_{\bar{x}} = \lceil T^{1/3} \rceil$ and $p = p_{\bar{x}} - 1$)																			
30	-1.72	-1.26	0.09	0.11	0.25	18.91	12.45	9.88	9.03	7.86	7.20	6.45	6.60	6.05	6.15	25.55	44.75	59.60	64.90	72.45
50	-0.83	-0.82	-0.08	0.09	0.05	14.33	10.13	7.55	6.75	6.27	4.90	6.25	5.75	4.80	5.85	33.70	58.60	78.05	83.05	87.85
100	-1.00	-0.91	-0.08	0.10	-0.06	10.50	7.00	5.23	5.00	4.37	5.30	5.15	4.75	5.70	4.95	54.90	85.70	95.85	96.95	99.20
150	-0.88	-0.90	-0.10	0.05	-0.05	8.57	5.62	4.43	4.01	3.63	6.00	5.35	6.05	4.50	4.75	71.40	95.80	99.25	99.70	100.00
200	-0.63	-0.85	-0.29	0.00	-0.01	7.47	4.92	3.72	3.36	3.12	5.50	5.00	4.90	4.55	4.70	80.85	98.15	100.00	100.00	100.00
	CS-DL pooled ($p_{\bar{y}} = 0$, $p_{\bar{x}} = \lceil T^{1/3} \rceil$ and $p = p_{\bar{x}} - 1$)																			
30	-1.57	-1.05	0.01	0.15	0.10	16.82	12.04	9.49	8.78	7.74	7.10	7.10	6.60	6.70	6.80	30.20	48.15	61.50	66.10	74.70
50	-1.13	-0.72	-0.04	0.10	0.06	12.71	9.34	7.25	6.48	6.09	4.80	6.20	5.70	5.05	5.75	41.60	63.65	80.55	84.65	89.50
100	-0.90	-0.78	-0.15	0.03	-0.02	9.11	6.53	5.03	4.80	4.27	5.55	5.45	4.80	5.65	5.50	63.95	89.15	96.85	98.20	99.55
150	-0.83	-0.79	-0.03	0.09	-0.03	7.40	5.37	4.23	3.86	3.52	5.45	5.75	5.55	4.70	4.75	81.20	96.60	99.40	99.85	99.95
200	-0.65	-0.78	-0.27	-0.01	-0.05	6.61	4.71	3.59	3.25	3.01	5.20	4.75	4.90	4.50	4.60	87.80	99.35	99.95	100.00	100.00
	Based on CS-ARDL estimates of short-run coefficients (ARDL(2,1) specifications with $p_{\bar{z}} = \lceil T^{1/3} \rceil$).																			
30	6.24	6.67	9.59	11.95	14.00	143.70	16.06	14.00	14.80	16.05	5.15	7.90	20.45	33.40	47.30	19.30	28.25	32.35	24.85	19.80
50	10.20	7.71	9.03	11.55	13.95	111.76	19.37	11.73	13.27	15.25	5.70	11.20	24.55	46.10	68.65	21.30	33.45	41.95	32.90	22.50
100	30.94	7.07	9.19	11.58	13.88	514.17	26.38	10.71	12.48	14.59	6.45	13.35	44.50	75.60	91.90	24.45	50.40	57.90	48.75	35.30
150	24.76	7.07	8.80	11.62	13.67	876.89	15.90	9.83	12.24	14.13	7.50	18.80	58.30	89.95	98.20	25.80	57.75	74.45	62.40	47.05
200	1.98	6.22	8.85	11.51	13.76	403.92	28.52	9.58	11.97	14.12	7.80	22.55	70.15	96.35	99.70	28.30	66.50	84.55	73.25	54.60
	Based on CS-ARDL estimates of short-run coefficients (ARDL(1,0) specifications with $p_{\bar{z}} = \lceil T^{1/3} \rceil$).																			
30	-28.03	-10.64	-5.13	-1.40	2.17	411.57	21.22	12.54	10.55	10.34	25.20	21.75	13.05	8.55	6.90	54.20	64.00	67.10	60.45	52.10
50	-15.90	-9.72	-5.33	-1.90	2.08	78.66	20.38	11.70	8.89	8.24	28.55	23.30	14.90	9.95	6.65	59.20	75.90	82.00	76.55	65.05
100	-30.60	-11.17	-5.39	-1.63	2.16	353.77	28.48	9.12	5.93	6.09	34.05	33.95	20.45	8.00	6.95	69.50	89.50	95.20	94.20	86.30
150	-15.01	381.68	-5.79	-1.55	2.13	497.45	17476	7.85	5.01	5.18	38.45	41.60	26.80	8.50	7.75	74.20	93.75	98.75	97.80	94.45
200	-14.40	-11.21	-5.69	-1.72	2.24	123.28	26.47	7.33	4.61	4.59	41.95	49.25	30.75	9.95	8.20	76.90	95.45	99.60	99.55	98.75

Notes: The dependent variable and regressors are generated according to (36)-(37) with correlated fixed effects, and with cross-sectionally weakly dependent and serially correlated heteroskedastic idiosyncratic innovations generated according to (40)-(41) with $a_\varepsilon = 0.6$. The knowledge of lag orders is not used in the estimation stage and the integer part of $T^{1/3}$ gives $3, 3, 4, 5$ and 5 for $T = 30, 50, 100, 150$ and 200, respectively.

Table 12: Monte Carlo Estimates of Bias, RMSE, Size and Power for Estimation of LR Coefficient (θ) in the Case of Feedback Effects

DGP is ARDL(2,1) model with heterogeneous coefficients, $\varphi_{\max} = 0.6$, stationary regressors, $m = 2$ factors, $\kappa_{yi} \sim IIDU(0, 0.2)$, and $\rho_{\varepsilon i} = 0$.

	Bias ($\times$100)					Root Mean Square Errors ($\times$100)					Size (5% level, $H_0: \theta = 1$)					Power (5% level, $H_1: \theta = 1.2$)				
(N,T)	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**
	CS-DL mean group ($p_{\bar{y}} = 0$, $p_{\bar{x}} = \lfloor T^{1/3} \rfloor$ and $p = p_{\bar{x}} - 1$)																			
30	-1.51	2.38	5.10	5.74	5.84	16.20	10.51	9.35	9.16	8.54	6.00	6.60	12.50	15.70	18.20	32.55	43.95	53.10	57.00	65.75
50	-1.45	2.38	5.11	5.48	6.06	12.54	8.54	7.91	7.68	7.79	6.50	8.10	14.45	19.00	25.55	45.25	62.35	70.55	77.40	81.15
100	-0.97	2.85	5.16	5.61	6.05	8.88	6.34	6.70	6.85	6.93	5.80	7.90	23.30	31.85	43.05	69.00	85.35	92.80	95.60	97.85
150	-1.23	2.64	5.05	5.68	6.15	7.50	5.41	6.12	6.45	6.76	5.80	8.70	30.90	46.05	59.55	84.60	95.60	98.90	99.50	99.90
200	-1.46	2.55	4.91	5.61	6.03	6.57	4.72	5.75	6.26	6.50	6.30	9.90	37.55	55.20	71.25	92.90	98.75	99.90	99.95	100.00
	CS-DL pooled ($p_{\bar{y}} = 0$, $p_{\bar{x}} = \lfloor T^{1/3} \rfloor$ and $p = p_{\bar{x}} - 1$)																			
30	2.28	4.73	6.80	7.10	7.00	14.57	10.87	10.37	10.04	9.43	7.25	8.90	17.20	19.90	21.25	30.75	40.75	45.90	50.95	57.55
50	2.26	4.83	6.77	6.89	7.40	11.41	9.10	8.96	8.75	8.94	6.55	10.70	21.25	26.60	33.25	41.50	54.60	61.40	69.95	72.05
100	2.96	5.30	6.89	7.10	7.42	8.53	7.65	8.07	8.11	8.17	7.20	16.60	37.05	47.05	58.10	61.00	75.60	85.40	89.95	92.90
150	2.70	5.07	6.72	7.20	7.50	7.21	6.83	7.55	7.84	8.03	8.05	20.65	50.65	63.80	75.60	76.60	90.55	95.50	97.85	98.55
200	2.47	5.10	6.68	7.15	7.36	6.21	6.40	7.30	7.68	7.78	7.60	25.45	62.00	75.80	85.45	87.80	95.70	99.05	99.50	99.80
	Based on CS-ARDL estimates of short-run coefficients (ARDL(2,1) specifications with $p_{\bar{z}} = \lfloor T^{1/3} \rfloor$).																			
30	-49.31	-6.64	-2.84	-1.50	-1.23	1613.71	11.90	7.13	6.12	5.29	17.15	14.35	8.40	8.70	7.50	54.75	80.05	93.55	95.05	97.95
50	15.94	-7.02	-3.05	-2.03	-1.28	1208.10	10.96	6.01	4.85	4.27	19.45	19.25	11.10	8.95	7.75	65.05	93.30	99.15	99.90	99.70
100	-17.83	-6.54	-3.06	-2.00	-1.50	120.30	9.22	4.83	3.81	3.25	26.80	26.90	15.40	12.55	10.15	78.50	99.35	100.00	100.00	100.00
150	-14.20	-6.84	-3.24	-2.06	-1.45	127.42	8.28	4.40	3.27	2.74	32.55	35.75	21.15	13.70	9.90	82.20	99.90	100.00	100.00	100.00
200	-17.89	-7.08	-3.39	-2.15	-1.57	283.24	8.22	4.28	3.15	2.56	37.60	47.30	27.80	19.30	12.55	83.90	99.85	100.00	100.00	100.00
	Based on CS-ARDL estimates of short-run coefficients (ARDL(1,0) specifications with $p_{\bar{z}} = \lfloor T^{1/3} \rfloor$).																			
30	-37.31	-29.77	-24.72	-23.08	-22.33	236.77	31.59	25.86	24.09	23.27	74.60	86.65	91.90	93.30	94.25	93.00	99.70	100.00	100.00	100.00
50	-39.45	-29.96	-25.01	-23.50	-22.67	100.20	30.98	25.63	24.07	23.19	88.15	97.15	99.00	99.55	99.80	96.60	100.00	100.00	100.00	100.00
100	-38.09	-30.16	-25.06	-23.56	-22.75	60.47	30.71	25.39	23.86	23.02	94.90	99.80	100.00	100.00	100.00	97.70	99.95	100.00	100.00	100.00
150	-37.23	-30.28	-25.21	-23.61	-22.85	52.96	30.76	25.43	23.80	23.03	95.80	99.85	100.00	100.00	100.00	97.80	99.90	100.00	100.00	100.00
200	-39.43	-30.46	-25.40	-23.73	-22.89	54.51	30.72	25.56	23.88	23.03	95.80	99.90	100.00	100.00	100.00	97.85	99.95	100.00	100.00	100.00

Notes: The dependent variable and regressors are generated according to (36)-(37) with correlated fixed effects, and with cross-sectionally weakly dependent and serially correlated heteroskedastic idiosyncratic innovations generated according to (40)-(41) with $a_{\varepsilon} = 0.6$. The knowledge of lag orders is not used in the estimation stage and the integer part of $T^{1/3}$ gives $3, 3, 4, 5$ and 5 for $T = 30, 50, 100, 150$ and 200, respectively.

Table 13: Monte Carlo Estimates of Bias, RMSE, Size and Power for Estimation of $\varphi_1 = E(\varphi_{i1})$

DGP is ARDL(1,0) model with homogeneous long-run, heterogeneous short-run, $\varphi_{\max} = 0.6$, stationary regressors, $m = 2$ factors, no feedback effects and $\rho_{\varepsilon i} = 0$.

	Bias (×100)					Root Mean Square Errors (×100)					Size (5% level, $H_0 : \varphi_1 = 0.3$)					Power (5% level, $H_1 : \varphi_1 = 0.4$)				
(N,T)	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**	**30**	**50**	**100**	**150**	**200**
	Imposing CS-DL pooled estimate of long-run coefficient																			
30	-8.38	-4.35	-1.90	-1.01	-0.86	10.26	6.22	4.18	3.68	3.64	51.45	27.60	12.50	9.10	9.35	96.25	95.00	93.45	90.70	90.80
50	-8.89	-4.78	-2.13	-1.36	-0.98	10.00	5.88	3.55	3.13	2.86	70.70	42.25	16.15	12.35	9.25	99.50	99.70	99.60	98.85	99.10
100	-9.27	-4.92	-2.30	-1.49	-1.15	9.85	5.51	3.10	2.44	2.24	91.55	64.60	24.90	15.75	11.70	100.00	100.00	100.00	100.00	100.00
150	-9.36	-5.10	-2.40	-1.48	-1.16	9.79	5.48	2.92	2.18	1.97	98.15	82.40	36.15	19.75	14.55	100.00	100.00	100.00	100.00	100.00
200	-9.37	-5.09	-2.36	-1.55	-1.10	9.72	5.40	2.76	2.08	1.74	99.15	89.95	45.75	24.10	16.00	100.00	100.00	100.00	100.00	100.00
	Infeasible estimator: Imposing knowledge of long run coefficients																			
30	-7.95	-4.01	-1.77	-0.94	-0.78	9.48	5.73	4.00	3.53	3.53	48.40	23.30	11.55	8.25	8.50	97.25	95.85	93.10	91.65	91.15
50	-8.42	-4.41	-1.98	-1.29	-0.88	9.31	5.43	3.38	3.01	2.77	70.60	36.55	14.45	11.00	8.00	99.80	99.85	99.65	98.90	99.20
100	-8.73	-4.50	-2.13	-1.44	-1.07	9.21	5.05	2.91	2.36	2.15	93.00	60.65	22.30	14.25	10.85	100.00	100.00	100.00	100.00	100.00
150	-8.77	-4.67	-2.21	-1.41	-1.08	9.13	5.03	2.71	2.10	1.88	98.20	77.70	31.50	18.25	13.00	100.00	100.00	100.00	100.00	100.00
200	-8.83	-4.67	-2.18	-1.48	-1.01	9.12	4.95	2.57	2.00	1.65	99.30	86.65	38.80	22.30	14.10	100.00	100.00	100.00	100.00	100.00
	Unconstrained CS-ARDL approach																			
30	-12.76	-6.34	-2.78	-1.58	-1.20	13.99	7.73	4.63	3.83	3.70	74.60	42.30	16.65	10.75	10.50	99.30	98.15	95.80	93.50	92.40
50	-13.34	-6.79	-3.04	-1.96	-1.37	14.11	7.59	4.14	3.40	3.01	92.00	63.95	24.05	14.95	10.55	99.95	99.95	99.80	99.30	99.60
100	-13.71	-6.96	-3.22	-2.13	-1.57	14.12	7.38	3.82	2.87	2.47	99.45	88.25	40.65	23.45	15.25	100.00	100.00	100.00	100.00	100.00
150	-13.77	-7.14	-3.30	-2.09	-1.59	14.10	7.41	3.68	2.63	2.23	99.85	96.95	57.80	30.35	21.40	100.00	100.00	100.00	100.00	100.00
200	-13.83	-7.13	-3.27	-2.17	-1.52	14.10	7.35	3.56	2.56	2.02	100.00	99.55	67.90	40.00	23.70	100.00	100.00	100.00	100.00	100.00

Notes: The dependent variable and regressors are generated according to (36)-(37) with correlated fixed effects, and with cross-sectionally weakly dependent and serially correlated heteroskedastic idiosyncratic innovations generated according to (40)-(41) with $a_{\varepsilon} = 0.6$.

Table 14: List of the 40 Countries in the Sample

Europe	**MENA Countries**	**Asia Pacific**	**Latin America**
Austria	Egypt	Australia	Argentina
Belgium	Iran	China	Brazil
Finland	Morocco	India	Chile
France	Syria	Indonesia	Ecuador
Germany	Tunisia	Japan	Peru
Italy	Turkey	Korea	Venezuela
Netherlands		Malaysia	
Norway	**North America**	New Zealand	**Rest of Africa**
Spain	Canada	Philippines	Nigeria
Sweden	Mexico	Singapore	South Africa
Switzerland	United States	Thailand	
United Kingdom			

Table 15: Fixed Effects (FE) Estimates of the Long-Run Effects Based on the ARDL Approach, 1966-2010

	ARDL (1 lag)			ARDL (2 lags)			ARDL (3 lags)		
	(a)	*(b)*	*(c)*	*(a)*	*(b)*	*(c)*	*(a)*	*(b)*	*(c)*
$\widehat{\theta}_{\Delta d}$	-0.075***		-0.069***	-0.061***		-0.054***	-0.055***		-0.044***
	(0.009)		(0.008)	(0.010)		(0.009)	(0.016)		(0.013)
$\widehat{\theta}_{\pi}$		-0.025***	-0.025***		-0.025***	-0.026***		-0.025***	-0.024***
		(0.007)	(0.004)		(0.007)	(0.006)		(0.008)	(0.006)
$\widehat{\lambda}$	-0.854***	-0.790***	-0.876***	-0.839***	-0.771***	-0.861***	-0.768***	-0.723***	-0.771***
	(0.052)	(0.064)	(0.051)	(0.045)	(0.051)	(0.048)	(0.045)	(0.039)	(0.049)
CD test statistics	24.52	34.72	26.35	23.20	34.90	24.96	21.85	32.68	23.31
$N \times T$	1642	1725	1642	1602	1685	1602	1562	1645	1562

Notes: The ARDL specification is given by: $\Delta y_{it} = c_i + \sum_{\ell=1}^{p} \varphi_{i\ell} \Delta y_{i,t-\ell} + \sum_{\ell=0}^{p} \boldsymbol{\beta}_{i\ell}' \mathbf{x}_{i,t-\ell} + u_{it}$, where y_{it} is the log of real GDP, $\mathbf{x}_{it} = (\Delta d_{it}, \pi_{it})'$, d_{it} is the log of debt to GDP ratio, π_{it} is the inflation rate, and $p = 1, 2,$ and 3. $\lambda_i = 1 - \sum_{\ell=1}^{p} \varphi_{i\ell}$, and $\boldsymbol{\theta}_i = \lambda_i^{-1} \sum_{\ell=0}^{p} \boldsymbol{\beta}_{i\ell}$. The reported standard errors are robust to cross-sectional heteroskedasticity and residual serial correlation as in Arellano (1987). Symbols ***, **, and * denote significance at 1%, 5%, and at 10% respectively.

Table 16: Mean Group (MG) Estimates of the Long-Run Effects Based on the ARDL Approach, 1966-2010

	ARDL (1 lag)			ARDL (2 lags)			ARDL (3 lags)		
	(a)	*(b)*	*(c)*	*(a)*	*(b)*	*(c)*	*(a)*	*(b)*	*(c)*
$\widehat{\theta}_{\Delta d}$	-0.070***	–	-0.070***	-0.061***	–	-0.076***	-0.066***	–	-0.083***
	(0.015)		(0.012)	(0.014)		(0.013)	(0.016)		(0.014)
$\widehat{\theta}_{\pi}$	–	-0.104***	-0.038*	–	-0.054**	0.021	–	-0.091***	0.040
		(0.021)	(0.023)		(0.024)	(0.030)		(0.032)	(0.040)
$\widehat{\lambda}$	-0.791***	-0.764***	-0.811***	-0.836***	-0.742***	-0.809***	-0.769***	-0.687***	-0.761***
	(0.028)	(0.037)	(0.030)	(0.039)	(0.044)	(0.047)	(0.043)	(0.041)	(0.053)
CD test statistics	19.15	33.62	21.39	16.99	31.21	16.63	16.42	30.39	15.98
$N \times T$	1642	1725	1642	1602	1685	1602	1562	1645	1562

Notes: The ARDL specification is given by: $\Delta y_{it} = c_i + \sum_{\ell=1}^{p} \varphi_{i\ell} \Delta y_{i,t-\ell} + \sum_{\ell=0}^{p} \boldsymbol{\beta}_{i\ell}' \mathbf{x}_{i,t-\ell} + u_{it}$, where y_{it} is the log of real GDP, $\mathbf{x}_{it} = (\Delta d_{it}, \pi_{it})'$, d_{it} is the log of debt to GDP ratio, π_{it} is the inflation rate, and $p = 1, 2,$ and 3. $\lambda_i = 1 - \sum_{\ell=1}^{p} \varphi_{i\ell}$, and $\boldsymbol{\theta}_i = \lambda_i^{-1} \sum_{\ell=0}^{p} \boldsymbol{\beta}_{i\ell}$. Symbols ***, **, and * denote significance at 1%, 5%, and at 10% respectively.

Table 17: Mean Group Estimates of the Long-Run Effects Based on the Cross-Sectionally Augmented ARDL (CS-ARDL) Approach, 1966-2010

	CS-ARDL (1 lag)			CS-ARDL (2 lags)			CS-ARDL (3 lags)		
	(a)	*(b)*	*(c)*	*(a)*	*(b)*	*(c)*	*(a)*	*(b)*	*(c)*
$\widehat{\theta}_{\Delta d}$	-0.087***	–	-0.087***	-0.090***	–	-0.079***	-0.096***	–	-0.120***
	(0.013)		(0.016)	(0.013)		(0.022)	(0.016)		(0.040)
$\widehat{\theta}_{\pi}$	–	-0.083**	-0.164***	–	-0.071**	-0.110***	–	-0.065	-0.080
		(0.034)	(0.038)		(0.031)	(0.035)		(0.041)	(0.059)
$\widehat{\lambda}$	-0.889***	-0.790***	-0.952***	-0.967***	-0.817***	-1.058***	-0.920***	-0.792***	-1.210***
	(0.031)	(0.041)	(0.039)	(0.042)	(0.053)	(0.053)	(0.047)	(0.058)	(0.201)
CD test statistics	-0.94	-0.30	0.55	-0.43	0.02	-0.11	-0.21	0.05	-0.56
$N \times T$	1599	1657	1599	1581	1652	1581	1562	1645	1562

Notes: The cross-sectionally augmented ARDL (CS-ARDL) regressions include the cross-sectional average of the dependent variable and the regressors together with three lags of these cross-sectional averages. The cross-sectionally augmented ARDL specification is given by: $\Delta y_{it} = c_i + \sum_{\ell=1}^{p} \varphi_{i\ell} \Delta y_{i,t-\ell} + \sum_{\ell=0}^{p} \boldsymbol{\beta}_{i\ell}' \mathbf{x}_{i,t-\ell} + \sum_{\ell=0}^{3} \boldsymbol{\psi}_{i\ell}' \overline{\mathbf{z}}_{t-\ell} + e_{it}$, where $\mathbf{x}_{it} = (\Delta d_{it}, \pi_{it})'$, $\overline{\mathbf{z}}_t = \left(\overline{\Delta y}_t, \overline{\mathbf{x}}_t'\right)'$, $\lambda_i = 1 - \sum_{\ell=1}^{p} \varphi_{i\ell}$ and $\boldsymbol{\theta}_i = \lambda_i^{-1} \sum_{\ell=0}^{p} \boldsymbol{\beta}_{i\ell}$, and $p = 1, 2$ and 3. See also the notes to Table 15.

Table 18: Mean Group Estimates of the Long-Run Effects Based on the Cross-Sectionally Augmented Distributed Lag (CS-DL) Approach, 1966-2010

	CS-DL (1 lag)			CS-DL (2 lags)			CS-DL (3 lags)		
	(a)	*(b)*	*(c)*	*(a)*	*(b)*	*(c)*	*(a)*	*(b)*	*(c)*
$\widehat{\theta}_{\Delta d}$	-0.084***	–	-0.087***	-0.078***	–	-0.084***	-0.068***	–	-0.082***
	(0.013)		(0.014)	(0.014)		(0.017)	(0.014)		(0.020)
$\widehat{\theta}_{\pi}$	–	-0.066***	-0.089***	–	-0.072***	-0.086**	–	-0.072**	-0.086**
		(0.021)	(0.026)		(0.024)	(0.037)		(0.030)	(0.040)
CD test statistics	-1.54	-0.21	1.16	-1.23	0.17	0.73	-1.09	-0.46	0.63
$N \times T$	1601	1661	1601	1586	1661	1586	1571	1660	1571

Notes: The cross-sectionally augmented distributed lag (CS-DL) regressions include the cross-sectional average of the dependent variable and three lags for the cross-sectional averages of the regressors. The estimates are based on the following specification: $\Delta y_{it} = c_i + \boldsymbol{\theta}_i' \mathbf{x}_{it} + \sum_{\ell=0}^{p-1} \boldsymbol{\delta}_{i\ell}' \Delta \mathbf{x}_{i,t-\ell} + \omega_{iy} \overline{\Delta y}_t + \sum_{\ell=0}^{3} \boldsymbol{\omega}_{i,x\ell}' \overline{\mathbf{x}}_{t-\ell} + e_{it}$. See also the notes to Table 15.

Table 19: Mean Group Estimates of the Long-Run Effects of the Log of Debt/GDP ratio and Inflation/CPI on the Log of Output Based on the Cross-Sectionally Augmented Distributed Lag (CS-DL) Approach, 1965-2010

	CS-DL (1 lag)		CS-DL (2 lags)		CS-DL (3 lags)	
	(i)	*(ii)*	*(i)*	*(ii)*	*(i)*	*(ii)*
$\widehat{\theta}_d$	-0.068***	-0.075***	-0.057***	-0.066***	-0.048*	-0.051*
	(0.018)	(0.020)	(0.019)	(0.024)	(0.025)	(0.027)
$\widehat{\theta}_\pi$	0.095	–	0.057	–	0.029	–
	(0.075)		(0.102)		(0.128)	
$\widehat{\theta}_p$	–	-0.008	–	-0.001	–	-0.008
		(0.042)		(0.052)		(0.057)
$N \times T$	1618	1641	1603	1626	1588	1611

Notes: The cross-sectionally augmented distributed lag (CS-DL) regressions include the cross-sectional average of the dependent variable and three lags for the cross-sectional averages of the regressors. The estimates are based on the following specification: $y_{it} = c_i + \boldsymbol{\theta}_i^{'}\mathbf{x}_{it} + \sum_{\ell=0}^{p-1} \boldsymbol{\delta}_{i\ell}^{'}\Delta\mathbf{x}_{i,t-\ell} + \omega_{iy}\overline{y}_t + \sum_{\ell=0}^{3} \boldsymbol{\omega}_{i,x\ell}^{'}\overline{\mathbf{x}}_{t-\ell} + e_{it}$, where in *(i)* y_{it} is the log of real GDP, $\mathbf{x}_{it} = (d_{it}, \pi_{it})'$, d_{it} is the log of the debt/GDP ratio, and π_{it} is the inflation rate and in *(ii)* y_{it} is the log of real GDP, $\mathbf{x}_{it} = (d_{it}, p_{it})'$, d_{it} is the log of the debt/GDP ratio, and p_{it} is the log of the CPI. See also the notes to Table 15.

Table 20: Estimates of the Average Threshold Effects on Output Growth, 1966-2010

τ	30	40	50	60	70	80	90
(i) Pooled OLS Estimates with $I_{it}(\tau)$, where $I_{it}(\tau) = I(d_{it} \geq \log(\tau))$							
$\widehat{\gamma}_\tau$	-0.008***	-0.009***	-0.009***	-0.009***	-0.009***	-0.009***	-0.011***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)
$\hat{c}_\tau$	0.043***	0.042***	0.041***	0.040***	0.039***	0.039***	0.039***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
N	40	40	40	40	40	40	40
$N \times T$	1696	1696	1696	1696	1696	1696	1696
(ii) Mean Group Estimates with $I_{it}(\tau)$							
$\widehat{\gamma}_\tau$	-0.008**	-0.010***	-0.012***	-0.011***	-0.016***	-0.020***	-0.021***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
$\hat{c}_\tau$	0.045***	0.046***	0.043***	0.041***	0.041***	0.044***	0.048***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
N	32	36	31	31	28	19	14
$N \times T$	1353	1531	1322	1332	1203	810	589
(iii) CS-DL Mean Group Estimates (3 lags) including $I_{it}(\tau)$							
$\widehat{\gamma}_\tau$	-0.006	-0.004	-0.008	-0.005	-0.009	-0.001	-0.006
	(0.009)	(0.005)	(0.009)	(0.006)	(0.006)	(0.009)	(0.007)
$\widehat{\theta}_{\tau,\Delta d}$	-0.071***	-0.087***	-0.076***	-0.063**	-0.076***	-0.089***	-0.109***
	(0.024)	(0.022)	(0.025)	(0.026)	(0.025)	(0.031)	(0.037)
$\widehat{\theta}_{\tau,\pi}$	-0.095*	-0.062	-0.090*	-0.079	-0.161***	-0.138**	-0.142
	(0.050)	(0.045)	(0.052)	(0.049)	(0.053)	(0.061)	(0.110)
N	32	35	31	31	28	18	14
$N \times T$	1251	1377	1226	1236	1115	710	547

Notes: The estimates are based on the following specifications:

$$\begin{aligned}
(i)\quad \Delta y_{it} &= c_\tau + \gamma_\tau I_{it}(\tau) + e_{it},\\
(ii)\quad \Delta y_{it} &= c_{i\tau} + \gamma_{i\tau} I_{it}(\tau) + e_{it},\\
(iii)\quad \Delta y_{it} &= c_{i\tau} + \gamma_{i\tau} I_{it}(\tau) + \boldsymbol{\theta}_i' \mathbf{x}_{it} + \sum_{\ell=0}^{2} \boldsymbol{\delta}_{i\ell}' \Delta \mathbf{x}_{i,t-\ell} + \omega_{iy} \overline{\Delta y}_t + \sum_{\ell=0}^{3} \boldsymbol{\omega}_{i,x\ell}' \overline{\mathbf{x}}_{t-\ell} + e_{it},
\end{aligned}$$

where $I_{it}(\tau) = I(d_{it} \geq \log(\tau))$, y_{it} is the log of real GDP, $\mathbf{x}_{it} = (\Delta d_{it}, \pi_{it})'$, d_{it} is the log of the debt GDP ratio, and π_{it} is the inflation rate. The cross-sectionally augmented distributed lag (CS-DL) regression (iii) include the cross-sectional average of the dependent variable and three lags for the cross-sectional averages of the regressors. We report heteroscedasticity-robust standard errors for specification (i). See also the notes to Table 15.

Table 21: Estimates of the Average Threshold Effects on Output Growth Based on the Cross-Sectionally Augmented Distributed Lag (CS-DL) Approach with Three Lags, 1966-2010

τ	30%	40%	50%	60%	70%	80%	90%
(iv) With $I_{it}(\tau)$ and $I_{it}(\tau) \times \max(0, \Delta d_{it})$							
$\widehat{\gamma}_{\tau}$	0.002	0.001	-0.006	-0.005	-0.018	-0.009	-0.001
	(0.005)	(0.005)	(0.007)	(0.006)	(0.011)	(0.015)	(0.018)
$\widehat{\gamma}_{\tau}^{+}$	-0.005	0.018	-0.028	-0.116***	-0.127	-0.192**	-0.140**
	(0.025)	(0.024)	(0.038)	(0.045)	(0.080)	(0.094)	(0.062)
$\widehat{\theta}_{\tau,\Delta d}$	-0.085***	-0.100***	-0.079***	-0.050*	-0.064**	-0.088***	-0.100***
	(0.031)	(0.025)	(0.028)	(0.027)	(0.028)	(0.034)	(0.038)
$\widehat{\theta}_{\tau,\pi}$	-0.119**	-0.073	-0.099**	-0.085**	-0.155***	-0.125*	-0.118
	(0.047)	(0.047)	(0.049)	(0.039)	(0.057)	(0.064)	(0.103)
N	30	33	31	31	25	18	14
$N \times T$	1184	1310	1226	1236	999	710	547
(v) With $I_{it}(\tau) \times \max(0, \Delta d_{it})$							
$\widehat{\gamma}_{\tau}^{+}$	-0.001	-0.001	-0.060	-0.113***	-0.158***	-0.171***	-0.159***
	(0.024)	(0.024)	(0.040)	(0.044)	(0.057)	(0.052)	(0.046)
$\widehat{\theta}_{\tau,\Delta d}$	-0.090***	-0.100***	-0.069***	-0.056**	-0.070***	-0.066**	-0.080**
	(0.025)	(0.024)	(0.025)	(0.024)	(0.021)	(0.028)	(0.035)
$\widehat{\theta}_{\tau,\pi}$	-0.087**	-0.083**	-0.085*	-0.096**	-0.135***	-0.061	-0.031
	(0.037)	(0.040)	(0.045)	(0.042)	(0.049)	(0.058)	(0.080)
N	38	36	32	31	28	18	14
$N \times T$	1487	1414	1263	1236	1115	710	547

Notes: The estimates are based on the following specifications:

$$
\begin{aligned}
(i) \quad \Delta y_{it} &= c_{i\tau} + \gamma_{i\tau} I_{it}(\tau) + \gamma_{i\tau}^{+}\left[I_{it}(\tau) \times \max(0, \Delta d_{it})\right] + \boldsymbol{\theta}_{i,\tau}^{'} \mathbf{x}_{it} + \sum_{\ell=0}^{2} \boldsymbol{\delta}_{i\ell,\tau}^{'} \Delta \mathbf{x}_{i,t-\ell} \\
&\quad + \omega_{iy,\tau} \overline{\Delta y}_{t} + \sum_{\ell=0}^{3} \boldsymbol{\omega}_{i,x\ell,\tau}^{'} \overline{\mathbf{x}}_{t-\ell} + e_{it}, \\
(ii) \quad \Delta y_{it} &= c_{i\tau} + \gamma_{i\tau}^{+}\left[I_{it}(\tau) \times \max(0, \Delta d_{it})\right] + \boldsymbol{\theta}_{i,\tau}^{'} \mathbf{x}_{it} + \sum_{\ell=0}^{2} \boldsymbol{\delta}_{i\ell,\tau}^{'} \Delta \mathbf{x}_{i,t-\ell} \\
&\quad + \omega_{iy,\tau} \overline{\Delta y}_{t} + \sum_{\ell=0}^{3} \boldsymbol{\omega}_{i,x\ell,\tau}^{'} \overline{\mathbf{x}}_{t-\ell} + e_{it},
\end{aligned}
$$

where $I_{it}(\tau) = I(d_{it} \geq \log(\tau))$, y_{it} is the log of real GDP, $\mathbf{x}_{it} = (\Delta d_{it}, \pi_{it})'$, d_{it} is the log of the debt GDP ratio, and π_{it} is the inflation rate. The cross-sectionally augmented distributed lag (CS-DL) regression (iii) include the cross-sectional average of the dependent variable and three lags for the cross-sectional averages of the regressors. See also the notes to Table 15.

References

Akerlof, G. A., W. T. Dickens, and G. L. Perry (2000). Near-rational Wage and Price Setting and the Long-run Phillips Curve. *Brookings Papers on Economic Activity 2000*(1), 1–44.

Arellano, M. (1987). Practitioner's Corner: Computing Robust Standard Errors for Within-groups Estimators. *Oxford Bulletin of Economics and Statistics 49*(4), 431–434.

Bai, J. (2009). Panel Data Models with Interactive Fixed Effects. *Econometrica 77*, 1229–1279.

Barro, R. J. (2001, May). Human Capital and Growth. *American Economic Review 91*(2), 12–17.

Bayoumi, T. and J. Gagnon (1996). Taxation and Inflation: A New Explanation for Capital Flows. *Journal of Monetary Economics 38*(2), 303–330.

Bruno, M. and W. Easterly (1998). Inflation Crises and Long-run Growth. *Journal of Monetary Economics 41*(1), 3 – 26.

Bullard, J. and J. W. Keating (1995). The Long-run Relationship between Inflation and Output in Postwar Economies. *Journal of Monetary Economics 36*(3), 477 – 496.

Chari, V. V., L. E. Jones, and R. E. Manuelli (1996). Inflation, Growth, and Financial Intermediation. *Federal Reserve Bank of St. Louis Review May*, 41–58.

Checherita-Westphal, C. and P. Rother (2012). The impact of High Government Debt on Economic Growth and its Channels: An Empirical Investigation for the Euro Area. *European Economic Review 56*(7), 1392 – 1405.

Chudik, A. and M. H. Pesaran (2013a). Common Correlated Effects Estimation of Heterogeneous Dynamic Panel Data Models with Weakly Exogenous Regressors. CESifo Working Paper No. 4232.

Chudik, A. and M. H. Pesaran (2013b). Large Panel Data Models with Cross-Sectional Dependence: A Survey. In B. H. Baltagi (Ed.), *forthcoming in The Oxford Handbook on Panel Data*. Oxford University Press.

Cochrane, J. H. (2011a). Understanding Policy in the Great Recession: Some Unpleasant Fiscal Arithmetic. *European Economic Review 55*(1), 2–30.

Cochrane, J. H. (2011b). Inflation and Debt. *National Affairs (Washington/DC),(Fall 2011) 9*, 56–78.

De Gregorio, J. (1993). Inflation, Taxation, and Long-run Growth. *Journal of Monetary Economics 31*(3), 271–298.

DeLong, J. B. and L. H. Summers (2012). Fiscal Policy in a Depressed Economy. *Brookings Papers on Economic Activity*, 233–297.

Dornbusch, R. and J. A. Frenkel (1973). Inflation and Growth: Alternative Approaches. *Journal of Money, Credit and Banking 5*(1), 141–156.

Eberhardt, M. and A. F. Presbitero (2013). This Time They Are Different: Heterogeneity and Nonlinearity in the Relationship between Debt and Growth. Mimeo, October 2013.

Elmendorf, D. W. and G. N. Mankiw (1999). Government Debt. In J. B. Taylor and M. Woodford (Eds.), *Handbook of Macroeconomics*, Volume 1, Part C, pp. 1615–1669. Elsevier.

Ericsson, N. R., J. S. Irons, and R. W. Tryon (2001). Output and Inflation in the Long Run. *Journal of Applied Econometrics 16*(3), 241–253.

Ghosh, A. and S. Phillips (1998). Warning: Inflation May Be Harmful to Your Growth. *IMF Staff Papers 45*(4), 672–710.

Ghosh, A. R., J. I. Kim, E. G. Mendoza, J. D. Ostry, and M. S. Qureshi (2013). Fiscal Fatigue, Fiscal Space and Debt Sustainability in Advanced Economies. *The Economic Journal 123*(566), F4–F30.

Gillman, M. and M. Kejak (2005). Contrasting Models of the Effect of Inflation on Growth. *Journal of Economic Surveys 19*(1), 113–136.

Gillman, M. and A. Nakov (2004). Granger Causality of the Inflation-Growth Mirror in Accession Countries. *Economics of Transition 12*(4), 653–681.

Gomme, P. (1993). Money and Growth Revisited: Measuring the Costs of Inflation in an Endogenous Growth Model. *Journal of Monetary Economics 32*(1), 51–77.

Gylfason, T. and T. T. Herbertsson (2001). Does Inflation Matter for Growth? *Japan and the World Economy 13*(4), 405 – 428.

Ireland, P. N. (1994). Money and Growth: An Alternative Approach. *The American Economic Review 84*(1), 47–65.

Judson, R. and A. Orphanides (1999). Inflation, Volatility and Growth. *International Finance 2*(1), 117–138.

Khan, M. S. and A. S. Senhadji (2001). Threshold Effects in the Relationship between Inflation and Growth. *IMF Staff Papers 48*(1), 1–21.

Kimbrough, K. P. (1986). The Optimum Quantity of Money Rule in the Theory of Public Finance. *Journal of Monetary Economics 18*(3), 277–284.

Krugman, P. (1988). Financing vs. Forgiving a Debt Overhang. *Journal of Development Economics 29*(3), 253–268.

Kumar, M. and J. Woo (2010). Public Debt and Growth. IMF Working Paper No. 10/174.

López-Villavicencio, A. and V. Mignon (2011). On the Impact of Inflation on Output Growth: Does the Level of Inflation Matter? *Journal of Macroeconomics 33*(3), 455 – 464.

Moon, H. R. and M. Weidner (2010). Dynamic Linear Panel Regression Models with Interactive Fixed Effects. Mimeo, July 2010.

Panizza, U. and A. F. Presbitero (2013). Public Debt and Economic Growth in Advanced Economies: A Survey. *Swiss Journal of Economics and Statistics 149*(II), 175–204.

Pesaran, M. H. (1997). The Role of Economic Theory in Modelling the Long Run. *Economic Journal 107*, 178–191.

Pesaran, M. H. (2004). General Diagnostic Tests for Cross Section Dependence in Panels. *IZA Discussion Paper No. 1240.*.

Pesaran, M. H. (2006). Estimation and Inference in Large Heterogeneous Panels with Multifactor Error Structure. *Econometrica 74*, 967–1012.

Pesaran, M. H. (2013). Testing Weak Cross-Sectional Dependence in Large Panels. *forthcoming in Econometric Reviews*.

Pesaran, M. H. and A. Chudik (2013). Aggregation in Large Dynamic Panels. *forthcoming in Journal of Econometrics*.

Pesaran, M. H. and Y. Shin (1999). An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis. In S. Strom (Ed.), *Econometrics and Economic Theory in 20th Century: The Ragnar Frisch Centennial Symposium*, Chapter 11, pp. 371–413. Cambridge: Cambridge University Press.

Pesaran, M. H. and R. Smith (1995). Estimating Long-run Relationships from Dynamic Heterogeneous Panels. *Journal of Econometrics 68*(1), 79–113.

Pesaran, M. H. and R. Smith (2013). Signs of Impact Effects in Time Series Regression Models. CAFE Research Paper No. 13.22.

Reinhart, C. M., V. R. Reinhart, and K. S. Rogoff (2012). Public Debt Overhangs: Advanced-Economy Episodes Since 1800. *The Journal of Economic Perspectives 26*(3), 69–86.

Reinhart, C. M. and K. S. Rogoff (2010). Growth in a Time of Debt. *American Economic Review 100*(2), 573–78.

Reinhart, C. M. and K. S. Rogoff (2011). From Financial Crash to Debt Crisis. *American Economic Review 101*(5), 1676–1706.

Roubini, N. and X. Sala-i-Martin (1992). Financial Repression and Economic Growth. *Journal of Development Economics 39*(1), 5 – 30.

Rousseau, P. L. and P. Wachtel (2001). Inflation, Financial Development and Growth. In *Economic Theory, Dynamics and Markets: Essays in Honor of Ryuzo Sato*. Springer.

Sarafidis, V. and T. Wansbeek (2012). Cross-Sectional Dependence in Panel Data Analysis. *Econometric Reviews 31*, 483–531.

Saving, T. R. (1971). Transactions Costs and the Demand for Money. *The American Economic Review 61*(3), 407–420.

Sidrauski, M. (1967). Rational Choice and Patterns of Growth in a Monetary Economy. *The American Economic Review 57*(2), 534–544.

Song, M. (2013). Asymptotic Theory for Dynamic Heterogeneous Panels with Cross-Sectional Dependence and Its Applications. Mimeo, January 2013.

Stockman, A. C. (1981). Anticipated Inflation and the Capital Stock in a Cash in-Advance Economy. *Journal of Monetary Economics 8*(3), 387–393.

Tobin, J. (1965). Money and Economic Growth. *Econometrica 33*(4), 671–684.

A Mathematical Appendix

We start by briefly summarizing the notations used in the paper, and introduce new notations which will prove useful in the proofs provided below. We use $\langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{a}'\mathbf{b}$ to denote the inner product (corresponding to the Euclidean norm) of vectors $\mathbf{a}$ and $\mathbf{b}$. $\|\mathbf{A}\|_1 \equiv \max_{1\leq j\leq n} \sum_{i=1}^n |a_{ij}|$, and $\|\mathbf{A}\|_\infty \equiv \max_{1\leq i\leq n} \sum_{j=1}^n |a_{ij}|$ denote the maximum absolute column and row sum norms of $\mathbf{A} \in \mathbb{M}^{n\times n}$, respectively, where $\mathbb{M}^{n\times n}$ is the space of real-valued $n\times n$ matrices. $\|\mathbf{A}\| = \sqrt{\varrho(\mathbf{A}'\mathbf{A})}$ is the spectral norm of $\mathbf{A}$, $\varrho(\mathbf{A})$ is the spectral radius of $\mathbf{A}$, $Col(\mathbf{A})$ denotes the space spanned by the column vectors of $\mathbf{A}$, and $\mathbf{A}^+$ is the Moore-Penrose pseudoinverse of $\mathbf{A}$. Note that $\|\mathbf{a}\| = \sqrt{\varrho(\mathbf{a}'\mathbf{a})} = \sqrt{\mathbf{a}'\mathbf{a}}$ corresponds to the Euclidean length of vector $\mathbf{a}$.

Let $\mathbf{z}_{it} = (y_{it}, \mathbf{x}_{it}')'$, $\bar{\mathbf{z}}_{wt} = (\bar{y}_{wt}, \bar{\mathbf{x}}_{wt}') = \sum_{i=1}^N w_i \mathbf{z}_{it}$, $\Delta = (1-L)$, L is the lag operator,

$$\underset{T-p\times 1}{\mathbf{y}_i} = \begin{pmatrix} y_{i,p+1} \\ y_{i,p+2} \\ \vdots \\ y_{iT} \end{pmatrix}, \underset{T-p\times k}{\mathbf{X}_i} = \begin{pmatrix} \mathbf{x}_{i,p+1}' \\ \mathbf{x}_{i,p+2}' \\ \vdots \\ \mathbf{x}_{iT}' \end{pmatrix}, \underset{(T-p)\times pk}{\Delta\mathbf{X}_{ip}} = \begin{pmatrix} \Delta\mathbf{x}_{i,p+1}' & \Delta\mathbf{x}_{i,p}' & \cdots & \Delta\mathbf{x}_{i2}' \\ \Delta\mathbf{x}_{i,p+2}' & \Delta\mathbf{x}_{i,p+1}' & \cdots & \Delta\mathbf{x}_{i3}' \\ \vdots & \vdots & & \vdots \\ \Delta\mathbf{x}_{iT}' & \Delta\mathbf{x}_{i,T-1}' & \cdots & \Delta\mathbf{x}_{i,T-p+1}' \end{pmatrix},$$

$$\underset{(T-p)\times k+1}{\bar{\mathbf{Z}}_w} = \begin{pmatrix} \bar{\mathbf{z}}_{i,p+1}' \\ \bar{\mathbf{z}}_{i,p+2}' \\ \vdots \\ \bar{\mathbf{z}}_{iT}' \end{pmatrix}, \underset{(T-p)\times pk}{\Delta\bar{\mathbf{X}}_{wp}} = \begin{pmatrix} \Delta\bar{\mathbf{x}}_{w,p+1}' & \Delta\bar{\mathbf{x}}_{w,p}' & \cdots & \Delta\bar{\mathbf{x}}_{w2}' \\ \Delta\bar{\mathbf{x}}_{w,p+2}' & \Delta\bar{\mathbf{x}}_{w,p+1}' & \cdots & \Delta\bar{\mathbf{x}}_{w3}' \\ \vdots & \vdots & & \vdots \\ \Delta\bar{\mathbf{x}}_{wT}' & \Delta\bar{\mathbf{x}}_{w,T-1}' & \cdots & \Delta\bar{\mathbf{x}}_{w,T-p+1}' \end{pmatrix}, \underset{T-p\times k}{\mathbf{V}_i} = \begin{pmatrix} \mathbf{v}_{i,p+1}' \\ \mathbf{v}_{i,p+2}' \\ \vdots \\ \mathbf{v}_{iT}' \end{pmatrix}$$

$\mathbf{Q}_{wi} = (\mathbf{Q}_w, \Delta\mathbf{X}_{ip})$, $\mathbf{Q}_w = \left(\bar{\mathbf{Z}}_w, \Delta\bar{\mathbf{X}}_{wp}\right)$,

$$\mathbf{M}_{qi} = \mathbf{I}_{T-p} - \mathbf{Q}_{wi}\left(\mathbf{Q}_{wi}'\mathbf{Q}_{wi}\right)^+ \mathbf{Q}_{wi}', \tag{A.1}$$

$\boldsymbol{\gamma}_{ip} = (\boldsymbol{\gamma}_i', \varphi_i\boldsymbol{\gamma}_i', ..., \varphi_i^p\boldsymbol{\gamma}_i')'$,

$$\underset{T-p\times mp}{\mathbf{F}_p} = \left(\mathbf{F}_{(0)}, \mathbf{F}_{(1)}, ..., \mathbf{F}_{(p)}\right), \underset{T-p\times m}{\mathbf{F}_{(\ell)}} \begin{pmatrix} \mathbf{f}_{p+1-\ell}' \\ \mathbf{f}_{p+2-\ell}' \\ \vdots \\ \mathbf{f}_{T-\ell}' \end{pmatrix}, \text{ for } \ell = 0, 1, 2, ..., p, \text{ and } \boldsymbol{\varepsilon}_i = \begin{pmatrix} \varepsilon_{i,p+1} \\ \varepsilon_{i,p+2} \\ \vdots \\ \varepsilon_{iT} \end{pmatrix}. \tag{A.2}$$

Using the above notations, model for the dependent variable can be written as

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\theta}_i + \Delta\mathbf{X}_{ip}\boldsymbol{\alpha}_{ip} + \mathbf{F}_p\boldsymbol{\gamma}_{ip} + \boldsymbol{\vartheta}_i + \boldsymbol{\varepsilon}_i,$$

for $i = 1, 2, ..., N$, where $\boldsymbol{\alpha}_{ip}$ is $pk \times 1$ vector containing the first p coefficients vectors of the polynomial $\boldsymbol{\alpha}_i(L)$ stacked into one single column vector, $\boldsymbol{\vartheta}_i = (\vartheta_{i,p+1}, \vartheta_{i,p+1}, ..., \vartheta_{i,T})'$, and

$$\vartheta_{it} = \sum_{\ell=p+1}^{\infty} \varphi_i^{\ell+1}\left(\boldsymbol{\beta}_i'\Delta\mathbf{x}_{i,t-\ell+1} + \gamma_i f_{t-\ell}\right),$$

for $i=1,2,...,N$ and $t=p+1,p+2,...,T$. The model for regressors can be written as

$$\mathbf{X}_i=\mathbf{F}_{(0)}\mathbf{\Gamma}_i+\mathbf{V}_i,$$

for $i=1,2,...,N$.

Define also the following projection matrix

$$\underset{T-p\times T-p}{\mathbf{M}_{hi}}=\mathbf{I}_{T-p}-\mathbf{H}_{wi}\left(\mathbf{H}_{wi}'\mathbf{H}_{wi}\right)^{+}\mathbf{H}_{wi}', \tag{A.3}$$

in which

$$\underset{T-p\times k(p+2)+1}{\mathbf{H}_{wi}}=\left(\mathbf{H}_w,\Delta\mathbf{X}_{ip}\right),\ \underset{T-p\times k(p+1)+1}{\mathbf{H}_w}=\begin{pmatrix}\mathbf{h}_{wp,p+1}'\\ \mathbf{h}_{w,p+2}'\\ \vdots\\ \mathbf{h}_{wp,T}'\end{pmatrix},$$

and

$$\underset{k(p+1)+1\times 1}{\mathbf{h}_{wpt}}=\begin{pmatrix}\bar{\boldsymbol{\theta}}_w'\bar{\mathbf{\Gamma}}_w'-\boldsymbol{\alpha}_w'\left(L\right)\bar{\mathbf{\Gamma}}_w'+\boldsymbol{\gamma}_w'\left(L\right)\\ \bar{\mathbf{\Gamma}}_w'\\ \left(1-L\right)\bar{\mathbf{\Gamma}}_w'\\ L\left(1-L\right)\bar{\mathbf{\Gamma}}_w'\\ \vdots\\ L^{p-1}\left(1-L\right)\bar{\mathbf{\Gamma}}_w'\end{pmatrix}\mathbf{f}_t,$$

where

$$\bar{\boldsymbol{\theta}}_w=\sum_{i=1}^N w_i\boldsymbol{\theta}_i,\ \bar{\mathbf{\Gamma}}_w=\sum_{i=1}^N w_i\mathbf{\Gamma}_i,\ \boldsymbol{\alpha}_w\left(L\right)=\sum_{i=1}^N w_i\boldsymbol{\alpha}_i\left(L\right),\ \boldsymbol{\gamma}_w\left(L\right)=\sum_{i=1}^N w_i\boldsymbol{\gamma}_i\left(L\right),$$

and $\boldsymbol{\gamma}_i\left(L\right)=\sum_{\ell=0}^{\infty}\varphi_i^p\boldsymbol{\gamma}_iL^p$.

A.1 Proofs of Theorems

Proof of Theorem 1. We have

$$\sqrt{N}\left(\widehat{\boldsymbol{\theta}}_{MG}-\boldsymbol{\theta}\right)=\frac{1}{\sqrt{N}}\sum_{i=1}^N\boldsymbol{\upsilon}_{\theta i}+\frac{1}{\sqrt{N}}\sum_{i=1}^N\widehat{\mathbf{\Psi}}_{iT}^{-1}\frac{\mathbf{X}_i'\mathbf{M}_{qi}\mathbf{F}_p\boldsymbol{\gamma}_{ip}}{T}+\frac{1}{\sqrt{N}}\sum_{i=1}^N\widehat{\mathbf{\Psi}}_{iT}^{-1}\frac{\mathbf{X}_i'\mathbf{M}_{qi}\boldsymbol{\vartheta}_i}{T}\frac{1}{\sqrt{N}}\sum_{i=1}^N\widehat{\mathbf{\Psi}}_{iT}^{-1}\frac{\mathbf{X}_i'\mathbf{M}_{qi}\boldsymbol{\varepsilon}_i}{T} \tag{A.4}$$

where $\widehat{\mathbf{\Psi}}_{iT}=T^{-1}\mathbf{X}_i'\mathbf{M}_{qi}\mathbf{X}_i$,

$$\underset{T-p\times m(p+1)}{\mathbf{F}_p}=\begin{pmatrix}\mathbf{f}_{p+1}' & \mathbf{f}_p' & \cdots & \mathbf{f}_1'\\ \vdots & \vdots & & \vdots\\ \mathbf{f}_T' & \mathbf{f}_{T-1}' & \cdots & \mathbf{f}_{T-p}'\end{pmatrix},$$

$\boldsymbol{\gamma}_{ip} = (\boldsymbol{\gamma}_i', \varphi_i \boldsymbol{\gamma}_i', ..., \varphi_i^p \boldsymbol{\gamma}_i')'$, $\boldsymbol{\vartheta}_i = (\vartheta_{i,p+1}, \vartheta_{i,p+1}, ..., \vartheta_{i,T})'$, and

$$\vartheta_{it} = \sum_{\ell=p+1}^{\infty} \varphi_i^{\ell+1} \left(\boldsymbol{\beta}_i' \Delta \mathbf{x}_{i,t-\ell+1} + \gamma_i f_{t-\ell} \right).$$

Consider the asymptotics $(N, T, p) \xrightarrow{j} \infty$ such that $\sqrt{N} p \rho^p \to 0$, for any constant $0 < \rho < 1$ and $p^3/T \to \varkappa$, $0 < \varkappa < \infty$. In what follows we establish convergence of the individual terms on the right side of (A.4).

It follows from (A.26) of Lemma A.1 and (A.27) of Lemma A.2 that

$$\widehat{\boldsymbol{\Psi}}_{\Xi,iT} - \boldsymbol{\Sigma}_i = o_p \left(N^{-1/2} \right) \text{ uniformly in } i. \tag{A.5}$$

(A.5), (A.28) of Lemma A.2, and (A.30) of Lemma A.3 imply

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \widehat{\boldsymbol{\Psi}}_{iT}^{-1} \frac{\mathbf{X}_i' \mathbf{M}_{qi} \boldsymbol{\varepsilon}_i}{T} \xrightarrow{p} \underset{k \times 1}{\mathbf{0}}. \tag{A.6}$$

Consider now the second term on the right side of (A.4), which involves common factors and their loadings. In the previous literature on CCE estimators, Pesaran (2006) established the asymptotic results for the term involving factors and their loadings in the expression for his CCEMG estimator by focusing on the properties of the matrix (using Pesaran (2006)'s notations) $\mathbf{X}_i' \bar{\mathbf{M}}_w \mathbf{F}/T$, see equation (40) in Pesaran (2006), in the full rank case, and by exploring the relation (still using Pesaran (2006)'s notations) $\mathbf{M}_q \mathbf{F} \bar{\mathbf{C}}_w = \mathbf{0}$, see p. 979 of Pesaran (2006), in the rank deficient case. But unlike in the set-up of Pesaran (2006), the dimension of $\mathbf{X}_i' \mathbf{M}_{qi} \mathbf{F}_p / T$ in this paper increases with the sample size, and furthermore $\mathbf{M}_{hi} \mathbf{F}_p \overline{\boldsymbol{\gamma}}_{wp}$ is not necessarily zero since $\mathbf{F}_p \overline{\boldsymbol{\gamma}}_{wp}$ (due to the truncation lag p) does not necessarily belong to the linear space spanned by the column vectors of $\mathbf{H}_{wi}$. We therefore focus on the elements of the vector $\mathbf{X}_i' \mathbf{M}_{qi} \mathbf{F}_p \boldsymbol{\gamma}_{ip} / T$ below, which has fixed (finite) dimensions, and we also take advantage of the exponential decay of certain coefficients below. Using (A.5), boundedness of $\boldsymbol{\Sigma}_i^{-1}$ (by Assumption 5), and result (A.29) of Lemma A.2 we obtain

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\frac{\mathbf{X}_i' \mathbf{M}_{qi} \mathbf{X}_i}{T} \right)^{-1} \frac{\mathbf{X}_i' \mathbf{M}_{qi} \mathbf{F}_p}{T} \boldsymbol{\gamma}_{ip} - \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\frac{\mathbf{X}_i' \mathbf{M}_{hi} \mathbf{X}_i}{T} \right)^{-1} \frac{\mathbf{X}_i' \mathbf{M}_{hi} \mathbf{F}_p}{T} \boldsymbol{\gamma}_{ip} \xrightarrow{p} \underset{k \times 1}{\mathbf{0}}.$$

Vector $\boldsymbol{\gamma}_{ip}$ can be written as $\boldsymbol{\gamma}_{ip} = \left(\overline{\boldsymbol{\gamma}}_{wp} - \overline{\boldsymbol{\eta}}_{\gamma wp} \right) + \boldsymbol{\eta}_{\gamma ip}$, and

$$T^{-1} \mathbf{X}_i' \mathbf{M}_{hi} \mathbf{F}_p \boldsymbol{\gamma}_{ip} = T^{-1} \mathbf{X}_i' \mathbf{M}_{hi} \mathbf{F}_p \overline{\boldsymbol{\gamma}}_{wp} + T^{-1} \mathbf{X}_i' \mathbf{M}_{hi} \mathbf{F}_p \left(\boldsymbol{\eta}_{\gamma ip} - \overline{\boldsymbol{\eta}}_{\gamma wp} \right).$$

Note again that $\mathbf{F}_p \overline{\boldsymbol{\gamma}}_{wp}$ does not necessarily belong to the linear space spanned by the column vectors of $\mathbf{H}_{wi}$ due to the truncation lag p. But Assumption 4 constraints the support of φ_i to fall strictly within the unit circle, which implies that there exists a positive constant $\rho < 1$ such that $|\varphi_i| < \rho < 1$ for all possible realizations of the random variable φ_i. Therefore, under Assumptions 3-4, the coefficients in the polynomials $\boldsymbol{\alpha}_w (L) = \sum_{i=1}^{N} w_i \boldsymbol{\alpha}_i (L)$ and $\boldsymbol{\gamma}_w (L) = \sum_{i=1}^{N} w_i \boldsymbol{\gamma}_i (L)$, where

$\boldsymbol{\alpha}_i(L) = \sum_{\ell=0}^{\infty} \varphi_i^{\ell+1}(1-\varphi_i)^{-1}\boldsymbol{\beta}_i L^\ell$ and $\boldsymbol{\gamma}_i(L) = \sum_{\ell=0}^{\infty} \varphi_i^p \boldsymbol{\gamma}_i L^p$, decay exponentially to zero[15] and we have

$$\bar{\boldsymbol{\gamma}}_w'(L,p)\mathbf{f}_t - E\left[\bar{\boldsymbol{\gamma}}_w'(L,p)\mathbf{f}_t \left|\mathbf{h}_{wpt}\right.\right] = O_p(\rho^p), \tag{A.7}$$

uniformly in t, where $\bar{\boldsymbol{\gamma}}_w(L,p) = \sum_{\ell=0}^{p}\sum_{i=1}^{N} w_i\varphi_i^\ell \boldsymbol{\gamma}_i L^\ell$ is the truncated polynomial of $\bar{\boldsymbol{\gamma}}_w(L)$ featuring only orders up to L^p. Using the properties of orthogonal projectors, we obtain[16]

$$\left\|\mathbf{M}_{hi}\mathbf{F}_p\overline{\boldsymbol{\gamma}}_{wp}\right\| \leq \left\|\mathbf{F}_p\overline{\boldsymbol{\gamma}}_{wp} - \mathbf{H}_{wi}\mathbf{c}\right\|, \tag{A.8}$$

for any $k(p+1)+1\times 1$ vector $\mathbf{c}$. Let $\mathbf{c}$ be defined by $E\left[\bar{\boldsymbol{\gamma}}_w'(L,p)\mathbf{f}_t\left|\mathbf{h}_{wpt}\right.\right] = \mathbf{c}'\mathbf{h}_{wpt}$. Then it follows from (A.7) that the individual elements of $T-p\times 1$ vector $\left(\mathbf{F}_p\overline{\boldsymbol{\gamma}}_{wp} - \mathbf{H}_{wi}\mathbf{c}\right)$ are uniformly $O_p(\rho^p)$ and using (A.8) we have

$$\left\|\mathbf{M}_{hi}\mathbf{F}_p\overline{\boldsymbol{\gamma}}_{wp}\right\| = O_p\left[(T-p)^{1/2}\rho^p\right].$$

Using now Cauchy-Schwarz inequality, we obtain[17]

$$T^{-1}\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{F}_p\overline{\boldsymbol{\gamma}}_{wp} = O_p(\rho^p). \tag{A.9}$$

Noting that $\sqrt{N}\rho^p \to 0$, and using (A.5) and boundedness of $\boldsymbol{\Sigma}_i^{-1}$ (by Assumption 5) we have

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\frac{\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{X}_i}{T}\right)^{-1}\frac{\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{F}_p}{T}\overline{\boldsymbol{\gamma}}_{wp} \xrightarrow{p} \mathbf{0},$$

and it now follows that

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\widehat{\boldsymbol{\Psi}}_{\Xi,iT}^{-1}\frac{\mathbf{X}_i'\mathbf{M}_{qi}\mathbf{F}_p}{T}\boldsymbol{\gamma}_{ip} - \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\frac{\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{X}_i}{T}\right)^{-1}\frac{\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{F}_p}{T}\left(\boldsymbol{\eta}_{\gamma ip} - \overline{\boldsymbol{\eta}}_{\gamma wp}\right) \xrightarrow{p} \underset{k\times 1}{\mathbf{0}}. \tag{A.10}$$

Now consider the term $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\frac{\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{X}_i}{T}\right)^{-1}\frac{\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{F}_p}{T}\overline{\boldsymbol{\eta}}_{\gamma wp}$. Let us denote individual columns of $\mathbf{F}_p$ as $\mathbf{f}_{p,[j]}$, for $j = 1,2,...,mp$, and individual elements of $\overline{\boldsymbol{\eta}}_{\gamma wp}$ and $\overline{\boldsymbol{\gamma}}_{wp}$ as $\overline{\eta}_{\gamma wp,j}$ and $\overline{\gamma}_{wp,j}$, respectively, for $j = 1,2,...,mp$. $\mathbf{F}_p\overline{\boldsymbol{\eta}}_{\gamma wp}$ thus can be written as $\sum_{j=1}^{mp}\mathbf{f}_{p,[j]}\overline{\eta}_{\gamma wp,j}$. Let

$$\pi_j = \frac{\overline{\eta}_{\gamma wp,j}}{\gamma_{p,j} + \overline{\eta}_{\gamma wp,j}},$$

where $\gamma_{p,j}$ is the j-th element of the vector $E\left(\boldsymbol{\gamma}_{ip}\right)$. Note that $p\lim_{N\to\infty}\pi_j = 1$ if $\gamma_{p,j} = 0$ and $p\lim_{N\to\infty}\pi_j = 0$ if $\gamma_{p,j} \neq 0$. Expression $\mathbf{F}_p\overline{\boldsymbol{\eta}}_{ywp}$ can now be written as $\mathbf{F}_p\overline{\boldsymbol{\eta}}_{ywp} = \sum_{j=1}^{mp}\mathbf{f}_{p,[j]}\overline{\gamma}_{wp,j}\pi_j$

[15] See Pesaran and Chudik (2013) for a related discussion.

[16] We use the following property. Let $\mathbf{A}$ be $s_1 \times s_2$ dimensional matrix, $s_1 > s_2$, and let $\mathbf{M}_A = \mathbf{I}_{s_1} - \mathbf{A}(\mathbf{A}'\mathbf{A})^{+}\mathbf{A}'$ be the corresponding orthogonal projector that projects on orthogonal complement of the space spanned by the column vectors of $\mathbf{A}$. Then for any $s_1 \times 1$ dimensional vector $\mathbf{x}$ and any $s_2 \times 1$ dimensional vector $\mathbf{c}$, $\|\mathbf{M}_A\mathbf{x}\| \leq \|\mathbf{x} - \mathbf{A}\mathbf{c}\|$.

[17] $\langle\mathbf{a},\mathbf{b}\rangle \leq \|\mathbf{a}\|\|\mathbf{b}\|$. We set $\mathbf{a} = T^{-1}\mathbf{X}_i$, and $\mathbf{b} = \mathbf{M}_h\mathbf{F}_p\overline{\boldsymbol{\gamma}}_{pw}$, where $\|\mathbf{a}\| = O_p\left[(T-p)^{-1/2}\right]$, and $\|\mathbf{b}\| = O_p\left[(T-p)^{1/2}\rho^p\right]$.

and

$$\frac{\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{F}_p}{T}\overline{\boldsymbol{\eta}}_{ywp} = \sum_{j=1}^{mp} \frac{\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{f}_{p,[j]}}{T}\overline{\gamma}_{wp,j}\pi_j.$$

Using the same arguments as in the derivation of (A.9), we obtain $\frac{\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{f}_{p,[j]}}{T}\overline{\gamma}_{wp,j} = O_p\left(\rho^p\right)$ and using the properties of π_j we have

$$\sum_{j=1}^{mp} \frac{\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{f}_{p,[j]}}{T}\overline{\gamma}_{wp,j}\pi_j = O_p\left(p\rho^p\right).$$

But $\sqrt{N}p\rho^p \to 0$ and therefore

$$\sqrt{N}\frac{\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{F}_p}{T}\overline{\boldsymbol{\eta}}_{ywp} \xrightarrow{p} \underset{k\times 1}{\mathbf{0}}. \tag{A.11}$$

Using this result in (A.10) together with (A.5) and the boundedness of $\left\|\boldsymbol{\Sigma}_i^{-1}\right\|$ we obtain

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\widehat{\boldsymbol{\Psi}}_{\Xi,iT}^{-1}\frac{\mathbf{X}_i'\mathbf{M}_{qi}\mathbf{F}_p}{T}\boldsymbol{\gamma}_i - \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\frac{\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{X}_i}{T}\right)^{-1}\frac{\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{F}_p}{T}\boldsymbol{\eta}_{\gamma ip} \xrightarrow{p} \underset{k\times 1}{\mathbf{0}}. \tag{A.12}$$

Consider now the third term on the right side of (A.4). Let $\tilde{\mathbf{x}}_{it}$ denote the column $(t-p)$ of the matrix $\mathbf{X}_i'\mathbf{M}_{qi}$, for $t = p+1, p+2, ..., T$. We have $\tilde{\mathbf{x}}_{it} = O_p(1)$ uniformly in i, $\widehat{\boldsymbol{\Psi}}_{iT}^{-1} = O_p(1)$ uniformly in i, and

$$E\left|\sqrt{N}\vartheta_{it}\right| \leq \sqrt{N}\sum_{\ell=p+1}^{\infty}\left|\varphi_i\right|^{\ell+1}E\left|\boldsymbol{\beta}_i'\Delta\mathbf{x}_{i,t-\ell+1} + \gamma_i f_{t-\ell}\right| < K\sqrt{N}\rho^p, \tag{A.13}$$

uniformly in i and t. It follows that $E\left|\sqrt{N}\vartheta_{it}\right| \xrightarrow{p} 0$ as $\sqrt{N}\rho^p \to 0$,

$$\frac{1}{T}\sum_{t=1}^{T}\tilde{\mathbf{x}}_{it}\vartheta_{it} \xrightarrow{p} 0 \text{ uniformly in } i, \tag{A.14}$$

and

$$\frac{1}{N}\sum_{i=1}^{N}\widehat{\boldsymbol{\Psi}}_{iT}^{-1}\left(\frac{\mathbf{X}_i'\overline{\mathbf{M}}_{wi}\left(\sqrt{N}\boldsymbol{\vartheta}_i\right)}{T}\right) \xrightarrow{p} \underset{k\times 1}{\mathbf{0}}. \tag{A.15}$$

Using (A.6), (A.12) and (A.15) in (A.4), we obtain

$$\sqrt{N}\left(\widehat{\boldsymbol{\theta}}_{MG} - \boldsymbol{\theta}\right) \overset{d}{\sim} \boldsymbol{\vartheta}_{\theta i}, ,$$

where

$$\boldsymbol{\vartheta}_{\theta i} = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\upsilon}_i + \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\frac{\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{X}_i}{T}\right)^{-1}\frac{\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{F}_p}{T}\boldsymbol{\eta}_{\gamma ip}, \tag{A.16}$$

and recall that $\boldsymbol{\upsilon}_i$ and $\boldsymbol{\eta}_{\gamma ip}$ are independently distributed across i. It now follows that when $\boldsymbol{\eta}_{\gamma i}$ is independently distributed from $\boldsymbol{\Gamma}_i$ and regardless whether the rank condition holds, $\sqrt{N}\left(\widehat{\boldsymbol{\theta}}_{MG} - \boldsymbol{\theta}\right) \xrightarrow{d}$

$N\left(\underset{k\times 1}{\mathbf{0}}, \mathbf{\Sigma}_{MG}\right)$, where

$$\mathbf{\Sigma}_{MG} = \mathbf{\Omega}_{\theta} + \lim_{p,N\to\infty}\left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{\Sigma}_{i\xi}^{-1}\mathbf{Q}_{if}\mathbf{\Omega}_{\gamma}\mathbf{Q}_{if}'\mathbf{\Sigma}_{i\xi}^{-1}\right], \tag{A.17}$$

in which $\mathbf{\Omega}_{\theta} = Var\left(\boldsymbol{\theta}_i\right)$, $\mathbf{\Omega}_{\gamma} = Var\left(\boldsymbol{\gamma}_i\right)$, and $\mathbf{\Sigma}_i = p\lim T^{-1}\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{X}_i$ and $\mathbf{Q}_{if} = p\lim T^{-1}\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{F}$. When the rank condition stated in assumptions of Theorem 1 holds then $\mathbf{Q}_{if} = \underset{k\times m}{\mathbf{0}}$, and therefore even if $\boldsymbol{\eta}_{\gamma i}$ is correlated with $\mathbf{\Gamma}_i$, $\sqrt{N}\left(\widehat{\boldsymbol{\theta}}_{MG} - \boldsymbol{\theta}\right) \overset{d}{\sim} \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\upsilon}_i$. Consistency of the nonparametric estimator can be established in the same way as in Chudik and Pesaran (2013a). ■

Proof of Theorem 2. Consider

$$\left(\sum_{i=1}^{N}w_i^2\right)^{-1/2}\left(\widehat{\boldsymbol{\theta}}_P - \boldsymbol{\theta}\right) = \left(\sum_{i=1}^{N}w_i\frac{\mathbf{X}_i'\mathbf{M}_{qi}\mathbf{X}_i}{T}\right)^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\widetilde{w}_i\frac{\mathbf{X}_i'\mathbf{M}_{qi}\left(\mathbf{X}_i\mathbf{v}_i + \mathbf{F}_p\boldsymbol{\gamma}_{ip} + \boldsymbol{\vartheta}_i + \boldsymbol{\varepsilon}_i\right)}{T}, \tag{A.18}$$

where $\boldsymbol{\vartheta}_i$ is defined below (A.4), $\widetilde{w}_i = \sqrt{N}w_i\left(\sum_{i=1}^{N}w_i^2\right)^{-1/2}$, and, by granularity conditions (19)-(20) there exists a constant $K<\infty$ (independent of i and N), such that

$$\left|\widetilde{w}_i\right| = \left|\sqrt{N}w_i\left(\sum_{i=1}^{N}w_i^2\right)^{-1/2}\right| < K. \tag{A.19}$$

We focus on the individual terms on the right side of (A.18) below and assume that $(N,T,p)\overset{j}{\to}\infty$ such that $\sqrt{N}p\rho^p\to 0$ for any constant $0<\rho<1$ and $p^3/T\to\varkappa$, $0<\varkappa<\infty$.

Using results (A.26) of Lemma A.1 we have

$$\sum_{i=1}^{N}w_i\frac{\mathbf{X}_i'\mathbf{M}_{qi}\mathbf{X}_i}{T} - \sum_{i=1}^{N}w_i\mathbf{\Sigma}_{iq}\overset{p}{\to}\underset{k\times 1}{\mathbf{0}},$$

for any weights $\{w_i\}$ satisfying granularity conditions (19)-(20). The limit $\lim_{N\to\infty}\sum_{i=1}^{N}w_i\mathbf{\Sigma}_{iq} = \mathbf{\Psi}^*$ exists by Assumption 5 and furthermore, by the same assumption, $\mathbf{\Psi}^*$ is nonsingular. It therefore follows that

$$\left(\sum_{i=1}^{N}w_i\frac{\mathbf{X}_i'\mathbf{M}_{qi}\mathbf{X}_i}{T}\right)^{-1}\overset{p}{\to}\mathbf{\Psi}^{*-1}. \tag{A.20}$$

Noting that $\boldsymbol{\gamma}_{ip}$ can be written as $\boldsymbol{\gamma}_{ip} = \overline{\boldsymbol{\gamma}}_{wp} + \boldsymbol{\eta}_{ip} - \overline{\boldsymbol{\eta}}_{wp}$, and using (A.9), (A.11), (A.19) and $\sqrt{N}\rho^p\to 0$ we obtain[18]

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\widetilde{w}_i\frac{\mathbf{X}_i'\mathbf{M}_{qi}\mathbf{F}_p}{T}\boldsymbol{\gamma}_{ip} - \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\widetilde{w}_i\frac{\mathbf{X}_i'\mathbf{M}_{qi}\mathbf{F}_p}{T}\boldsymbol{\eta}_{ip}\overset{p}{\to}\underset{k\times 1}{\mathbf{0}}. \tag{A.21}$$

[18](A.21) can also be established by noting that the column vectors of $\mathbf{X}_w = \sum_{i=1}^{N}w_i\mathbf{X}_i$ are included in $\mathbf{Q}_{wi}$ and therefore $\mathbf{X}_w'\mathbf{M}_{qi} = \mathbf{0}$.

(A.14) and (A.19) imply

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\widetilde{w}_i\frac{\mathbf{X}_i'\mathbf{M}_{qi}\boldsymbol{\vartheta}_i}{T}\xrightarrow{p}\underset{k\times 1}{\mathbf{0}}. \tag{A.22}$$

Result (A.28) of Lemma A.2 and result (A.30) of Lemma A.3 establish

$$\sqrt{N}\frac{\mathbf{X}_i'\mathbf{M}_{qi}\boldsymbol{\varepsilon}_i}{T}\xrightarrow{p}\underset{k\times 1}{\mathbf{0}} \text{ uniformly in } i,$$

and therefore (noting that $\widetilde{w}_i$ is uniformly bounded in i, see (A.19)),

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\widetilde{w}_i\frac{\mathbf{X}_i'\mathbf{M}_{qi}\boldsymbol{\varepsilon}_i}{T}=\frac{1}{N}\sum_{i=1}^{N}\widetilde{w}_i\left(\sqrt{N}\frac{\mathbf{X}_i'\mathbf{M}_{qi}\boldsymbol{\varepsilon}_i}{T}\right)\xrightarrow{p}\underset{k\times 1}{\mathbf{0}}. \tag{A.23}$$

Using (A.20), (A.21), (A.22), (A.23) and result (A.27) of Lemma A.2 in (A.18), we obtain

$$\left(\sum_{i=1}^{N}w_i^2\right)^{-1/2}\left(\widehat{\boldsymbol{\theta}}_P-\boldsymbol{\theta}\right)\overset{d}{\sim}\boldsymbol{\Psi}^{*-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\widetilde{w}_i\frac{\mathbf{X}_i'\mathbf{M}_{hi}\left(\mathbf{X}_i\mathbf{v}_i+\mathbf{F}_p\boldsymbol{\eta}_{ip}\right)}{T}.$$

Assumption 5 is sufficient for the bounded second moments of $\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{X}_i/T$ and $\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{F}_p/T$. In particular, condition $E\left(\widetilde{x}_{ist}^4\right)<K$, for $s=1,2,..,k$, is sufficient for the bounded second moment of $\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{X}_i/T$. To see this note that

$$\frac{\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{X}_i}{T}=\frac{1}{T}\sum_{t=1}^{T}\widetilde{\mathbf{x}}_{it}\widetilde{\mathbf{x}}_{it}',$$

and, by Minkowski's inequality,

$$\left\|\frac{1}{T}\sum_{t=1}^{T}\widetilde{x}_{ist}\widetilde{x}_{ipt}'\right\|_{L_2}\leq\frac{1}{T}\sum_{t=1}^{T}\left\|\widetilde{x}_{ist}\widetilde{x}_{ipt}'\right\|_{L_2},$$

for any $s,p=1,2,..,k$. But by Cauchy-Schwarz inequality, we have $E\left(\widetilde{x}_{ist}^2\widetilde{x}_{ipt}^2\right)\leq\left[E\left(\widetilde{x}_{ist}^4\right)E\left(\widetilde{x}_{ipt}^4\right)\right]^{1/2}$, and therefore bounded fourth moments of the elements of $\widetilde{\mathbf{x}}_{it}$ are sufficient for the existence of an upper bound for the second moments of $\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{X}_i/T$. Similar arguments can be used to establish that $\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{F}_p/T$ has bounded second moments. Note also that $\mathbf{v}_i$ and $\boldsymbol{\eta}_{ip}$ are independently distributed across i; and, independently distributed of $\mathbf{M}_{hi}$, $\mathbf{F}_p$ and, assuming that $\boldsymbol{\gamma}_i$ is independently distributed of $\boldsymbol{\Gamma}_i$, also $\mathbf{X}_i$. It therefore follows, using similar arguments as in Lemma 4 of Pesaran (2006), that

$$\left(\sum_{i=1}^{N}w_i^2\right)^{-1/2}\left(\widehat{\boldsymbol{\theta}}_P-\boldsymbol{\theta}\right)\xrightarrow{d}N\left(\mathbf{0},\boldsymbol{\Sigma}_P\right),$$

where

$$\boldsymbol{\Sigma}_P=\boldsymbol{\Psi}^{*-1}\mathbf{R}^{*}\boldsymbol{\Psi}^{*-1}, \tag{A.24}$$

in which

$$\mathbf{\Psi}^* = \lim_{N\to\infty}\sum_{i=1}^N w_i \mathbf{\Sigma}_i,\ \mathbf{R}^* = \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^N \widetilde{w}_i^2\left(\mathbf{\Sigma}_i\mathbf{\Omega}_\theta\mathbf{\Sigma}_i + \mathbf{Q}_{if}\mathbf{\Omega}_\gamma\mathbf{Q}_{if}'\right),$$

$\mathbf{\Omega}_\theta = Var\left(\boldsymbol{\theta}_i\right)$, $\mathbf{\Omega}_\gamma = Var\left(\boldsymbol{\gamma}_i\right)$, $\mathbf{\Sigma}_i = p\lim T^{-1}\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{X}_i$ and $\mathbf{Q}_{if} = p\lim T^{-1}\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{F}$. $\mathbf{\Sigma}_P$ can be estimated as

$$\widehat{\mathbf{\Sigma}}_P = \left(\sum_{i=1}^N w_i^2\right)\hat{\mathbf{\Psi}}^{*-1}\hat{\mathbf{R}}^*\hat{\mathbf{\Psi}}^{*-1}, \tag{A.25}$$

where

$$\hat{\mathbf{\Psi}}^* = \sum_{i=1}^N w_i\left(\frac{\mathbf{X}_i'\mathbf{M}_{qi}\mathbf{X}_i}{T}\right),$$

and

$$\hat{\mathbf{R}}^* = \frac{1}{N-1}\sum_{i=1}^N \tilde{w}_i^2\left(\frac{\mathbf{X}_i'\mathbf{M}_{qi}\mathbf{X}_i}{T}\right)\left(\widehat{\boldsymbol{\theta}}_i - \widehat{\boldsymbol{\theta}}_{MG}\right)\left(\widehat{\boldsymbol{\theta}}_i - \widehat{\boldsymbol{\theta}}_{MG}\right)'\left(\frac{\mathbf{X}_i'\bar{\mathbf{M}}_{wi}\mathbf{X}_i}{T}\right).$$

When the rank condition holds, then column vectors of $\mathbf{F}_p$ belong to the space spanned by the column vectors of $\mathbf{H}_w$, and therefore regardless whether $\boldsymbol{\eta}_{\gamma i}$ is correlated with $\mathbf{\Gamma}_i$ or not, $\left(\sum_{i=1}^N w_i^2\right)^{-1/2}\left(\widehat{\boldsymbol{\theta}}_P - \boldsymbol{\theta}\right) \xrightarrow{d} N\left(\mathbf{0}, \mathbf{\Sigma}_P\right)$ in the full rank case with $\mathbf{\Sigma}_P$ reduced to $\mathbf{\Psi}^{*-1}\mathbf{R}_\theta^*\mathbf{\Psi}^{*-1}$ and $\mathbf{Q}_{if} = \underset{k\times m}{\mathbf{0}}$. Consistency of $\widehat{\mathbf{\Sigma}}_P$ can be established using similar arguments as in Pesaran (2006). ■

A.2 Lemmas

Lemma A.1 *Suppose Assumptions 1- 5 hold and* $(N,T,p) \xrightarrow{j} \infty$ *such that* $p^3/T \to \varkappa$, $0 < \varkappa < \infty$. *Then,*

$$\frac{\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{X}_i}{T} \xrightarrow{p} \mathbf{\Sigma}_i,\ \textit{uniformly in } i. \tag{A.26}$$

Proof. Let $\boldsymbol{\xi}_{hit}'$ denote the individual rows of $\mathbf{M}_{hi}\mathbf{X}_i$ so that

$$\frac{\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{X}_i}{T} = \frac{T-p}{T}\frac{1}{T-p}\sum_{t=p+1}^T \boldsymbol{\xi}_{hit}\boldsymbol{\xi}_{hit}'.$$

Ergodicity in mean of $\boldsymbol{\xi}_{hit}$ has been established in Chudik and Pesaran, (2013a, Lemma A3). This completes the proof of (A.26). ■

Lemma A.2 *Suppose Assumptions 1- 5 hold and* $(N,T,p) \xrightarrow{j} \infty$ *such that* $p^3/T \to \varkappa$, $0 < \varkappa < \infty$. *Then,*

$$\sqrt{N}\frac{\mathbf{X}_i'\mathbf{M}_{qi}\mathbf{X}_i}{T} - \sqrt{N}\frac{\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{X}_i}{T} \xrightarrow{p} \underset{k\times k}{\mathbf{0}},\ \textit{uniformly in } i. \tag{A.27}$$

$$\sqrt{N}\frac{\mathbf{X}_i'\mathbf{M}_{qi}\boldsymbol{\varepsilon}_i}{T} - \sqrt{N}\frac{\mathbf{X}_i'\mathbf{M}_{hi}\boldsymbol{\varepsilon}_i}{T} \xrightarrow{p} \underset{k\times 1}{\mathbf{0}},\ \textit{uniformly in } i. \tag{A.28}$$

$$\left\|\sqrt{N}\frac{\mathbf{X}_i'\mathbf{M}_{qi}\mathbf{F}_p}{T} - \sqrt{N}\frac{\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{F}_p}{T}\right\|_1 \xrightarrow{p} 0,\ \textit{uniformly in } i. \tag{A.29}$$

Proof. Results (A.27) and (A.28) can be established in the same way as Chudik and Pesaran, (2013a, results A.21 and A.22 of Lemma A6). Consider now (A.29). $\mathbf{F}_p$ can be written as $\mathbf{F}_p = \left[\mathbf{F}_{(0)}, \mathbf{F}_{(1)}, ..., \mathbf{F}_{(p)}\right]$, where $\mathbf{F}_{(\ell)} = (\mathbf{f}_{p+1-\ell}, \mathbf{f}_{p+2-\ell}, ..., \mathbf{f}_{T-\ell})'$ for $\ell = 0, 1, 2, ..., p$. Using the same arguments as in Chudik and Pesaran, (2013a, results A.23 of Lemma A6), it can be shown that

$$\sqrt{N}\frac{\mathbf{X}_i'\mathbf{M}_{qi}\mathbf{F}_{(\ell)}}{T} - \sqrt{N}\frac{\mathbf{X}_i'\mathbf{M}_{hi}\mathbf{F}_{(\ell)}}{T} \xrightarrow{p} \underset{k\times m}{\mathbf{0}},$$

uniformly in i and ℓ. This is sufficient for (A.29) to hold. ■

Lemma A.3 *Suppose Assumptions 1- 5 hold and* $(N, T, p) \xrightarrow{j} \infty$ *such that* $p^3/T \to \varkappa$, $0 < \varkappa < \infty$. *Then,*

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{\mathbf{X}_i'\mathbf{M}_{hi}\boldsymbol{\varepsilon}_i}{T} \xrightarrow{p} \underset{k\times 1}{\mathbf{0}}, \text{ uniformly in } i. \tag{A.30}$$

Proof. Results (A.27) can be established in the same way as Chudik and Pesaran, (2013a, results A.26). ■