

Centre for Energy Economics Research and Policy

Estimating long run effects in models with cross-sectional dependence using xtdcce2

Jan Ditzen

CEERP Working Paper No. 7
March 2019



Heriot-Watt University
Edinburgh, Scotland
EH14 4AS
ceerp.hw.ac.uk

Estimating long run effects in models with cross-sectional dependence using `xtdcce2`

Jan Ditzen*

Centre for Energy Economics Research and Policy (CEERP)
Heriot-Watt University, Edinburgh, UK

as of March 8, 2019

Abstract

This paper describes how to estimate long run effects in a large heterogeneous panel data model with cross sectional dependence in Stata using the user written command `xtdcce2`. It builds on Chudik et al. (2016) and explains how to estimate models using the CS-DL and CS-ARDL estimator. In addition it includes a method how to estimate an error correction model.

keywords `st0536`, `xtdcce2`, parameter heterogeneity, dynamic panels, cross section dependence, common correlated effects, pooled mean-group estimator, mean-group estimator, error correction model, `ardl`, long run coefficients.

DRAFT

1 Introduction

The estimation of long run relationships is important in the empirical application of macroeconomics models. Long run relationships describe the steady state solution and how a change in the steady state affect the long run relationship between variables. Examples are the the relationship between macroeconomic variables such as GDP and inflation or investments, exchange rates of different countries, effects of education or technological progress on economic growth.

In an econometric model, estimating a long run relationship implies to estimate coefficients which capture this relationship. A standard model in time series econometrics is the error correction model, in which one term of the equation

* *Correspondence:* Jan Ditzen, Centre for Energy Economics Research and Policy (CEERP) and Spatial Economics and Econometrics Centre (SEEC), Heriot-Watt University, Edinburgh EH14 4AS, Scotland, United Kingdom; Email: j.ditzen@hw.ac.uk, Web: www.jan.ditzen.net

captures the short and the other the long run movements (Engle and Granger, 1987). In a panel data model with heterogeneous slope coefficients, the model is estimated by the pooled mean group estimator (Shin et al., 1999), which was implemented into Stata using the user written command `xtpmg` (Blackburne and Frank, 2007). A related model is the ARDL model, implemented by the user written `ardl` command in Stata (Kripfganz and Schneider, 2018).

A common problem in the estimation of panels with a large number of observations across time and cross-sectional units is cross-sectional dependence. If not taken care of, it causes estimates to be inconsistent. Pesaran (2006) and Chudik and Pesaran (2015) developed methods to estimate static and dynamic models. The idea of the so called common correlated effects estimators is to add cross-sectional averages which approximate the cross-sectional dependence. This estimator was implemented into Stata in the static version by the user written command `xtmg` (Eberhardt, 2012) and in the dynamic version by `xtdcce2` (Ditzen, 2018).

Chudik et al. (2016) describe two methods to estimate the long run coefficients in large dynamic panel data models with cross-sectional dependence. This paper builds on their work and introduces an extended version of `xtdcce2`, which allows the estimation of the long run coefficients.¹

The remainder of the paper is structured as follows. The next section describes the general model and the common correlated effects estimator. Then three different methods to estimate the long run coefficients are discussed, first from a theoretical perspective, then from an applied perspective. Examples on how to estimate the equations using `xtdcce2` are given. The validity of the estimators is shown by a Monte Carlo Simulation and the paper closes with a conclusion.

2 Panel Model and CCE

There are several ways to introduce independence across cross-sectional units. If the dependence is observable, it is obvious to use spatial econometric methods. This paper focuses on unobservable dependence across units. Therefore the cross-sectional dependence is modelled as a common factor, which is part of the error term. Chudik et al. (2011) and Chudik and Pesaran (2015) assume a dynamic ARDL(1,1) panel model with heterogeneous coefficients in the form of:

$$y_{i,t} = \alpha_i + \lambda_i y_{i,t-1} + \beta_{0,i} x_{i,t} + \beta_{1,i} x_{i,t-1} + u_{i,t} \quad (1)$$

$$u_{i,t} = \sum_{l=1}^{m_f} \gamma_{i,l} f_{t,l} + e_{i,t} \quad (2)$$

with $i = 1, \dots, N$ and $t = 1, \dots, T_i$,

¹The estimation of long run coefficients is possible with `xtcce2` version 1.33 and later. This paper refers to version 1.35 or later. Please see the authors webpage www.jan.ditzen.net for updates.

where $e_{i,t}$ is a cross-section unit-specific IID error term. $\mathbf{f}_t = (f_{t,1}, \dots, f_{t,m_f})'$ is an unobserved common factors and m_f is the number of factors. The common factors are covariance stationary, have absolute summable autocovariances, are distributed independently over $e_{i,t}$ and the fourth order moments are bounded. Assume that $f_{t,l}$ is a strong common factor which is possibly correlated with the regressor $x_{i,t}$. The factor loadings $\boldsymbol{\gamma}_i = (\gamma_{i,1}, \dots, \gamma_{i,m_f})$ are heterogeneous and α_i is a unit-specific fixed effect. The heterogeneous coefficients are randomly distributed around a common mean, such that $\beta_i = \beta + v_i$, $v_i \sim IID(0, \boldsymbol{\Omega}_v)$, and $\lambda_i = \lambda + a_i$, $a_i \sim IID(0, \boldsymbol{\Omega}_a)$, where $\boldsymbol{\Omega}_v$ and $\boldsymbol{\Omega}_a$ are the variance covariance matrices. λ_i lies strictly inside the unit circle to ensure a non explosive series. In addition, the random deviation of λ_i and β_i are independently distributed of the error term and the common factors.

In case the explanatory variables are correlated with the common factors, leaving the factors out leads to an omitted variable bias. This renders the OLS residuals not iid anymore and OLS becomes inconsistent (Everaert and De Groote, 2016). Pesaran (2006) and Chudik and Pesaran (2015) show that Eq. (3) can be estimated consistently by approximating the common factors with cross sectional averages. In a dynamic model the floor of $\sqrt[3]{T}$ is added. The estimated equation becomes:

$$y_{i,t} = \alpha_i + \lambda_i y_{i,t-1} + \beta_{0,i} x_{i,t} + \beta_{1,i} x_{i,t-1} + \sum_{l=0}^{p_T} \delta'_{i,l} \bar{z}_{t-l} + e_{i,t} \quad (3)$$

where $\bar{z}_t = (\bar{y}_{t-1}, \bar{x}_t)$ are the cross sectional averages of the dependent and independent variables. This estimator is commonly called the common correlated effects mean group estimator (CCEMG estimator). Important to note is, that for the cross sectional averages, only the base of the variables are added, but not further lags in order to avoid multicollinearity problems. The mean group coefficient can then be estimated as the unweighted average of the cross sectional individual coefficients:

$$\hat{\pi}_{MG} = \left(\hat{\lambda}_{MG}, \hat{\beta}_{0,MG}, \hat{\beta}_{1,MG} \right) = \frac{1}{N} \sum_{i=1}^N \hat{\pi}_i \quad (4)$$

The asymptotic variance of the mean group coefficients is the unweighted variance of the cross sectional specific coefficients:

$$Var(\hat{\pi}_{MG}) = \frac{1}{N} \sum_{i=1}^N (\hat{\pi}_i - \hat{\pi}_{MG})(\hat{\pi}_i - \hat{\pi}_{MG})'$$

3 Estimating Long Run Relationships

Dynamic models allow the estimation of long run relationships. Long run relationships origin from an underlying steady state of a macroeconomic model. In

econometric terms, they measure the effect of an explanatory variable on the steady state value of the dependent variable. Three approaches how to estimate long run relationships will be explained next.

Following the notation from equation 1 and assuming that model is in its steady state with $y_t^* = y_{t-1}^* = y^*$ and $x_t^* = x_{t-1}^* = x^*$, the long run effect of variable x is defined as:

$$\theta_i = \frac{\beta_{0,i} + \beta_{1,i}}{1 - \lambda_i}.$$

Following the common correlated effects approach, the long run effects in equation 1 can be estimated in three ways, as a ECM or pooled mean group (PMG) model if it is an ARDL(1,1), or if it is a more general ARDL(p_y, p_x), directly without the short run coefficients (CS-DL) or indirectly with the short run coefficients (CS-ARDL).

3.1 Error Correction Approach

The first approach follows on the lines of Lee et al. (1997) and Shin et al. (1999), in which equation (3) is transformed into an error correction model:

$$\Delta y_{i,t} = \phi_i [y_{i,t-1} - \theta_{0,i} - \theta_{1,i}x_{i,t-1}] + \beta_{0,i}\Delta x_{i,t} + \sum_{l=0}^{p_T} \delta'_{i,l} \bar{z}_{t-l} + e_{i,t} \quad (5)$$

with

$$\begin{aligned} \theta_i &= \frac{\beta_{0,i} + \beta_{1,i}}{1 - \lambda_i} \\ \phi_i &= (1 - \lambda_i). \end{aligned} \quad (6)$$

In the case without dependence in the residuals and homogeneous long run coefficients ($\theta_i = \theta \ \forall \ i$), the model can be estimated in Stata by the pooled mean group (PMG) estimator, implemented by Blackburne and Frank (2007). With cross sectional averages, `xtddcce2` can estimated the equation using the `lr()` option. The ECM is transferred back into:

$$\begin{aligned} \Delta y_{i,t} &= \phi_i [y_{i,t-1} - \theta_{0,i} - \theta_{1,i}x_{i,t-1}] + \beta_{0,i}\Delta x_{i,t} + \sum_{l=0}^{p_T} \delta'_{i,l} \bar{z}_{t-l} + e_{i,t} \\ &= \alpha_i + \gamma_{1,i}y_{i,t-1} + \gamma_{2,i}x_{i,t-1} + \beta_{0,i}\Delta x_{i,t} + \sum_{l=0}^{p_T} \delta'_{i,l} \bar{z}_{t-l} + e_{i,t} \end{aligned} \quad (7)$$

where

$$\gamma_{1,i} = -\phi_i \quad \theta_{0,i} = \frac{\alpha_i}{-\gamma_{1,i}} \quad \theta_i = \frac{\gamma_{2,i}}{-\gamma_{1,i}} \quad (8)$$

$\phi_i = (1 - \lambda_i)$ is the error-correction speed of adjustment parameter and $[y_{i,t-1} - \theta_{0,i} - \theta_{1,i}x_{i,t-1}]$ is the error correction term. In general, a long run

relationship exists if $\phi \neq 0$ (Shin et al., 1999). $\beta_{0,i}$ captures the immediate or short run effect of $x_{i,t}$ on $y_{i,t}$. The long run or equilibrium effect is captured by θ_i . The long run effect measures how the equilibrium changes and ϕ represents how fast the adjustment occurs.

Besides the inclusion of the cross sectional averages, there is another important difference. `xtpmg` calculates the long run coefficients using maximum likelihood. `xtdcce2` treats the long run coefficients, defined in `lr()`, as further covariates and estimates equation 8 entirely by OLS. To calculate the long run coefficients, the unit specific coefficients are divided by the negative of the long run cointegration vector such that $\theta_{1,i} = -\gamma_{2,i}/\phi_i$. The variances are calculated using the Delta method. After the long run unit specific coefficients are obtained, the mean groups are calculated.

The PMG estimator assumes homogeneous long run and heterogeneous short run coefficients. `xtdcce2` is build to handle both coefficients to be heterogeneous. If the long run coefficient is pooled, then it takes the mean group estimate of the error speed of correction term. Thus the long run coefficient then becomes $\theta_1^p = -\gamma_2^p/\phi_{MG}$.

Following Lee et al. (1997) and Shin et al. (1999), Blackburne and Frank (2007) explain the use of `xtpmg` by estimating a long-run consumption function in the form of:

$$c_{i,t} = \theta_{0t} + \theta_{1t}y_{i,t} + \theta_{2t}\pi_{i,t} + \mu_i + \epsilon_{i,t}, \quad (9)$$

where $c_{i,t}$ log of consumption per capita, $y_{i,t}$ is log of real per capita income and $\pi_{i,t}$ is the inflation rate. The ARDL(1,1,1) of the specification is

$$c_{i,t} = \delta_{0,i} + \lambda_i c_{i,t-1} + \delta_{10,i} y_{i,t} + \delta_{1,i} y_{i,t-1} + \delta_{20,i} \pi_{i,t} + \delta_{2,i} \pi_{i,t-1} + \epsilon_{i,t}$$

and ECM representation of the model reads:²

$$\Delta c_{i,t} = \phi_i (c_{i,t-1} - \theta_{1,i} y_{i,t} - \theta_{2,i} \pi_{i,t}) + \delta_{0,i} + \delta_{1,i} \Delta y_{i,t} + \delta_{2,i} \Delta \pi_{i,t} + \epsilon_{i,t}. \quad (10)$$

`xtdcce2` internally estimates (leaving out any cross sectional means):

$$\Delta c_{i,t} = \phi_i c_{i,t-1} + \gamma_{1,i} y_{i,t} + \gamma_{2,i} \pi_{i,t} + \delta_{0,i} + \delta_{1,i} \Delta y_{i,t} + \delta_{2,i} \Delta \pi_{i,t} + \epsilon_{i,t}, \quad (11)$$

Using the *jasa2* dataset, the syntax for `xtdcce2` is then:

```
xtdcce2 depvar[sr_indepvars] , lr(lr_indepvars) [lr_options(string)]
```

where *sr_indepvars* are the short run variables and the long run are defined in the option `lr()`. As the underlying model is an ARDL(1,1,1) model, the first difference of the explanatory variables is added to the short run part, directly behind the dependent variable. For an ARDL(1,0,0) model, the differences would be omitted, for an ARDL(1,2,2) model, the second difference would be added. The appendix 7.2 gives further details. Cross sectional averages are defined by the `crosssectional()` option. As all variables are included, the

²The details of the transformation from the ARDL model to the ECM can be found in the appendix in 7.2.

shortcut `_all` is used, which adds the base of all variables. The entire command line and the output are:³⁴

```
. xtdcce2135 d.c d.pi d.y if year >= 1962, ///
> lr(L.c pi y) p(l.c pi y) cr(_all) cr_lags(2)
(Dynamic) Common Correlated Effects Estimator - Pooled Mean Group
Panel Variable (i): id                Number of obs   =    719
Time Variable (t): year                Number of groups =    24
Degrees of freedom per group:          Obs per group:
without cross-sectional avg. min = 23      min = 29
                                max = 24      avg = 30
with cross-sectional avg. min = 14         max = 30
                                max = 15
Number of                               F(291, 428)     =    3.61
cross-sectional lags                    = 2            Prob > F       =    0.00
variables in mean group regression = 51            R-squared      =    0.71
variables partialled out                 = 240          Adj. R-squared =    0.51
                                           Root MSE      =    0.02
                                           CD Statistic  =   -0.25
                                           p-value      =    0.8062
```

D.c	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
Short Run Est.						
Mean Group:						
D.pi	.0341953	.0343376	1.00	0.319	-.0331051	.1014958
D.y	.1658292	.0361585	4.59	0.000	.0949598	.2366986
Long Run Est.						
Pooled:						
L.c	-.4691509	.1162393	-4.04	0.000	-.6969757	-.241326
pi	-.1557539	.3383306	-0.46	0.645	-.8188697	.5073619
y	.8547452	.0984223	8.68	0.000	.661841	1.047649

```
Pooled Variables: L.c pi y
Mean Group Variables: D.pi D.y
Cross Sectional Averaged Variables: pi y c
Long Run Variables: L.c pi y
Cointegration variable(s): L.c
Heterogenous constant partialled out.
```

`xtdcce2` estimates equation 8 entirely by OLS. To calculate the long run coefficients, the coefficients are divided by the negative of the long run cointegration vector to match equation 10, $\theta_{1,i} = -\gamma_{1,i}/\phi_i$. The variances are calculated using the Delta method as described in the Appendix.

First the long run coefficients for each cross section are computed and in a second step the individual long run coefficients are averaged. As an example, the average long run coefficient for $\hat{\theta}_1$ is calculated as: $\hat{\theta}_1 = 1/N \sum_{i=1}^N \hat{\theta}_{1,i} = 1/N \sum_{i=1}^N (-\hat{\gamma}_{1,i}/\hat{\phi}_i)$. In the example here, ϕ is heterogeneous, but θ_1 homogeneous. Thus the long run coefficient θ_1 is calculated as $\theta_1 = -\gamma_1/(1/N \sum_{i=1}^N \phi_i)$.

³Following $p = [T^{1/3}]$ would result in 3 lags. However as the panel is relatively short with $T = 31$ and the CD test statistic is in a non rejection region, the number of lags is set to 2.

⁴The syntax for `xtdcce2` is discussed in the appendix 7.1.

An estimation of the same model, but with heterogeneous long and short run coefficients leads to the following result:

```
. xtddce2135 d.c d.pi d.y if year >= 1962, ///
> lr(L.c pi y) cr(_all) cr_lags(2)
(Dynamic) Common Correlated Effects Estimator - Mean Group
Panel Variable (i): id          Number of obs   =      719
Time Variable (t): year        Number of groups =      24
Degrees of freedom per group:  Obs per group:
without cross-sectional avg.  min   = 23          min   = 29
                             max   = 24          avg   = 30
with cross-sectional avg.     min   = 14          max   = 30
                             max   = 15
Number of                      F(360, 359)     = 3.82
cross-sectional lags           = 2             Prob > F       = 0.00
variables in mean group regression = 120          R-squared      = 0.79
variables partialled out       = 240          Adj. R-squared = 0.59
                             Root MSE       = 0.01
                             CD Statistic    = -0.17
                             p-value        = 0.8638
```

D.c	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
Short Run Est.						
Mean Group:						
D.pi	.0701836	.0424275	1.65	0.098	-.0129727	.15334
D.y	.0630114	.0514939	1.22	0.221	-.0379147	.1639375
Long Run Est.						
Mean Group:						
L.c	-.5849174	.0453523	-12.90	0.000	-.6738062	-.4960286
pi	-.3760294	.1470868	-2.56	0.011	-.6643143	-.0877445
y	.7848242	.0779883	10.06	0.000	.6319699	.9376785

```
Mean Group Variables: D.pi D.y
Cross Sectional Averaged Variables: pi y c
Long Run Variables: L.c pi y
Cointegration variable(s): L.c
Heterogenous constant partialled out.
```

The long run coefficients slightly changed and are in absolute value higher.

3.2 ARDL Approach

The more general representation of eq (1) as an ARDL(p_y, p_x) model is:

$$y_{i,t} = \alpha_i + \sum_{l=1}^{p_y} \lambda_{l,i} y_{i,t-l} + \sum_{l=0}^{p_x} \beta_{l,i} x_{i,t-l} + u_{i,t}. \quad (12)$$

Next, two further methods to estimate the long run coefficients are explained. The first one is the CS-DL method and directly estimates the long run coefficients. The second method, the CS-ARDL estimator, uses an auxiliary regression.

3.2.1 CS-DL

Under the assumption that λ_i lies in the unit circle, the general representation of an ARDL(p_y, p_x) model can be written as:

$$y_{i,t} = \theta_{0,i} + \theta_{1,i}x_{i,t} + \delta_i(L)\Delta x_{i,t} + \tilde{u}_{i,t} \quad (13)$$

where

$$\delta_i(L) = - \sum_{l=0}^{\infty} \left[\lambda_i^{l+1} (1 - \lambda_i)^{-1} \beta_{1,i} \right] L^l \quad (14)$$

$$\theta_{0,i} = (1 - \lambda_i L)^{-1} \alpha_i \quad (15)$$

$$\tilde{u}_{i,t} = (1 - \lambda_i L)^{-1} u_{i,t} \quad (16)$$

$$(17)$$

and L is the lag operator. Chudik et al. (2016) show that Estimation (13) can be directly estimated by the common correlated effects estimator. The regression is augmented by the differences of the explanatory variables (x), their lags and the cross-sectional averages. Following Pesaran (2006) the estimation is consistent even if the errors are serially correlated.

For a general ARDL(p_y, p_x) model with added cross-sectional averages to take out the cross-sectional dependence, the estimated equation is:

$$y_{i,t} = \theta_{0,i} + \theta_{1,i}x_{i,t} + \sum_{l=0}^{p_x-1} \delta_{i,l}\Delta x_{i,t-l} + \sum_{l=0}^{p_{\bar{y}}} \gamma_{y,i,l}\bar{y}_{i,t-l} + \sum_{l=0}^{p_{\bar{x}}} \gamma_{x,i,l}\bar{x}_{i,t-l} + e_{i,t}, \quad (18)$$

where $\bar{y}_{i,t-l}$ and $\bar{x}_{i,t-l}$ are the cross sectional averages and $p = p_{\bar{x}} = [T^{1/3}]$ and $p_{\bar{y}} = 0$. The mean group coefficients are calculated as

$$\hat{\theta}_{0,MG} = \frac{1}{N} \sum_{i=1}^N \hat{\theta}_{0,i} \quad \text{and} \quad \hat{\theta}_{1,MG} = \frac{1}{N} \sum_{i=1}^N \hat{\theta}_{1,i}. \quad (19)$$

The variance estimator is the same as in equation (4), but with the mean group of the long run coefficients rather than of the short run coefficients:

$$Var(\hat{\theta}_{MG}) = \frac{1}{N} \sum_{i=1}^N \left(\hat{\theta}_i - \hat{\theta}_{MG} \right) \left(\hat{\theta}_i - \hat{\theta}_{MG} \right)'. \quad (20)$$

A pooled estimation is possible as well. The estimator is the same as in Pesaran (2006) and explained for `xtcce2` in Ditzen (2018).

Estimating the consumption model from equation (10), implies the following equation with an ARDL(1,1,1) model with $p_y = p_c = p_\pi = 2$ and $p = 1$:

$$c_{i,t} = \alpha_i + \theta_{1,i}y_{i,t} + \theta_{2,i}\pi_{i,t} + \sum_{l=0}^2 \delta_{i,l}\Delta \mathbf{V}_{i,t-l} + \sum_{l=0}^2 \gamma_{y,i,l}\bar{c}_{i,t-l} + \sum_{l=0}^2 \gamma_{x,i,l}\bar{\mathbf{V}}_{i,t-l} + e_{i,t}$$

where $\mathbf{V}_{i,t} = (y_{i,t}, \pi_{i,t})$. The model can be estimated using `xtdcce2`, by adding the first difference of the variables aggregated in $\mathbf{V}_{i,t}$, namely $\Delta y_{i,t}$ and $\Delta \pi_{i,t}$ as further explanatory variables, but they are excluded from the cross-sectional averages. As the long run effects are directly estimated, setting `lr()` is not necessary. The output is:

```
. xtdcce2135 c pi y d.pi d.y d.(d.pi d.y) if year >= 1962, ///
> p(pi y d.pi d.y d.(d.pi d.y)) cr(c pi y) cr_lags(2) pooledc
(Dynamic) Common Correlated Effects Estimator - Pooled

Panel Variable (i): id                Number of obs   =       719
Time Variable (t): year                Number of groups =        24
Degrees of freedom per group:          Obs per group:
  without cross-sectional avg. min     = 22                min =       29
                                max     = 23                avg  =       30
  with cross-sectional avg.   min     = 13                max  =       30
                                max     = 14
Number of cross-sectional lags         = 2                F(223, 496)   =       5.23
variables in mean group regression = 7                Prob > F      =       0.00
variables partialled out              = 216             R-squared     =       0.70
                                Adj. R-squared =       0.57
                                Root MSE     =       0.03
                                CD Statistic  =      -0.89
                                p-value       =       0.3728
```

	c	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
Pooled:							
	pi	-.3054825	.0831854	-3.67	0.000	-.4685228	-.1424422
	y	.7540553	.1297665	5.81	0.000	.4997176	1.008393
	D.pi	.2381957	.147126	1.62	0.105	-.050166	.5265574
	D.y	-.3622873	.1202688	-3.01	0.003	-.5980097	-.1265648
	D2.pi	-.0745891	.0879009	-0.85	0.396	-.2468716	.0976934
	D2.y	.0766291	.0846078	0.91	0.365	-.0891992	.2424573

```
Pooled Variables: pi y D.pi D.y D2.pi D2.y
Cross-sectional Averaged Variables: pi y c
Homogenous constant not displayed.
```

The coefficient on variable π remains insignificant, the coefficient on y is positive and significant. The results for the two coefficients are similar as those obtained in (Blackburne and Frank, 2007, p. 203). The variables in first differences are only auxiliary and can be ignored.

A second example can be found in Chudik et al. (2013). The authors estimate

the long run effect of public debt on output growth with the following equation:

$$\Delta y_{i,t} = c_i + \theta'_i \mathbf{x}_{i,t} + \sum_{l=0}^{p-1} \gamma_{i,l} \Delta \mathbf{x}_{i,t-l} + \omega_{i,y} \Delta \bar{y}_t + \sum_{l=0}^3 \omega_{i,l} \bar{\mathbf{x}}_{i,t-l} e_{i,t}$$

where $y_{i,t}$ is the log of real GDP, and $\mathbf{x}_{i,t} = (\Delta d_{i,t}, \pi_{i,t})'$, $d_{i,t}$ is log of debt to GDP ratio and π is the inflation rate and p the number of lags. The results from Chudik et al. (2013, Table 18) with 1 lag ($p = 1$) in the form of an ARDL(1,1,1) model can be replicated as follows, where variable dp is π and debt is gd :⁵

```
. xtccce2135 d.y dp d.gd d.(dp d.gd) ///
> , cr(d.y dp d.gd) cr_lags(0 3 3) fullsample
(Dynamic) Common Correlated Effects Estimator - Mean Group
Panel Variable (i): ccode                Number of obs   =       1601
Time Variable (t): year                  Number of groups =         40
Degrees of freedom per group:
without cross-sectional averages         = 35.025
with cross-sectional averages            = 26.025
Number of                                F(560, 1041)     =         0.90
cross-sectional lags                     0 to 3          Prob > F         =         0.93
variables in mean group regression        = 160            R-squared        =         0.33
variables partialled out                  = 400           Adj. R-squared   =        -0.04
                                           Root MSE       =         0.03
                                           CD Statistic   =         1.11
                                           p-value       =         0.2667
```

	D.y	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
Mean Group:						
dp		-.0889339	.0256445	-3.47	0.001	-.1391961 -.0386717
D.gd		-.0865123	.0143	-6.05	0.000	-.1145398 -.0584849
D.dp		.0053284	.0413629	0.13	0.897	-.0757413 .0863981
D2.gd		.0068065	.0148306	0.46	0.646	-.022261 .035874

```
Mean Group Variables: dp D.gd D.dp D2.gd
Cross Sectional Averaged Variables: D.y(0) dp(3) D.gd(3)
Heterogenous constant partialled out.
```

The different lag structure is imposed adding a *numlist* to the `cr_lags()` option, rather than a single number. The first differences as part of the vector $\Delta \mathbf{x}_{i,t}$ are added as `d.(dp d.gd)`. The `fullsample` option is used to make use of the entire sample. Using three rather than one lag for the differences, hence an ARDL(3,3,3) model can be estimated as well, `d.(dp d.gd)` replaces `L(0/2).d.(dp d.gd)`:

```
. xtccce2135 d.y dp d.gd L(0/2).d.(dp d.gd) ///
> , cr(d.y dp d.gd) cr_lags(0 3 3) fullsample
(Dynamic) Common Correlated Effects Estimator - Mean Group
Panel Variable (i): ccode                Number of obs   =       1571
Time Variable (t): year                  Number of groups =         40
Degrees of freedom per group:
without cross-sectional averages         = 30.275
Obs per group (T) =                      39
```

⁵Note that the dependent variable $\delta y_{i,t}$ and the inflation rate $\delta \pi_{i,t}$ are in first differences. Those differences are not generated as new variables, Stata's time series operators are used.

with cross-sectional averages	= 21.275	F(720, 851)	=	1.12
Number of cross-sectional lags	0 to 3	Prob > F	=	0.06
variables in mean group regression	= 320	R-squared	=	0.49
variables partialled out	= 400	Adj. R-squared	=	0.05
		Root MSE	=	0.03
		CD Statistic	=	0.73
		p-value	=	0.4681

D.y	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
Mean Group:						
dp	-.0855843	.0400844	-2.14	0.033	-.1641483	-.0070202
D.gd	-.0816584	.0196252	-4.16	0.000	-.1201231	-.0431936
D.dp	.0183591	.0478696	0.38	0.701	-.0754636	.1121817
LD.dp	.0015577	.0373619	0.04	0.967	-.0716704	.0747857
L2D.dp	.0034021	.029477	0.12	0.908	-.0543718	.0611759
D2.gd	.0045224	.0144741	0.31	0.755	-.0238464	.0328912
LD2.gd	-.0129676	.0134554	-0.96	0.335	-.0393396	.0134044
L2D2.gd	-.0095151	.0090814	-1.05	0.295	-.0273142	.0082841

Mean Group Variables: dp D.gd D.dp LD.dp L2D.dp D2.gd LD2.gd L2D2.gd
Cross Sectional Averaged Variables: D.y(0) dp(3) D.gd(3)
Heterogenous constant partialled out.

The number of variables becomes rather large, but only the first two coefficients are of interest.

3.2.2 CS-ARDL

An alternative to the CS-DL estimator is the CS-ARDL approach. In a first step the short run coefficients are estimated and then the long run coefficients calculated. The advantage of this approach is that a full set of estimates for the long and the short run coefficients is obtained. The CS-ARDL model can be rewritten as an ECM and therefore the long run estimates are numerically equivalent.

Extending equation (12) with the cross-sectional averages to accommodate cross-sectional dependence leads to:

$$y_{i,t} = \alpha_i + \sum_{l=1}^{p_y} \lambda_{l,i} y_{i,t-l} + \sum_{l=0}^{p_x} \beta_{l,i} x_{i,t-l} + \sum_{l=0}^p \bar{v}_{t-l} + e_{i,t}. \quad (21)$$

with

$$\bar{v}_{t-l} = (\bar{y}_{i,t-l}, \bar{x}_{i,t-l})$$

The long run coefficients are calculated as:

$$\hat{\theta}_{CS-ARDL,i} = \frac{\sum_{l=0}^{p_x} \hat{\beta}_{l,i}}{1 - \sum_{l=1}^{p_y} \hat{\lambda}_{l,i}} \quad (22)$$

A disadvantage of this approach is that p_x and p_y need to be known. `xtdcce2` supports only the calculation of heterogeneous coefficients. It can calculate

pooled coefficients in an ARDL model, however no standard errors have been derived for the long run coefficient. `xtdcce2` uses the delta method to calculate the standard errors of the pooled long run coefficients.

Besides the CS-DL model, Chudik et al. (2013) estimate an ARDL model.⁶ Using `xtdcce2`, all short run variables are added to the `lr` option and the `ardl` routine is invoked by using `lr_options(ardl)`. There is no need to specify the long run variables separately, because `xtdcce2` automatically detects the common base of variables if time series operators are used. If variables are created for the lags via `gen lx = L.x`, then the variables with the same base which form a long run coefficient need to be enclosed in parenthesis, for example `lr((y ly) (x lx))`.

To replicate the results from Table 17 with 1 lag, hence an ARDL(1,1,1) model, the first lag of the dependent and the base and the first lag of the dependent variables are added:

```
. xtdcce2135 d.y , lr(L.d.y L.dp dp L.d.gd d.gd) ///
> lr_options(ardl) cr(d.y dp d.gd) cr_lags(3) ///
> fullsample
(Dynamic) Common Correlated Effects Estimator - (CS-ARDL)
Panel Variable (i): ccode                Number of obs   =       1599
Time Variable (t): year                  Number of groups =        40
Degrees of freedom per group:
without cross-sectional averages         = 33.975
with cross-sectional averages           = 21.975
Number of
cross-sectional lags                     = 3
variables in mean group regression       = 200
variables partialled out                 = 520
F(720, 879)                             =        0.79
Prob > F                                 =        1.00
R-squared                                =        0.39
Adj. R-squared                           =       -0.11
Root MSE                                 =        0.03
CD Statistic                             =        0.57
p-value                                  =       0.5690
```

D.y	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
Short Run Est.						
Mean Group:						
LD.y	.0475614	.0393514	1.21	0.227	-.0295659	.1246888
dp	-.1036029	.0402888	-2.57	0.010	-.1825675	-.0246383
D.gd	-.0745686	.0122305	-6.10	0.000	-.0985398	-.0505973
L.dp	-.0199465	.0462873	-0.43	0.667	-.1106679	.0707749
LD.gd	-.0132482	.0156115	-0.85	0.396	-.0438463	.0173498
Long Run Est.						
Mean Group:						
lr_dp	-.1639748	.0378594	-4.33	0.000	-.2381778	-.0897718
lr_gd	-.0873991	.0164432	-5.32	0.000	-.1196271	-.0551711
lr_y	-.9524386	.0393514	-24.20	0.000	-1.029566	-.8753112

Mean Group Variables:
Cross Sectional Averaged Variables: D.y dp D.gd

⁶Chudik et al. (2013) use `xtpmg`. To do so they re-write the ARDL(p_y, p_x) model into an error correction model.

Long Run Variables: lr_dp lr_gd lr_y
 Cointegration variable(s): lr_y
 Heterogenous constant partialled out.

The long run coefficients are labelled by the prefix lr_. The mean group coefficients are calculated as the average of the cross-section individual coefficients. For the ARDL(3,3,3) the 3 lags of the explanatory variables and the dependent variable are added. To improve readability, the different bases are enclosed into parenthesis:

```
. xtccce2135 d.y , cr_lags(3) fullsample ///
> lr(L(1/3).(d.y) (L(0/3).dp) (L(0/3).d.gd) ) ///
> lr_options(ardl) cr(d.y dp d.gd)
(Dynamic) Common Correlated Effects Estimator - (CS-ARDL)
Panel Variable (i): ccode                Number of obs   =       1562
Time Variable (t): year                  Number of groups =        40
Degrees of freedom per group:
  without cross-sectional averages      = 27.05
  with cross-sectional averages         = 15.05
Number of cross-sectional lags          = 3
variables in mean group regression      = 440
variables partialled out                = 520
F(960, 602)                             =       0.96
Prob > F                                 =       0.71
R-squared                                 =       0.61
Adj. R-squared                           =      -0.03
Root MSE                                 =       0.02
CD Statistic                             =      -0.51
p-value                                  =       0.6108
```

D.y	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
Short Run Est.						
Mean Group:						
LD.y	.0123776	.0349374	0.35	0.723	-.0560984	.0808536
L2D.y	-.1395721	.0948493	-1.47	0.141	-.3254733	.046329
L3D.y	-.0829106	.1072972	-0.77	0.440	-.2932092	.1273881
dp	-.0707066	.0503045	-1.41	0.160	-.1693015	.0278883
D.gd	-.0853072	.0137595	-6.20	0.000	-.1122754	-.0583391
L.dp	-.0312738	.0513445	-0.61	0.542	-.1319071	.0693595
L2.dp	.098219	.101743	0.97	0.334	-.1011937	.2976317
L3.dp	-.0424672	.0581718	-0.73	0.465	-.1564818	.0715474
LD.gd	-.0270313	.0204755	-1.32	0.187	-.0671624	.0130999
L2D.gd	-.0114101	.012726	-0.90	0.370	-.0363525	.0135324
L3D.gd	.0283559	.0177672	1.60	0.110	-.0064671	.0631789
Long Run Est.						
Mean Group:						
lr_dp	-.0795232	.0587003	-1.35	0.176	-.1945738	.0355274
lr_gd	-.1198351	.0402246	-2.98	0.003	-.1986738	-.0409964
lr_y	-1.210105	.2006012	-6.03	0.000	-1.603276	-.8169339

Mean Group Variables:
 Cross Sectional Averaged Variables: D.y dp D.gd
 Long Run Variables: lr_dp lr_gd lr_y
 Cointegration variable(s): lr_y
 Heterogenous constant partialled out.

As a final exercise, the ARDL(1,1,1) model used by Blackburne and Frank (2007) is estimated, but using the CS-ARDL estimator and heterogeneous co-

efficients:

```
. xtdcce2135 c if year >= 1962, ///
> lr(L.c pi L.pi L(0/1).y) lr_options(ardl) ///
> cr(_all) cr_lags(2)
(Dynamic) Common Correlated Effects Estimator - (CS-ARDL)
Panel Variable (i): id                Number of obs   =       719
Time Variable (t): year                Number of groups =        24
Degrees of freedom per group:
without cross-sectional avg. min   = 23
                                max   = 24
with cross-sectional avg. min     = 14
                                max   = 15
Number of
cross-sectional lags                = 2
variables in mean group regression = 120
variables partialled out            = 240
F(360, 359)                         = 4.96
Prob > F                             = 0.00
R-squared                            = 0.83
Adj. R-squared                       = 0.66
Root MSE                             = 0.01
CD Statistic                         = -0.17
p-value                              = 0.8638
```

	c	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
Short Run Est.							
Mean Group:							
	L.c	.4150826	.0453523	9.15	0.000	.3261938	.5039714
	pi	-.0948748	.0586788	-1.62	0.106	-.2098832	.0201335
	y	.5397594	.0527499	10.23	0.000	.4363715	.6431474
	L.pi	-.0701836	.0424275	-1.65	0.098	-.15334	.0129727
	L.y	-.0630114	.0514939	-1.22	0.221	-.1639375	.0379147
Long Run Est.							
Mean Group:							
	lr_c	-.5849174	.0453523	-12.90	0.000	-.6738062	-.4960286
	lr_pi	-.3760294	.1470868	-2.56	0.011	-.6643143	-.0877445
	lr_y	.7848242	.0779883	10.06	0.000	.6319699	.9376785

```
Mean Group Variables:
Cross Sectional Averaged Variables: pi y c
Long Run Variables: lr_c lr_pi lr_y
Cointegration variable(s): lr_c
Heterogenous constant partialled out.
```

All long run coefficients are the same. The error speed of correction is displayed as variable `lr_c`, which is $1 - 0.4150$, the coefficient of `L.c`.

4 Monte Carlo Simulation

The aims of the exercise are several fold. First of all the bias of the long run coefficients is assessed. Additionally the standard errors of the long run coefficients are compared to their true values. For an empirical analysis, standard errors are important.

The simulation is carried out on the lines of Chudik et al. (2016) and the

underlying DGP is an ARDL(2,1) model:⁷

$$\begin{aligned} y_{i,t} &= c_{yi} + \lambda_{1,i}y_{i,t-1} + \lambda_{2,i}y_{t-2} + \beta_{0,i}x_{i,t} + \beta_{1,i}x_{i,t-1} + u_{i,t} \\ u_{i,t} &= \gamma'_i f_t + \epsilon_{i,t} \\ x_{i,t} &= c_{xi} + \alpha_{xi}y_{i,t-1} + \gamma_{xi}f_t + v_{xi,t} \end{aligned}$$

$y_{i,t}$ is the dependent variable and $x_{i,t}$ the only independent variable. For a matter of ease, it is assumed that only one explanatory variable exists and only one common factor exists. To generate the long run coefficients, first the long run coefficient θ_i is drawn and then the short run coefficients are backed out, such that:⁸

$$\begin{aligned} \theta_i &\sim IIDN(1, \sigma_\theta^2), & \lambda_{1,i} &= (1 + \xi_{\lambda i})\eta_{\lambda i}, & \lambda_{2,i} &= -\xi_{\lambda i}\eta_{\lambda i} \\ \beta_{0,i} &= \xi_{\beta i}\eta_{\beta i}, & \beta_{1,i} &= (1 - \xi_{\beta i})\eta_{\beta i}, & \eta_{\lambda i} &= IIDU(0, \lambda_{max}) \\ \eta_{\beta i} &= \theta_i(1 - \lambda_{i,1} - \lambda_{2,i}), & \xi_{\lambda i} &\sim IIDU(0.2, 0.3), & \xi_{\beta i} &\sim IIDU(0, 1) \end{aligned}$$

To allow for a wide range of different results, two sets of coefficients are produced. One in which the long run coefficients vary a lot and another one with little variation. This implies that $\eta_{\lambda,i}$ is varied as well. The set is varied between $(\sigma_\theta^2, \lambda_{max}) = (0.2, 0.6)$ or $(0.8, 0.8)$. The fixed effects are drawn as $c_{yi} \sim IIDN(0, 1)$ and $c_{xi} = c_{yi} + \varrho_{c_{xi}}, \varrho_{c_{xi}} \sim IIDN(0, 1)$.

The common factors are potentially correlated over time and calculated as $f_t = \rho_f f_{t-1} + \varsigma_{ft}, \varsigma_{ft} \sim IIDN(0, 1 - \rho_f^2)$. The heterogeneous factor loading is generated as $\gamma_i = \gamma + \eta_{i\gamma}, \eta_{i\gamma} \sim IIDN(0, \sigma_\gamma^2)$ and $\gamma_x = \sqrt{b_x}, b_x = \frac{2}{m(m+1)} - \frac{2}{m+1}\sigma_{\gamma_x}^2$ with $m = 1$ and $\sigma_\gamma^2 = \sigma_{\gamma_x}^2 = 0.2^2$. Two scenarios for ρ_f are considered, one with correlated factors $\rho_f = 0.6$ and one with uncorrelated factors, $\rho_f = 0$.

The error terms $\epsilon_{i,t}$ and $v_{xi,t}$ are generated such that they allow for autocorrelation. $\epsilon_{i,t}$ also can inhibit cross-sectional dependence. For a further discussion of the parametrisation see the Appendix 7.4.

Multiple potential biases matter for the estimation of the coefficients. In a small, finite sample the bias of the short run CCE estimators potentially arises from three sources: the length of the time series, cross-sectional dependence and heterogeneous slope coefficients. The first source relates to small T and the time series bias of order T^{-1} (Hurwicz bias) and is expected to decrease with $T \rightarrow \infty$. The heterogeneous coefficients are assumed to be randomly distributed around a common mean. As N converges to infinity, the mean group estimate converges to its true parameter, ignoring any influence from the other two biases. Therefore the bias due to heterogeneous coefficients is expected to decrease with an increase in N . The bias due to cross-sectional dependence needs to be separated further. For a given T and an increase in N , weak cross sectional

⁷This paper focuses on the baseline cases with heterogeneous slopes and stationary factors.

⁸Chudik et al. (2016, p. 103) have a for $\eta_{\beta i} = \theta_i / (1 - \lambda_{i,1} - \lambda_{2,i})$, thus they divide instead of multiply. However given that $\eta_{\beta i} = \beta_{0,i} + \beta_{1,i}$ and $\theta_i = \eta_{\beta i} / (1 - \lambda_{i,1} - \lambda_{2,i})$, a multiplier is correct.

dependence declines. Strong cross-sectional dependence does not decline when N increases, but should not pose a problem as the cross-sectional averages take it out.

As the long run coefficients of the CS-ECM and the CS-ARDL are the same, the Monte Carlo focuses on the CS-DL and CS-ARDL estimator. Four different specifications are considered. The baseline specification has correlated factors and low values for the variance of the long run coefficients ($\sigma_\theta^2 = 0.2$) and is equally spread across the autocorrelation coefficients ($\lambda_{max} = 0.6$). The second specification includes low values of $(\sigma_\theta^2, \lambda_{max})$, but correlated factors ($\rho_f = 0.6$). The last two specifications have high values for $(\sigma_\theta^2, \lambda_{max}) = (0.8, 0.8)$ and include correlated and none correlated factors. The CS-DL and CS-ARDL, and for Specification 1 results for the CS-ECM estimator are presented. Table 1 summarizes the specifications:

[TABLE 1 here]

4.1 Results

Table 2 shows the bias of the mean group coefficient of the long run coefficient and Table 3 the bias of the standard errors. To allow a complete picture, Table 2 shows the results for the CS-DL, CS-ARDL and CS-ECM estimator. In order to shed light on a misspecified model, the bias of short run coefficients of a CS-ARDL estimation is presented in Table 4.

[TABLE 2 here]

The first panel of Table 2 displays the bias of the mean group estimates of the long run coefficients using the CS-DL estimator. The bias decreases with an increase in T , however remains relatively stable with an increase in the number of cross sections. An improvement of the estimates with an increase in T is confirmed by the decrease of the root mean squared error. The CS-ARDL and CS-ECM results are identical. The level of the bias of those is much smaller, in a region well below 5%. The root mean squared error benefits from a larger number of time periods and points that in large panels the estimates become more stable.

[TABLE 3 here]

The next table presents the bias of the estimates of the standard error. The bias is calculated as the estimated standard error $\widehat{SE}(\hat{\theta}_{MG})$ minus its true value. The standard errors turn out to be upwards biased and increase slightly with a larger number of time periods. The upward bias is preferable over an underestimation. However the bias reaches almost 100%, implying that t-statistics become small and type 2 errors are more likely.

[TABLE 4 here]

The bias for the long run estimations of the CS-ARDL is influenced by the bias of the short run estimates. Therefore a difference in the bias between the two estimators should give some insights into the importance of estimating the short run coefficients as well. Table 4 displays the short run coefficients of the CS-ARDL estimator. The bias for the contemporaneous value of β_{MG} is small, the bias of the coefficient of the lagged variable is larger. Both coefficients are exposed to a smaller bias if the number of time periods is increased. The coefficients of the lags of the dependent variable are biased towards different directions and both decline with T .

Overall the results are interesting in the sense that the CS-ARDL estimator outperforms the CS-DL estimator. This is despite the fact that the CS-DL estimator estimates directly the long run coefficients, while the CS-ARDL requires an estimation of the short run coefficients first. A possibility is, that the bias of the short run coefficients cancel each other out, a possibility seen by the different directions of the biases of the short run coefficient of the lagged dependent variable. Additionally the CS-ARDL estimator needs to be specified more carefully with respect to the number of lags of the independent and dependent variables. These additional set of information might benefit the bias of the estimator. This topic will be more discussed in the next section.

The results show, that all three estimator produce reasonable results, even in small samples. However, two drawbacks remain, first of all the CS-DL estimator performs worse than the CS-ARDL estimator and secondly the standard errors are overestimated for all three estimators.⁹

4.1.1 Misspecified Model

As a robustness check to shed light on the behaviour of the estimators if the number of lags of the dependent and independent variable is accidentally larger than of the DGP. Thus instead of the following command line:

```
xtdcce2135 y , lr(L(1/2).y x L.x) lr_options(ardl) cr_lags(cr_lags) cr(y x )
xtdcce2135 y x d.x , cr_lags(0 cr_lags) cr(y x )
```

for the CS-ARDL and CS-DL estimator, the command lines are:

```
xtdcce2135 y , lr(L(1/3).y x L.x L2.x) lr_options(ardl) cr_lags(cr_lags) cr(y x )
xtdcce2135 y x d.x d2.x , cr_lags(0 cr_lags) cr(y x )
```

An additional lag of the dependent and independent variables is added for each model. Table 14 shows the results of the simulation with the misspecified estimation model. The bias for the CS-DL estimator is closer to zero than for the correctly specified model. The CS-ARDL estimator suffers more, however the bias remains substantially small. This points that the CS-ARDL estimator

⁹Tables 5 - 10 show the results for Specification 2 to 4 and are available in the appendix. In general results are similar to those from Specification 1 and therefore detailed analysis is left to the reader.

is more vulnerable to misspecifications, while the CS-DL estimator seems to be more robust.

[TABLE 14 here]

5 Conclusion

This paper reviewed three different methods to estimate long run coefficients in dynamic panels with a large number of observations over time and cross-sectional units and with cross-sectional dependence. It uses an extended version of `xtdcce2` (Ditzen, 2018) which allows for the estimation of long run coefficients using the CS-DL, CS-ARDL and CS-ECM estimator. Examples on how to apply `xtdcce2` are given and options explained. A Monte Carlo simulation assesses the bias of the mean group coefficient of the long and the short run estimates. All three estimators perform well in terms bias. However standard errors are overestimated. An alternative approach is to bootstrap standard errors or confidence intervals. However the difficulties are three fold. First of all the dynamics of the model need to be preserved. Secondly the correlation of the errors across time and cross-sectional units should not change as well. Finally, the question arises how to treat the common factors. Are the common factors treated as determined or are they bootstrapped as well. All difficulties have been discussed on their own, but a combination of all has not been done and is left for further research.

6 Acknowledgments

I am grateful to all participants of the Stata User Group Meeting in Zürich in 2018, in particular Achim Ahrens and David Drukker for valuable comments and feedback. I am grateful for help and comments from Kamiar Mohaddes, Mark Schaffer, Gregorio Tullio and plenty of users of `xtdcce2`, who gave valuable feedback. All remaining errors are my own.

7 Appendix

7.1 The xtdcce2 command

7.1.1 Syntax

```
xtdcce2 depvar [indepvars] [varlist2 = varlist_iv] [if]  
    crosssectional(varlist_cr) [, nocrosssectional pooled(varlist_p)  
    cr_lags(#) ivreg2options(string) e.ivreg2 ivslow lr(varlist_lr)  
    lr_options(string) pooledconstant noconstant reportconstant  
    trend pooledtrend jackknife recursive noomitted nocd  
    fullsample showindividual fast blockdiaguse nodimcheck]
```

Data has to be [XT] `xtset` before using `xtdcce2`. `depvar`, `indepvars`, `varlist2`, `varlist_iv`, `varlist_cr`, `varlist_p` and `varlist_lr` may contain time-series operators, see [TS] `tsvarlist`, and factor variables, see [U] **11.4.3 Factor variables**. `xtdcce2` requires the `moremata` package by Jann (2005). `varlist2` are the endogenous variables and `varlist_iv` are the instruments.

7.1.2 Options

crosssectional(*varlist*) specifies cross sectional means, i.e. variables in \bar{z}_t .

These variables are partialled out. `cr_lags`(#) defines number of lags (p_T) of the cross sectional averages. The number of lags can be variable specific.

The same order as in `cr()` applies, hence if `cr(y x)`, then `cr_lags`($p_{\bar{y}}$ $p_{\bar{x}}$).

pooled(*varlist*) constraints coefficients to be homogeneous ($\beta_i = \beta, \forall i \in N$).

reportconstant reports constant and pooledconstant pools it.

pooled(*varlist*) specifies homogeneous coefficients. For these variables the estimated coefficients are constrained to be equal across all units ($\beta_i = \beta \forall i$). Variable may occur in `indepvars`. Variables in `exogenous_vars()`, `endogenous_vars()` and `lr()` may be pooled as well.

crosssectional(*varlist*) defines the variables which are included in z_t and added as cross sectional averages (\bar{z}_{t-l}) to the equation. Variables in `crosssectional()` may be included in `pooled()`, `exogenous_vars()`, `endogenous_vars()` and `lr()`. Variables in `crosssectional()` are partialled out, the coefficients not estimated and reported. `crosssectional(_all)` adds adds all variables as cross sectional averages. No cross sectional averages are added if `crosssectional(_none)` is used, which is equivalent to `nocrosssectional`. `crosssectional()` is a required option but can be substituted by `nocrosssectional`.

`cr_lags`(#) specifies the number of lags of the cross sectional averages. If not defined but `crosssectional()` contains *varlist*, then only contemporaneous cross sectional averages are added, but no lags. `cr_lags(0)` is equivalent to. The number of lags can be different for different variables, following the order defined in `cr()`.

nocrosssectional prevents adding cross sectional averages. Results will be equivalent to the Pesaran and Smith (1995) Mean Group estimator, or if **lr(varlist)** specified to the Shin et al. (1999) Pooled Mean Group estimator.

xtdcce2 supports instrumental variable regression using **ivreg2**. The IV specific options are:

- ivreg2options** passes further options on to **ivreg2**. See **ivreg2**, **options** for more information.
- fulliv** posts all available results from **ivreg2** in **e()** with prefix *ivreg2_*.
- noisily** shows the output of wrapped **ivreg2** regression command.
- ivslow** For the calculation of standard errors for pooled coefficients an auxiliary regressions is performed. In case of an IV regression, **xtdcce2** runs a simple IV regression for the auxiliary regressions. This is faster. If option is used **ivslow**, then **xtdcce2** calls **ivreg2** for the auxiliary regression. This is advisable as soon as **ivreg2** specific options are used.

xtdcce2 is able to estimate long run coefficients. Three models are supported, an error correction model, the CS-DL and CS-ARDL method. No options for the CS-DL method are necessary.

lr(varlist): Variables to be included in the long-run cointegration vector. The first variable(s) is/are the error-correction speed of adjustment term. The default is to use the pmg model. In this case each estimated coefficient is divided by the negative of the long-run cointegration vector (the first variable). If the option **ardl** is used, then the long run coefficients are estimated as the sum over the coefficients relating to a variable, divided by the sum of the coefficients of the dependent variable.

lr_options(string) Options for the long run coefficients. Options may be:

- ardl** estimates the CS-ARDL estimator.
- nodivide**, coefficients are not divided by the error correction speed of adjustment vector.
- xtpmgnames**, coefficients names in **e(b_p_mg)** and **e(V_p_mg)** match the name convention from **xtpmg**.

noconstant suppress constant term.

pooledconstant restricts the constant to be the same across all groups ($\beta_{0,i} = \beta_0, \forall i$).

reportconstant reports the constant. If not specified the constant is treated as a part of the cross sectional averages.

trend adds a linear unit specific trend. May not be combined with **pooledtrend**.

pooledtrend a linear common trend is added. May not be combined with **trend**.

jackknife applies the jackknife bias correction for small sample time series bias. May not be combined with **recursive**.

recursive applies recursive mean adjustment method to correct for small sample time series bias. May not be combined with **jackknife**.

nocd suppresses calculation of CD test statistic.

nomitted suppress checks for collinearity.

showindividual reports unit individual estimates in output.

fast omit calculation of unit specific standard errors.

`fullsample` uses entire sample available for calculation of cross sectional averages. Any observations which are lost due to lags will be included calculating the cross sectional averages (but are not included in the estimation itself). The option does not remove any if statements. This means, that `if` an if removes certain cross-sectional units from the estimation sample, `xtdcce2` will not use those (as specified by `if`), even if `fullsample` is used. `blockdiaguse` uses mata `blockdiag` rather than an alternative algorithm. `mata blockdiag` is slower, but might produce more stable results.

7.1.3 Stored Values

Scalars

<code>e(N)</code>	number of observations	<code>e(N_g)</code>	number of groups
<code>e(T)</code>	number of time periods	<code>e(K_mg)</code>	number of regressors
<code>e(N_partial)</code>	number of variables partialled out	<code>e(N_omitted)</code>	number of omitted variables
<code>e(N_pooled)</code>	number of pooled variables	<code>e(mss)</code>	model sum of square
<code>e(rss)</code>	residual sum of squares	<code>e(F)</code>	F statistic
<code>e(l1)</code>	log-likelihood (only IV)	<code>e(rmse)</code>	root mean squared error
<code>e(df_m)</code>	model degrees of freedom	<code>e(df_r)</code>	residual degree of freedom
<code>e(r2)</code>	R -squared	<code>e(r2_a)</code>	R -squared adjusted
<code>e(cd)</code>	CD test statistic	<code>e(cdp)</code>	p-value of CD test statistic
<code>e(cr_lags)</code>	number of lags of cross sectional averages		

Scalars

	(unbalanced panel)		
<code>e(Tmin)</code>	minimum time	<code>e(Tmax)</code>	maximum time
<code>e(Tbar)</code>	average time		

Macros

<code>e(tvar)</code>	name of time variable	<code>e(idvar)</code>	name of unit variable
<code>e(depvar)</code>	name of dependent variable	<code>e(indepvar)</code>	name of independent variables
<code>e(omitted)</code>	name of omitted variables	<code>e(lr)</code>	long run variables
<code>e(pooled)</code>	name of pooled variables	<code>e(cmd)</code>	command line
<code>e(cmdline)</code>	command line including options	<code>e(version)</code>	<code>xtdcce2</code> version, if <code>xtdcce2</code> , version used
<code>e(insts)</code>	instruments (exogenous) variables	<code>e(instd)</code>	instrumented (endogenous) variables

Matrices

<code>e(b)</code>	coefficient vector (mean group or individual)	<code>e(V)</code>	variance-covariance matrix (mean group or individual)
<code>e(bi)</code>	coefficient vector (individual and pooled)	<code>e(Vi)</code>	variance-covariance matrix (individual and pooled)

Functions

<code>e(sample)</code>	marks estimation sample
------------------------	-------------------------

7.2 ECM calculation

An general ARDL model is transformed to an ECM by adding and subtracting $y_{i,t-1}$ and cross products of the contemporaneous value of x and β coefficients

it is not associated with in the level equation. For an ARDL(1,0):

$$\begin{aligned}
y_{i,t} &= \alpha_i + \lambda_{1,i}y_{i,t-1} + \beta_{0,i}x_{i,t} + e_{i,t} \\
\Delta y_{i,t} &= \alpha_i + \lambda_{1,i}y_{i,t-1} - y_{i,t} + \beta_{0,i}x_{i,t} + e_{i,t} \\
&= (\lambda_{1,i} - 1) \left[y_{i,t-1} + \frac{\alpha_i}{\lambda_{1,i} - 1} + \frac{\beta_{0,i}}{\lambda_{1,i} - 1}x_{i,t} \right] + e_{i,t} \\
&= -(1 - \lambda_{1,i}) \left[y_{i,t-1} - \frac{\alpha_i}{1 - \lambda_{1,i}} - \frac{\beta_{0,i}}{1 - \lambda_{1,i}}x_{i,t} \right] + e_{i,t} \\
&= \phi_i [y_{i,t-1} - \theta_{0,i} - \theta_{1,i}x_{i,t}] + e_{i,t}
\end{aligned}$$

where

$$\phi_i = -(1 - \lambda_{1,i}), \quad \theta_{0,i} = \frac{\alpha_i}{1 - \lambda_{1,i}}, \quad \theta_{1,i} = \frac{\beta_{0,i}}{1 - \lambda_{1,i}}$$

For an ARDL(1,1):

$$\begin{aligned}
y_{i,t} &= \alpha_i + \lambda_{1,i}y_{i,t-1} + \beta_{0,i}x_{i,t} + \beta_{1,i}x_{i,t-1} + e_{i,t} \\
\Delta y_{i,t} &= \alpha_i + \lambda_{1,i}y_{i,t-1} - y_{i,t} + \beta_{0,i}x_{i,t} + \beta_{1,i}x_{i,t-1} + \beta_{1,i}x_{i,t} - \beta_{1,i}x_{i,t} + e_{i,t} \\
&= \alpha_i + (\lambda_{1,i} - 1)y_{i,t-1} + x_{i,t}(\beta_{0,i} + \beta_{1,i}) + \beta_{1,i}x_{i,t-1} - \beta_{1,i}x_{i,t} + e_{i,t} \\
&= (\lambda_{1,i} - 1) \left[y_{i,t-1} + \frac{\alpha_i}{\lambda_{1,i} - 1} + \frac{\beta_{0,i} + \beta_{1,i}}{\lambda_{1,i} - 1}x_{i,t} \right] + \beta_{1,i}(x_{i,t-1} - x_{i,t}) + e_{i,t} \\
&= \phi_i [y_{i,t-1} - \theta_{0,i} - \theta_{1,i}x_{i,t}] - \beta_{1,i}\Delta x_{i,t} + e_{i,t}
\end{aligned}$$

where

$$\phi_i = -(1 - \lambda_{1,i}), \quad \theta_{0,i} = \frac{\alpha_i}{1 - \lambda_{1,i}}, \quad \theta_{1,i} = \frac{\beta_{0,i} + \beta_{1,i}}{1 - \lambda_{1,i}}$$

For an ARDL(1,2):

$$\begin{aligned}
y_{i,t} &= \alpha_i + \lambda_{1,i}y_{i,t-1} + \beta_{0,i}x_{i,t} + \beta_{1,i}x_{i,t-1} + \beta_{2,i}x_{i,t-2} + e_{i,t} \\
\Delta y_{i,t} &= \alpha_i + \lambda_{1,i}y_{i,t-1} - y_{i,t} + \beta_{0,i}x_{i,t} + \beta_{1,i}x_{i,t-1} \\
&\quad + \beta_{1,i}x_{i,t} - \beta_{1,i}x_{i,t} + \beta_{2,i}x_{i,t} - \beta_{2,i}x_{i,t} + e_{i,t} \\
&= \alpha_i + (\lambda_{1,i} - 1)y_{i,t-1} + x_{i,t}(\beta_{0,i} + \beta_{1,i} + \beta_{2,i}) + \beta_{1,i}(x_{i,t-1} - x_{i,t}) \\
&\quad + \beta_{2,i}(x_{i,t-2} - x_{i,t}) + e_{i,t} \\
&= (\lambda_{1,i} - 1) \left[y_{i,t-1} + \frac{\alpha_i}{\lambda_{1,i} - 1} + \frac{\beta_{0,i} + \beta_{1,i} + \beta_{2,i}}{\lambda_{1,i} - 1}x_{i,t} \right] - \beta_{1,i}\Delta x_{i,t} - \beta_{2,i}\Delta_2 x_{i,t} + e_{i,t} \\
&= \phi_i [y_{i,t-1} - \theta_{0,i} - \theta_{1,i}x_{i,t}] - \beta_{1,i}\Delta x_{i,t} - \beta_{2,i}\Delta_2 x_{i,t} + e_{i,t}
\end{aligned}$$

where

$$\phi_i = -(1 - \lambda_{1,i}), \quad \theta_{0,i} = \frac{\alpha_i}{1 - \lambda_{1,i}}, \quad \theta_{1,i} = \frac{\beta_{0,i} + \beta_{1,i} + \beta_{2,i}}{1 - \lambda_{1,i}}$$

For an ARDL(2,1)

$$\begin{aligned}
y_{i,t} &= \alpha_i + \lambda_{1,i}y_{i,t-1} + \lambda_{2,i}y_{i,t-2} + \beta_{0,i}x_{i,t} + \beta_{1,i}x_{i,t-1} + e_{i,t} \\
\Delta y_{i,t} &= \alpha_i + (\lambda_{1,i} - 1)y_{i,t-1} + \lambda_{2,i}y_{i,t-2} + \beta_{0,i}x_{i,t} + \beta_{1,i}x_{i,t-1} \\
&\quad + \lambda_{2,i}y_{i,t-1} - \lambda_{2,i}y_{i,t-2} + \beta_{1,i}x_{i,t} - \beta_{1,i}x_{i,t-1} + e_{i,t} \\
&= (\lambda_{0,i} + \lambda_{1,i} - 1)y_{i,t-1} + x_{i,t}(\beta_{0,i} + \beta_{1,i})x_{i,t} \\
&\quad + \beta_{2,i}(x_{i,t-1} - x_{i,t}) + \lambda_{2,i}(y_{i,t-2} - y_{i,t-1}) + e_{i,t} \\
&= \phi_i [y_{i,t-1} - \theta_{0,i} - \theta_{1,i}x_{i,t}] - \beta_{1,i}\Delta x_{i,t} - \lambda_{2,i}\Delta y_{i,t-1} + e_{i,t}
\end{aligned}$$

where

$$\phi_i = -(1 - \lambda_{1,i} - \lambda_{2,i}), \quad \theta_{0,i} = \frac{\alpha_i}{1 - \lambda_{1,i} - \lambda_{2,i}}, \quad \theta_{1,i} = \frac{\beta_{0,i} + \beta_{1,i}}{1 - \lambda_{1,i} - \lambda_{2,i}}$$

7.3 Delta Method

7.3.1 ARDL Long Run Estimates

The delta method allows the calculation of an approximate probability distribution for a matrix function $a(\beta)$ based on a random vector with a known variance (see for example Hayashi, 2000, p. 93). Suppose that for the random vector $\beta_i \rightarrow_p \beta$ and $\sqrt{n}(\beta_i - \beta) \rightarrow_d N(0, \sigma)$. Denote the first derivatives of $a(\beta)$ as

$$A(\beta) \equiv \frac{\partial a(\beta)}{\partial \beta'}$$

Then the distribution of the function $a()$ is

$$\sqrt{n}[a(\beta_i) - a(\beta)] \rightarrow_d N(0, A(\beta)\Sigma A(\beta)')$$

The following ARDL(2,2) model, omitted the common factors without loss of generality, is estimated:

$$y_{i,t} = \alpha_i + \lambda_{1,i}y_{i,t-1} + \lambda_{2,i}y_{i,t-2} + \beta_{1,i}x_{i,t} + \beta_{2,i}x_{i,t-1} + e_{i,t}$$

The long run coefficients are

$$\begin{aligned}
\hat{\phi}_i &= -(1 - \hat{\lambda}_{1,i} - \hat{\lambda}_{2,i}) \\
\hat{\theta}_{1,i} &= \frac{\hat{\beta}_{1,i} + \hat{\beta}_{2,i}}{1 - \hat{\lambda}_{1,i} - \hat{\lambda}_{2,i}}
\end{aligned}$$

Stacking the short run coefficients into $\pi_i = (\lambda_{1,i}, \lambda_{2,i}, \beta_{1,i}, \beta_{2,i})$, then the vector function $a(\pi_i)$ maps the short run coefficients into a vector of the short run and long run coefficients $a(\pi_i) = (\lambda_{1,i}, \lambda_{2,i}, \beta_{1,i}, \beta_{2,i}, \phi_i, \theta_{1,i})$, where $\phi_i = 1 + \lambda_{1,i} +$

$\lambda_{2,i}$ and $\theta_{1,i} = \frac{\beta_{1,i} + \beta_{2,i}}{1 - \lambda_{1,i} - \lambda_{2,i}}$. The covariance matrix is

$$\Sigma_i = \begin{pmatrix} \text{Var}(\lambda_{1,i}) & \text{Cov}(\lambda_{1,i}, \lambda_{2,i}) & \text{Cov}(\lambda_{1,i}, \beta_{1,i}) & \text{Cov}(\lambda_{1,i}, \beta_{2,i}) \\ & \ddots & & \\ & & \ddots & \\ & & & \text{Var}(\beta_{2,i}) \end{pmatrix}$$

The first derivative of vector $a(\pi_i)$ is:

$$A(\pi_i) = \begin{pmatrix} \frac{\partial \lambda_{1,i}}{\partial \lambda_{1,i}} & \frac{\partial \lambda_{1,i}}{\partial \lambda_{2,i}} & \frac{\partial \lambda_{1,i}}{\partial \beta_{1,i}} & \frac{\partial \lambda_{1,i}}{\partial \beta_{2,i}} \\ \frac{\partial \lambda_{2,i}}{\partial \lambda_{1,i}} & \frac{\partial \lambda_{2,i}}{\partial \lambda_{2,i}} & \frac{\partial \lambda_{2,i}}{\partial \beta_{1,i}} & \frac{\partial \lambda_{2,i}}{\partial \beta_{2,i}} \\ \frac{\partial \beta_{1,i}}{\partial \lambda_{1,i}} & \frac{\partial \beta_{1,i}}{\partial \lambda_{2,i}} & \frac{\partial \beta_{1,i}}{\partial \beta_{1,i}} & \frac{\partial \beta_{1,i}}{\partial \beta_{2,i}} \\ \frac{\partial \beta_{2,i}}{\partial \lambda_{1,i}} & \frac{\partial \beta_{2,i}}{\partial \lambda_{2,i}} & \frac{\partial \beta_{2,i}}{\partial \beta_{1,i}} & \frac{\partial \beta_{2,i}}{\partial \beta_{2,i}} \\ \frac{\partial \phi_i}{\partial \lambda_{1,i}} & \frac{\partial \phi_i}{\partial \lambda_{2,i}} & \frac{\partial \phi_i}{\partial \beta_{1,i}} & \frac{\partial \phi_i}{\partial \beta_{2,i}} \\ \frac{\partial \theta_{1,i}}{\partial \lambda_{1,i}} & \frac{\partial \theta_{1,i}}{\partial \lambda_{2,i}} & \frac{\partial \theta_{1,i}}{\partial \beta_{1,i}} & \frac{\partial \theta_{1,i}}{\partial \beta_{2,i}} \end{pmatrix}$$

The derivatives are

$$\begin{aligned} \frac{\partial \phi_i}{\partial \lambda_{1,i}} &= \frac{\partial \phi_i}{\partial \lambda_{2,i}} = 1 \\ \frac{\partial \theta_{1,i}}{\partial \beta_{1,i}} &= \frac{\partial \theta_{1,i}}{\partial \beta_{2,i}} = \frac{1}{1 - \lambda_{1,i} - \lambda_{2,i}} \\ \frac{\partial \theta_{1,i}}{\partial \lambda_{1,i}} &= \frac{\partial \theta_{1,i}}{\partial \lambda_{2,i}} = \frac{\beta_{1,i} + \beta_{2,i}}{(1 - \lambda_{1,i} - \lambda_{2,i})^2} \end{aligned}$$

and thus

$$A(\pi_i) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ \frac{\beta_{1,i} + \beta_{2,i}}{(1 - \lambda_{1,i} - \lambda_{2,i})^2} & \frac{\beta_{1,i} + \beta_{2,i}}{(1 - \lambda_{1,i} - \lambda_{2,i})^2} & \frac{1}{1 - \lambda_{1,i} - \lambda_{2,i}} & \frac{1}{1 - \lambda_{1,i} - \lambda_{2,i}} \end{pmatrix}$$

Then the covariance matrix including the long run coefficients is

$$\Sigma_i^{lr} = A(\pi_i) \Sigma_i A(\pi_i)'$$

The same method as for ϕ_i is used for the calculation of variance for the long run pooled coefficients. Assume that $\theta_2 = \lambda_2 + \lambda_3$, where λ_2 and λ_3 are pooled coefficients. Then the row corresponding to θ_2 in $A(\cdot)$ includes a 1 in the columns corresponding to the coefficients the long run coefficient is estimated from.

7.4 Monte Carlo GDP

The underlying DGP is:

$$\begin{aligned} y_{i,t} &= c_{yi} + \lambda_{1,i}y_{i,t-1} + \lambda_{2,i}y_{i,t-2} + \beta_{0,i}x_{i,t} + \beta_{1,i}x_{i,t-1} + u_{i,t} \\ u_{i,t} &= \gamma'_i f_t + \epsilon_{i,t} \\ x_{i,t} &= c_{xi} + \alpha_{xi}y_{i,t-1} + \gamma_{xi}f_t + v_{xi,t} \end{aligned}$$

The cross-section specific fixed effects are generated as:

$$\begin{aligned} c_{yi} &\sim IIDN(1, 1) \\ c_{xi} &= c_{yi} + \varsigma_{c_x i}, \varsigma_{c_x i} \sim IIDN(0, 1). \end{aligned}$$

Dependence between $x_{i,t}$, $g_{i,t}$ and c_{yi} is introduced by adding c_{yi} to the equations for c_{xi} and c_{gi} . First the long run coefficient θ is drawn and then the short run coefficients are backed out.

$$\begin{aligned} \theta_i &\sim IIDN(1, \sigma_\theta^2) \\ \lambda_{1,i} &= (1 + \xi_{\lambda i})\eta_{\lambda i}, & \lambda_{2,i} &= -\xi_{\lambda i}\eta_{\lambda i} \\ \beta_{0,i} &= \xi_{\beta i}\eta_{\beta i}, & \beta_{1,i} &= (1 - \xi_{\beta i})\eta_{\beta i} \\ \eta_{\lambda i} &= IIDU(0, \lambda_{max}), & \eta_{\beta i} &= \theta_i (1 - \lambda_{i,1} - \lambda_{2,i}) \\ \xi_{\lambda i} &\sim IIDU(0.2, 0.3), & \xi_{\beta i} &\sim IIDU(0, 1) \end{aligned}$$

The common factors are calculated as:

$$\begin{aligned} \gamma_i &= \gamma + \eta_{i\gamma}, & \eta_{i\gamma} &\sim IIDN(0, \sigma_\gamma^2) \\ \gamma_{xi} &= \gamma_x + \eta_{i\gamma x}, & \eta_{i\gamma x} &\sim IIDN(0, \sigma_{\gamma x}^2) \\ \sigma_\gamma^2 &= \sigma_{\gamma x}^2 = 0.2^2 \\ \gamma &= \sqrt{b_\gamma}, & b_\gamma &= \frac{1}{m} - \sigma_\gamma^2 \end{aligned}$$

where m is the number of unobserved factors. In comparison to Chudik and Pesaran (2015) it is restricted to 1.

The error component of the variable x is generated as:

$$\begin{aligned} v_{xi,t} &= \rho_{xi}v_{xi,t-1} + \varsigma_{xi,t}, \varsigma_{xi,t} \sim IIDN(0, \sigma_{vxi}^2) \\ \rho_{xi} &\sim IIDU(0, 0.95) \\ \rho_f &= 0 \text{ if serially uncorrelated factors, or if correlated } \rho_f = 0.6 \\ \sigma_{vxi}^2 &= \sigma_{vi}^2 = \left(\beta_{0i} \sqrt{1 - [E(\rho_{xi})]^2} \right)^2 \end{aligned}$$

The errors are generated such that heteroskedasticity, autocorrelation and weakly

cross-sectional dependence is allowed.

$$\begin{aligned}
\epsilon_{i,t} &= \rho_{ei}\epsilon_{i,t-1} + \zeta_{i,t} \\
\zeta_t &= (\zeta_{1,t}, \zeta_{2,t}, \dots, \zeta_{N,t}) = \alpha_{CSD} S_\epsilon e_{\epsilon t} + e_{\epsilon t} \\
\Rightarrow \zeta_t &= (1 - \alpha_{CSD} S_\epsilon)^{-1} e_{\epsilon t} \\
e_{\epsilon t} &\sim IIDN(0, \frac{1}{2} \sigma_i^2 (1 - \rho_{ei}^2)), \text{ with } \sigma_i^2 \sim \chi^2(2) \\
\rho_{ei} &\sim IIDU(0, 0.8) \\
S_\epsilon &= \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & & 0 \\ 0 & \frac{1}{2} & 0 & \ddots & & \vdots \\ 0 & 0 & \ddots & \ddots & \frac{1}{2} & 0 \\ \vdots & & & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & \dots & 0 & 1 & 0 \end{pmatrix}
\end{aligned}$$

7.5 Monte Carlo Results

Table				
	$(\sigma_\theta^2, \lambda_{max})$	ρ_f	LR	SR
1	(0.2,0.6)	0	Table 2, 3	4
2	(0.2,0.6)	0.6	Table 5, 6	7
3	(0.8,0.8)	0	Table 8, 9	10
4	(0.8,0.8)	0.6	Table 11, 12	13

Table 1: Overview of Specifications. The DGP is $y_{i,t} = c_{yi} + \lambda_{1,i}y_{i,t-1} + \lambda_{2,i}y_{i,t-2} + \beta_{0,i}x_{i,t} + \beta_{1,i}x_{i,t-1} + \gamma'_i f_t + \epsilon_{i,t}$, where $\theta_i = (\beta_{0,i} + \beta_{1,i}) / (1 - \lambda_{1,i} - \lambda_{2,i}) \sim IIDN(1, \sigma_\theta^2)$. $c_{yi} \sim IIDN(0, 1)$ and $\gamma_i = \gamma + \eta_{i\gamma}$, $\eta_{i\gamma} \sim IIDN(0, \sigma_\gamma^2)$. For a further description see Section 7.4.

(N,T)	Bias of $\hat{\theta}_{MG}$ (x100)					RMSE of $\hat{\theta}_{MG}$ (x100)				
	40	50	100	150	200	40	50	100	150	200
CS-DL										
40	-21.57	-21.04	-19.52	-18.73	-18.26	23.50	22.48	20.10	19.04	18.46
50	-19.41	-19.15	-17.09	-16.64	-16.42	21.12	20.19	17.51	16.84	16.52
100	-20.04	-18.76	-17.40	-17.08	-16.93	20.39	19.02	17.25	16.81	16.61
150	-16.99	-16.41	-15.06	-14.72	-14.56	17.35	16.64	15.05	14.62	14.46
200	-20.73	-19.62	-18.20	-17.72	-17.37	21.04	19.80	18.24	17.70	17.31
CS-ARDL										
40	-2.63	-1.64	-1.94	-0.64	-0.48	192.31	13.65	8.01	5.58	4.80
50	-2.13	-186.07	-1.45	-0.75	-0.58	40.85	4049.97	6.53	5.47	4.36
100	-3.53	-0.43	-1.21	-0.94	-0.65	182.04	24.21	4.64	3.46	2.96
150	-4.93	-2.29	-1.31	-0.95	-0.59	34.46	7.20	3.69	2.69	2.48
200	-2.63	-2.29	-1.63	-1.11	-0.61	23.47	8.54	3.76	2.73	2.22
CS-ECM										
40	-2.63	-1.64	-1.94	-0.64	-0.48	192.31	13.65	8.01	5.58	4.80
50	-2.13	-186.07	-1.45	-0.75	-0.58	40.85	4049.97	6.53	5.47	4.36
100	-3.53	-0.43	-1.21	-0.94	-0.65	182.04	24.21	4.64	3.46	2.96
150	-4.93	-2.29	-1.31	-0.95	-0.59	34.46	7.20	3.69	2.69	2.48
200	-2.63	-2.29	-1.63	-1.11	-0.61	23.47	8.54	3.76	2.73	2.22

Table 2: Monte Carlo results for $\hat{\theta}_{MG} = 1/N \sum_{i=1}^N \hat{\theta}_i$ for Specification 1 with $p_T = [T^{1/3}]$, $\rho_f = 0$ and $(\sigma_\theta^2, \lambda_{max}) = (0.2, 0.6)$. See Table 1 and Section 7.4 for a description of the DGP.

(N,T)	Bias of $SE(\hat{\theta}_{MG})$ (x100)					RMSE of $SE(\hat{\theta}_{MG})$ (x100)				
	40	50	100	150	200	40	50	100	150	200
CS-DL										
40	77.60	78.13	79.62	80.41	80.87	12.06	13.54	15.85	16.68	17.19
50	79.26	79.52	81.55	81.99	82.20	11.40	12.63	14.87	15.66	16.13
100	77.65	78.89	80.21	80.52	80.67	12.91	13.75	15.26	15.79	16.07
150	81.66	82.23	83.56	83.89	84.05	14.17	14.81	16.01	16.39	16.60
200	78.67	79.76	81.18	81.65	82.00	14.77	15.40	16.51	16.88	17.09
CS-ARDL										
40	96.34	97.31	97.02	98.30	98.47	187.57	10.94	14.57	15.84	16.59
50	96.26	-84.65	96.93	97.62	97.79	36.00	4048.46	13.72	14.91	15.67
100	93.68	96.69	95.94	96.20	96.48	180.31	24.47	14.53	15.41	15.85
150	93.52	96.12	97.09	97.44	97.79	32.86	13.31	15.56	16.21	16.53
200	96.62	96.96	97.62	98.14	98.63	21.64	14.47	15.98	16.60	16.93
CS-ECM										
40	96.34	97.31	97.02	98.30	98.47	187.57	10.94	14.57	15.84	16.59
50	96.26	-84.65	96.93	97.62	97.79	36.00	4048.46	13.72	14.91	15.67
100	93.68	96.69	95.94	96.20	96.48	180.31	24.47	14.53	15.41	15.85
150	93.52	96.12	97.09	97.44	97.79	32.86	13.31	15.56	16.21	16.53
200	96.62	96.96	97.62	98.14	98.63	21.64	14.47	15.98	16.60	16.93

Table 3: Monte Carlo results for $SE(\hat{\theta}_{MG}) = \sqrt{1/N \sum_{i=1}^N (\hat{\theta}_i - \hat{\theta}_{MG})^2}$ for Specification 1 with $p_T = [T^{1/3}]$, $\rho_f = 0$ and $(\sigma_\theta^2, \lambda_{max}) = (0.2, 0.6)$. See Table 2 and Section 7.4 for a description of the DGP.

(N,T)	Bias of $\hat{\beta}_{0,MG}$ (x100)					Bias of $\hat{\beta}_{1,MG}$ (x100)				
	40	50	100	150	200	40	50	100	150	200
40	0.92	0.36	0.59	0.21	0.35	15.69	13.26	4.67	3.13	2.20
50	-1.63	0.20	-0.10	0.12	0.20	19.23	13.09	5.89	4.14	2.51
100	-1.01	-0.12	0.08	0.14	0.01	16.95	13.06	5.90	3.44	2.50
150	0.48	-0.76	-0.32	-0.04	0.04	16.59	13.39	6.04	3.66	2.73
200	-1.38	-0.92	-0.22	-0.01	0.09	17.89	13.07	5.95	3.78	2.73
Bias of $\hat{\phi}_{1,MG}$ (x100)										
40	-15.46	-10.69	-4.34	-2.22	-1.77	13.70	10.78	6.32	3.42	2.42
50	-15.59	-11.18	-4.43	-2.73	-1.83	12.25	10.31	5.12	3.75	2.50
100	-16.52	-11.33	-4.65	-2.87	-2.19	16.12	12.37	6.62	4.22	2.43
150	-17.08	-11.95	-5.10	-3.01	-2.30	21.70	18.34	7.88	5.94	3.59
200	-16.13	-11.35	-4.88	-3.03	-2.15	16.96	14.00	7.13	4.62	3.10

Table 4: Monte Carlo results for the mean group coefficients of the short run coefficients for Specification 1 with $p_T = [T^{1/3}]$, $\rho_f = 0$ and $(\sigma_\theta^2, \lambda_{max}) = (0.2, 0.6)$. See Table 1 and Section 7.4 for a description of the DGP.

(N,T)	Bias of $\hat{\theta}_{MG}$ (x100)					RMSE of $\hat{\theta}_{MG}$ (x100)				
	40	50	100	150	200	40	50	100	150	200
CS-DL										
40	-22.20	-21.55	-19.72	-18.85	-18.39	24.09	22.95	20.30	19.17	18.59
50	-20.00	-19.53	-17.33	-16.78	-16.53	21.68	20.62	17.75	16.98	16.63
100	-20.71	-19.27	-17.62	-17.22	-17.02	21.02	19.53	17.46	16.94	16.70
150	-17.57	-16.84	-15.24	-14.84	-14.65	17.92	17.06	15.23	14.74	14.55
200	-21.37	-20.11	-18.42	-17.85	-17.46	21.68	20.28	18.47	17.83	17.40
CS-ARDL										
40	-8.65	-1.64	-2.43	-1.01	-0.72	87.34	31.89	8.18	5.67	4.81
50	0.49	-1.36	-1.98	-1.08	-0.79	46.41	35.50	6.66	5.54	4.39
100	-23.34	-2.60	-1.69	-1.27	-0.87	394.98	9.34	4.77	3.54	3.01
150	-4.58	-3.30	-1.80	-1.25	-0.82	108.86	6.97	3.85	2.81	2.53
200	-10.41	-3.29	-2.16	-1.45	-0.85	74.29	9.23	4.03	2.88	2.29

Table 5: Monte Carlo results for $\hat{\theta}_{MG} = 1/N \sum_{i=1}^N \hat{\theta}_i$ for Specification 2 with $p_T = [T^{1/3}]$, $\rho_f = 0.6$ and $(\sigma_\theta^2, \lambda_{max}) = (0.2, 0.6)$. See Table 2 and Section 7.4 for a description of the DGP.

(N,T)	Bias of $SE(\hat{\theta}_{MG})$ (x100)					RMSE of $SE(\hat{\theta}_{MG})$ (x100)				
	40	50	100	150	200	40	50	100	150	200
CS-DL										
40	76.98	77.62	79.43	80.29	80.74	11.93	13.46	15.81	16.67	17.18
50	78.68	79.15	81.31	81.85	82.10	11.24	12.54	14.84	15.64	16.13
100	77.00	78.39	80.00	80.39	80.58	12.82	13.69	15.24	15.78	16.07
150	81.09	81.80	83.38	83.77	83.96	14.08	14.76	16.00	16.38	16.59
200	78.03	79.28	80.96	81.52	81.91	14.70	15.35	16.49	16.88	17.08
CS-ARDL										
40	90.38	97.32	96.54	97.94	98.23	83.61	29.57	14.58	15.85	16.59
50	98.84	97.02	96.40	97.29	97.58	42.69	33.93	13.73	14.91	15.67
100	74.44	94.59	95.46	95.87	96.26	393.67	11.59	14.52	15.41	15.85
150	93.86	95.13	96.61	97.15	97.56	107.24	13.16	15.56	16.21	16.53
200	88.90	95.98	97.09	97.80	98.40	71.66	14.18	15.97	16.60	16.93

Table 6: Monte Carlo results for $SE(\hat{\theta}_{MG}) = \sqrt{1/N \sum_{i=1}^N (\hat{\theta}_i - \hat{\theta}_{MG})^2}$ for Specification 2 with $p_T = [T^{1/3}]$, $\rho_f = 0.6$ and $(\sigma_\theta^2, \lambda_{max}) = (0.2, 0.6)$. See Table 2 and Section 7.4 for a description of the DGP.

(N,T)	40	50	100	150	200	40	50	100	150	200
	Bias of $\hat{\beta}_{0,MG}$ (x100)					Bias of $\hat{\beta}_{1,MG}$ (x100)				
40	0.61	0.34	0.57	0.17	0.32	18.56	15.73	5.92	3.89	2.80
50	-2.10	0.04	-0.20	0.12	0.19	22.65	15.93	7.16	4.96	3.11
100	-1.34	-0.33	0.03	0.14	0.00	19.86	15.39	7.04	4.16	3.04
150	0.28	-0.82	-0.38	-0.05	0.02	19.86	15.85	7.31	4.48	3.32
200	-1.61	-1.13	-0.24	-0.02	0.09	20.74	15.48	7.10	4.56	3.27
	Bias of $\hat{\phi}_{1,MG}$ (x100)					Bias of $\hat{\phi}_{2,MG}$ (x100)				
40	-19.22	-13.28	-5.48	-2.91	-2.28	13.89	11.14	6.45	3.43	2.39
50	-19.50	-13.96	-5.61	-3.48	-2.35	12.98	10.43	5.13	3.77	2.52
100	-20.58	-14.19	-5.84	-3.62	-2.72	17.22	12.94	6.69	4.31	2.48
150	-21.56	-15.11	-6.43	-3.86	-2.92	23.25	19.32	8.30	6.23	3.75
200	-20.21	-14.29	-6.08	-3.82	-2.69	17.90	14.60	7.35	4.70	3.14

Table 7: Monte Carlo results for the mean group coefficients of the short run coefficients for Specification 3 with $p_T = [T^{1/3}]$, $\rho_f = 0.6$ and $(\sigma_\theta^2, \lambda_{max}) = (0.2, 0.6)$. See Table 2 and Section 7.4 for a description of the DGP.

(N,T)	Bias of $\hat{\theta}_{MG}$ (x100)					RMSE of $\hat{\theta}_{MG}$ (x100)				
	40	50	100	150	200	40	50	100	150	200
CS-DL										
40	-30.10	-29.30	-27.24	-26.12	-25.54	31.22	29.94	27.08	25.63	24.95
50	-27.31	-26.84	-24.11	-23.63	-23.29	27.88	26.75	23.38	22.68	22.24
100	-26.59	-25.02	-23.28	-22.85	-22.52	24.60	23.14	21.03	20.47	20.14
150	-20.69	-20.03	-18.33	-17.94	-17.81	20.15	19.37	17.43	16.98	16.83
200	-28.43	-26.93	-25.01	-24.40	-23.92	28.05	26.48	24.49	23.79	23.28
CS-ARDL										
40	1.41	-0.06	-2.08	-0.69	-0.55	66.86	36.72	11.30	7.84	6.85
50	-23.32	-6.02	-2.00	-1.00	-0.84	340.46	143.68	9.82	8.47	6.67
100	-12.22	-8.48	-1.41	-1.16	-0.65	194.50	147.12	6.47	4.63	4.03
150	-6.73	0.25	-1.34	-1.05	-0.69	75.90	34.84	5.31	3.70	3.55
200	-37.15	-17.69	-2.16	-1.58	-0.87	545.06	380.23	5.41	3.98	3.17

Table 8: Monte Carlo results for $\hat{\theta}_{MG} = 1/N \sum_{i=1}^N \hat{\theta}_i$ for Specification 3 with $p_T = [T^{1/3}]$, $\rho_f = 0$ and $(\sigma_\theta^2, \lambda_{max}) = (0.8, 0.8)$. See Table 2 and Section 7.4 for a description of the DGP.

(N,T)	Bias of $SE(\hat{\theta}_{MG})$ (x100)					RMSE of $SE(\hat{\theta}_{MG})$ (x100)				
	40	50	100	150	200	40	50	100	150	200
CS-DL										
40	66.94	67.70	69.68	70.75	71.31	73.57	75.01	76.76	77.25	77.62
50	67.91	68.34	70.90	71.35	71.67	68.99	70.12	72.06	72.69	73.03
100	64.92	66.30	67.85	68.23	68.52	67.40	68.11	69.26	69.63	69.82
150	74.14	74.76	76.35	76.71	76.84	69.13	69.66	70.53	70.79	70.93
200	69.39	70.84	72.70	73.29	73.76	70.64	71.16	72.06	72.33	72.47
CS-ARDL										
40	97.12	95.71	93.77	95.10	95.23	79.39	71.01	71.97	72.99	73.55
50	71.64	87.80	91.55	92.49	92.63	338.88	148.27	67.81	68.95	69.67
100	77.62	80.94	87.19	87.41	87.86	195.75	156.44	67.04	67.89	68.23
150	87.19	93.72	92.23	92.51	92.84	86.13	71.57	69.08	69.74	69.99
200	60.93	79.79	94.85	95.41	96.10	541.60	381.86	70.37	71.01	71.30

Table 9: Monte Carlo results for $SE(\hat{\theta}_{MG}) = \sqrt{1/N \sum_{i=1}^N (\hat{\theta}_i - \hat{\theta}_{MG})^2}$ for Specification 3 with $p_T = [T^{1/3}]$, $\rho_f = 0$ and $(\sigma_\theta^2, \lambda_{max}) = (0.8, 0.8)$. See Table 2 and Section 7.4 for a description of the DGP.

(N,T)	Bias of $\hat{\beta}_{0,MG}$ (x100)					Bias of $\hat{\beta}_{1,MG}$ (x100)				
	40	50	100	150	200	40	50	100	150	200
40	0.84	0.46	0.72	0.29	0.43	16.84	14.75	4.66	3.11	2.01
50	-2.14	0.06	-0.10	0.20	0.29	18.58	12.34	5.32	3.78	2.13
100	-1.39	-0.25	0.10	0.24	0.06	14.00	11.01	4.97	2.73	1.98
150	0.54	-0.78	-0.34	-0.02	0.06	16.62	13.43	5.93	3.46	2.62
200	-1.62	-1.14	-0.24	-0.00	0.11	18.00	13.09	5.92	3.77	2.70
Bias of $\hat{\phi}_{1,MG}$ (x100)										
40	-14.47	-9.95	-3.93	-1.92	-1.48	8.22	7.03	4.14	2.30	1.69
50	-14.70	-10.48	-4.04	-2.45	-1.58	7.24	6.87	3.71	2.73	1.82
100	-15.61	-10.61	-4.29	-2.62	-1.96	10.90	8.87	4.92	3.36	1.75
150	-16.04	-11.18	-4.65	-2.75	-2.05	15.73	13.65	6.07	4.66	2.83
200	-15.30	-10.71	-4.55	-2.81	-1.97	11.33	9.67	5.29	3.44	2.35

Table 10: Monte Carlo results for the mean group coefficients of the short run coefficients for Specification 3 with $p_T = [T^{1/3}]$, $\rho_f = 0$ and $(\sigma_\theta^2, \lambda_{max}) = (0.8, 0.8)$. See Table 2 and Section 7.4 for a description of the DGP.

(N,T)	Bias of $\hat{\theta}_{MG}$ (x100)					RMSE of $\hat{\theta}_{MG}$ (x100)				
	40	50	100	150	200	40	50	100	150	200
CS-DL										
40	-30.98	-29.90	-27.48	-26.29	-25.67	32.06	30.48	27.31	25.80	25.09
50	-28.01	-27.34	-24.38	-23.77	-23.39	28.67	27.33	23.65	22.80	22.35
100	-27.27	-25.71	-23.53	-22.99	-22.62	25.21	23.77	21.26	20.60	20.23
150	-21.32	-20.50	-18.49	-18.07	-17.92	20.75	19.81	17.59	17.10	16.94
200	-29.22	-27.52	-25.28	-24.55	-24.03	28.84	27.08	24.75	23.93	23.40
CS-ARDL										
40	-5.01	-2.56	-2.62	-1.09	-0.82	64.97	22.93	11.44	7.94	6.86
50	-6.79	-6.08	-2.65	-1.35	-1.00	81.88	28.84	9.90	8.47	6.67
100	14.20	-3.47	-1.90	-1.49	-0.86	381.80	22.48	6.51	4.68	4.07
150	-0.99	-2.16	-1.88	-1.31	-0.90	64.07	43.98	5.28	3.77	3.57
200	-5.86	-5.53	-2.72	-1.92	-1.14	56.87	71.13	5.67	4.10	3.23

Table 11: Monte Carlo results for $\hat{\theta}_{MG} = 1/N \sum_{i=1}^N \hat{\theta}_i$ for Specification 4 with $p_T = [T^{1/3}]$, $\rho_f = 0.6$ and $(\sigma_\theta^2, \lambda_{max}) = (0.8, 0.8)$. See Table 2 and Section 7.4 for a description of the DGP.

(N,T)	Bias of $SE(\hat{\theta}_{MG})$ (x100)					RMSE of $SE(\hat{\theta}_{MG})$ (x100)				
	40	50	100	150	200	40	50	100	150	200
CS-DL										
40	66.10	67.13	69.45	70.59	71.18	73.35	74.93	76.74	77.25	77.63
50	67.26	67.88	70.65	71.22	71.57	68.78	69.97	72.00	72.66	73.02
100	64.32	65.70	67.62	68.10	68.43	67.31	68.04	69.24	69.62	69.82
150	73.55	74.32	76.20	76.60	76.74	69.04	69.60	70.52	70.78	70.92
200	68.61	70.26	72.44	73.15	73.64	70.57	71.11	72.04	72.32	72.47
CS-ARDL										
40	90.96	93.31	93.25	94.71	94.98	76.81	66.77	72.04	73.04	73.57
50	87.08	87.74	90.94	92.16	92.48	85.22	62.87	67.91	69.00	69.68
100	100.99	85.37	86.76	87.12	87.67	381.30	63.53	67.07	67.91	68.24
150	92.56	91.46	91.73	92.26	92.64	79.58	74.77	69.13	69.76	70.00
200	91.27	91.58	94.30	95.08	95.84	73.61	90.47	70.41	71.03	71.31

Table 12: Monte Carlo results for $SE(\hat{\theta}_{MG}) = \sqrt{1/N \sum_{i=1}^N (\hat{\theta}_i - \hat{\theta}_{MG})^2}$ for Specification 4 with $p_T = [T^{1/3}]$, $\rho_f = 0.6$ and $(\sigma_\theta^2, \lambda_{max}) = (0.8, 0.8)$. See Table 2 and Section 7.4 for a description of the DGP.

(N,T)	40	50	100	150	200	40	50	100	150	200
	Bias of $\hat{\beta}_{0,MG}$ (x100)					Bias of $\hat{\beta}_{1,MG}$ (x100)				
40	0.60	0.32	0.68	0.27	0.39	19.75	17.44	5.93	3.89	2.69
50	-2.67	-0.17	-0.23	0.19	0.29	21.56	14.87	6.46	4.52	2.68
100	-1.66	-0.56	0.03	0.25	0.04	16.31	12.88	5.90	3.33	2.42
150	0.32	-0.84	-0.39	-0.04	0.04	19.89	15.89	7.16	4.26	3.21
200	-1.84	-1.34	-0.26	-0.01	0.11	20.75	15.36	7.03	4.52	3.20
	Bias of $\hat{\phi}_{1,MG}$ (x100)					Bias of $\hat{\phi}_{2,MG}$ (x100)				
40	-17.85	-12.33	-4.95	-2.53	-1.93	7.21	6.35	3.87	2.09	1.48
50	-18.22	-12.99	-5.08	-3.09	-2.03	6.50	5.98	3.26	2.49	1.61
100	-19.30	-13.18	-5.35	-3.28	-2.42	10.59	8.53	4.60	3.19	1.63
150	-20.07	-14.02	-5.82	-3.49	-2.59	15.92	13.55	6.00	4.65	2.78
200	-18.98	-13.35	-5.62	-3.49	-2.44	10.80	9.11	5.03	3.26	2.19

Table 13: Monte Carlo results for the mean group coefficients of the short run coefficients for Specification 2 with $p_T = [T^{1/3}]$, $\rho_f = 0.6$ and $(\sigma_{\theta}^2, \lambda_{max}) = (0.8, 0.8)$. See Table 2 and Section 7.4 for a description of the DGP.

(N,T)	Bias of $\hat{\theta}_{MG}$ (x100)					RMSE of $\hat{\theta}_{MG}$ (x100)				
	40	50	100	150	200	40	50	100	150	200
CS-DL										
40	-7.67	-7.28	-6.95	-6.28	-6.06	14.54	12.26	9.47	7.96	7.38
50	-7.18	-7.52	-6.33	-6.13	-5.95	13.26	11.69	8.47	7.57	7.04
100	-6.86	-5.91	-5.65	-5.56	-5.58	10.17	8.84	6.73	6.17	6.05
150	-5.51	-5.49	-4.86	-4.73	-4.65	7.97	7.23	5.64	5.22	5.07
200	-7.76	-7.06	-6.54	-6.36	-6.07	9.36	8.27	7.17	6.70	6.29
CS-ARDL										
40	-0.27	1.92	-0.74	-0.19	-0.29	191.07	20.76	9.12	6.19	5.40
50	7.66	2.26	-0.49	-0.53	-0.46	102.70	36.44	7.39	6.17	4.90
100	7.11	2.93	-0.43	-0.42	-0.52	153.49	33.15	5.16	3.80	3.35
150	14.19	-1.35	-0.63	-0.54	-0.30	123.62	34.92	4.13	3.08	2.63
200	4.84	0.88	-0.65	-0.53	-0.22	107.79	50.73	4.08	2.94	2.42

Table 14: Monte Carlo results for $\hat{\theta}_{MG} = 1/N \sum_{i=1}^N \hat{\theta}_i$ for Specification 1 with $p_T = [T^{1/3}]$, $\rho_f = 0$ and $(\sigma_{\theta}^2, \lambda_{max}) = (0.2, 0.6)$, but with misspecified command line. See Table 1 and Section 7.4 for a description of the DGP.

References

- Blackburne, Edward F. and Mark W. Frank. 2007. “Estimation of nonstationary heterogeneous panels.” *The Stata Journal* 7(2):197–208.
- Chudik, Alexander, Kamiar Mohaddes, M. Hashem Pesaran and Mehdi Raissi. 2013. “Debt, Inflation and Growth: Robust Estimation of Long-Run Effects in Dynamic Panel Data Models.” *Cambridge Working Papers in Economics Debt*.
- Chudik, Alexander, Kamiar Mohaddes, M. Hashem Pesaran and Mehdi Raissi. 2016. Long-Run Effects in Large Heterogeneous Panel Data Models with Cross-Sectionally Correlated Errors. In *Essays in Honor of Aman Ullah (Advances in Econometrics, Volume 36)*, ed. R. Carter Hill, Gloria González-Rivera and Tae-Hwy Lee. Emerald Group Publishing Limited pp. 85–135.
- Chudik, Alexander and M. Hashem Pesaran. 2015. “Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors.” *Journal of Econometrics* 188(2):393–420.
- Chudik, Alexander, M. Hashem Pesaran and Elisa Tosetti. 2011. “Weak and strong cross-section dependence and estimation of large panels.” *The Econometrics Journal* 14(1):C45–C90.
- Ditzen, Jan. 2018. “Estimating dynamic common-correlated effects in Stata.” *The Stata Journal* 18(3):585 – 617.
- Eberhardt, Markus. 2012. “Estimating panel time series models with heterogeneous slopes.” *The Stata Journal* 12(1):61–71.
- Engle, Robert F. and C. W. J. Granger. 1987. “Co-Integration and Error Correction: Representation, Estimation, and Testing.” *Econometrica* 55(2):251.
- Everaert, Gerdie and Tom De Groote. 2016. “Common Correlated Effects Estimation of Dynamic Panels with Cross-Sectional Dependence.” *Econometric Reviews* 35(3):1–31.
- Hayashi, Fumio. 2000. *Econometrics*. Princeton: Princeton University Press.
- Jann, Ben. 2005. “moremata: Stata module (Mata) to provide various functions.”
URL: <https://ideas.repec.org/c/boc/bocode/s455001.html>
- Kripfganz, Sebastian and Daniel Schneider. 2018. “ardl: Estimating autoregressive distributed lag and equilibrium correction models.”
URL: <https://ideas.repec.org/p/boc/usug18/09.html>
- Lee, Kevin, M. Hashem Pesaran and Ron Smith. 1997. “Growth and Convergence in a Multi-Country Empirical Stochastic Solow Model.” *Journal of Applied Economics* 12(4):357–392.

- Pesaran, M. Hashem. 2006. "Estimation and inference in large heterogeneous panels with a multifactor error structure." *Econometrica* 74(4):967–1012.
- Pesaran, M. Hashem and Ron Smith. 1995. "Estimating long-run relationships from dynamic heterogeneous panels." *Journal of Econometrics* 68(1):79–113.
- Shin, Yongcheol, M. Hashem Pesaran and Ron P. Smith. 1999. "Pooled Mean Group Estimation of Dynamic Heterogeneous Panels." *Journal of the American Statistical Association* 94(446):621–634.