Yale University
EliScholar – A Digital Platform for Scholarly Publishing at Yale

Cowles Foundation Discussion Papers | Cowles Foundation

12-1-2018

Dynamic Panel Modeling of Climate Change

Peter C.B. Phillips

Follow this and additional works at: https://elischolar.library.yale.edu/cowles-discussion-paper-series

 Part of the Economics Commons

Recommended Citation

Phillips, Peter C.B., "Dynamic Panel Modeling of Climate Change" (2018). *Cowles Foundation Discussion Papers*. 115.
https://elischolar.library.yale.edu/cowles-discussion-paper-series/115

This Discussion Paper is brought to you for free and open access by the Cowles Foundation at EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Cowles Foundation Discussion Papers by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.

DYNAMIC PANEL MODELING OF CLIMATE CHANGE

By

Peter C. B. Phillips

December 2018

COWLES FOUNDATION DISCUSSION PAPER NO. 2150



COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Dynamic Panel Modeling of Climate Change*

Peter C. B. Phillips
Yale University, University of Auckland,
University of Southampton and Singapore Management University

December 11, 2018

Abstract

We discuss some conceptual and practical issues that arise from the presence of global energy balance effects on station level adjustment mechanisms in dynamic panel regressions with climate data. The paper provides asymptotic analyses, observational data computations, and Monte Carlo simulations to assess the use of various estimation methodologies, including standard dynamic panel regression and cointegration techniques that have been used in earlier research. The findings reveal massive bias in system GMM estimation of the dynamic panel regression parameters, which arise from fixed effect heterogeneity across individual station level observations. Difference GMM and Within Group (WG) estimation have little bias and WG estimation is recommended for practical implementation of dynamic panel regression with highly disaggregated climate data. Intriguingly from an econometric perspective and importantly for global policy analysis, it is shown that despite the substantial differences between the estimates of the regression model parameters, estimates of global transient climate sensitivity (of temperature to a doubling of atmospheric CO_2) are robust to the estimation method employed and to the specific nature of the trending mechanism in global temperature, radiation, and CO_2.

Keywords: Climate modeling, Cointegration, Difference GMM, Dynamic panel, Spatio-temporal modeling, System GMM, Transient climate sensitivity, Within group estimation.

JEL Classification: C32, C33

1 Introduction

A natural and near universal condition in modeling climate is the use of an energy balance relationship that links average global temperature to average

*Support from a Kelly Fellowship at the University of Auckland is acknowledged. Special thanks to Chirok Han and Donggyu Sul for discussions and for helping with the panel regression Gauss code.

global downwellling radiation and greenhouse gas influences. This balance suggests the existence of a long run cointegrating econometric relation among these variables, a relation that is now supported by considerable empirical evidence (Storelvmo et al, 2016; Kaufmann et al, 2011, 2013). While such global balancing relations are of considerable interest in themselves, they are also useful in the specification of more detailed models that relate to station level behavior and adjustments that must necessarily take global influences into account. Panel models of this type have been used recently in climate studies by Magnus et al. (2011) and Storelvmo et al. (2016). These studies help to assess, inter alia, the impact that atmospheric aerosols have on measurements of greenhouse gas (GHG) effects on global warming and thereby the measurement of transient climate sensitivity (TCS) to CO_2, which is arguably the 'holy grail' of modern climate science. These econometric models are now also being employed as a window through which global climate models can be calibrated against observational data (Phillips et al, 2018).

The present contribution raises some conceptual issues and provides analyses that are useful in understanding the manner in which the Earth's mechanism of global energy balance (or imbalance) affects the dynamic mechanism of local station level adjustments in temperature. As shown in Phillips et al (2018) and discussed below, station level dynamic adjustments that are impacted by the time path of the equilibrium energy balance can, under the seemingly natural condition of a stationary error correction formulation, imply a further long run cointegrating relationship between average global temperature and radiation. That relation in turn implies a long run relationship between downwelling radiation and CO_2.

A second objective of this note is to report simulations that compare the use of standard panel econometric methods for estimating dynamic panel regressions with disaggregated station level data. The methods examined are Within Group (WG) least squares, difference GMM (diff-GMM; Arellano and Bond, 1993), and system GMM (sys-GMM; Blundell and Bond, 1998). The simulation design is based on the empirical model used in Magnus et al (2011) and Storelvmo et al (2016) with observational data on both CO_2 and downwelling radiation employed in the data generating mechanism and with sample sizes that correspondingly match the observed data.

The simulation findings show substantial bias in system GMM estimation, particularly in the panel autoregressive coefficient estimates which are biased upwards almost sixfold and thereby provide a hugely distorted picture of station level temperature dynamics and the manner in which these are impacted by trends in global averages in radiation and CO_2. These biases correspond closely to the empirical differences between the estimates using the data of Storelvmo et al (2016) and Phillips et al (2018). They are also predicted by earlier simulations and by stationary panel asymptotic theory (Bun and Weijmeyer, 2010; Hayakawa, 2007, 2015;), which show how system GMM limit theory is affected by the magnitude of the ratio of the variance of individual station level fixed effects to the equation error variance. Global climate data naturally display substantial heterogeneity across station location, so that fixed effect heterogeneity is

a prominent characteristic in modeling this data. As a result, sys-GMM estimation is deemed unreliable in parametric dynamic panel regressions with climate data of this highly disaggregated type. For the cross section and time series sample sizes that are presently available, WG and diff-GMM methods both perform well although diff-GMM manifests some bias and has greater variance than WG estimation. The findings therefore indicate a preference for WG estimation of dynamic panels with substantially disaggregated climate data. The present paper gives a complete asymptotic theory for WG estimation of such models in the presence of potentially cointegrated nonstationary climate data. This limit theory enables inference about individual parameters in the panel regression model and assists in forecasting exercises.

A third objective of this paper is to investigate the estimation of TCS. This parameter measures the effect on temperature of a doubling of atmospheric CO_2 levels from pre-industrial time levels. It is therefore a global parameter that is expressed as a function of both dynamic adjustment parameters in the panel regression and the parameters of the global energy balance relationship. Estimation of TCS may be conducted based on full system estimation of the dynamic panel model. Despite the substantial differences between WG, diff-GMM and sys-GMM estimates of the regression model parameters, estimates of global TCS are shown to be identical, and therefore completely robust to the estimation method employed as well as the specific nature of the trending mechanism that is present in the key variables of the system: global temperature, radiation, and CO_2. This robustness extends to the asymptotic theory of the TCS estimates and therefore provides some measure of assurance of reliability concerning both the TCS estimate and its associated asymptotic confidence intervals for this important parameter. This reassurance is important to policy makers in the consideration of GHG abatement measures designed to control the effects of anthropogenic-driven climate forcing.

A second method of estimation of TCS is to conduct a simple single equation cointegrating regression to capture the long-run impact of atmospheric CO_2 levels on global temperature. This procedure was explored in Phillips et al. (2018) and shown to allow for energy imbalance, so that sustained rises in atmospheric CO_2 may impact station level temperature while continuing to influence rising global temperature, a situation that approximates prevailing climate conditions and accords with earlier empirical studies with aggregate data (Kaufmann et al, 2011, 2013). The cointegration approach allows for the use of standard methods of estimation, such as fully-modified least squares and dynamic ordinary least squares, accounts for the presence of both deterministic and stochastic trends in the global variables as well as the cointegrating link, and is convenient to apply in practical work.

The paper is organized as follows. The dynamic panel model and assumptions on its various components are given in Section 2. Section 3 shows invariance of the estimate of the TCS parameter to the specific method employed in estimation of the panel regression. Asymptotic theory for the panel regression coefficient estimates and the TCS parameter are given in Section 4. Simulations are reported in Section 5 and Section 6 concludes. Proofs are given in the

Appendix B and additional figures in Appendix A.

2 Model and Assumptions

Throughout the paper we use the following dynamic panel model from Magnus et al. (2011) and Storelvmo et al (2016), which relates station-level temperature (T_{it+1}) at time $t+1$ to local temperature (T_{it}), local downwelling surface radiation (R_{it}), and global factors (λ_t), all at time t. The base model has the following two equations

$$T_{i,t+1} = \alpha_i + \beta_1 T_{i,t} + \beta_2 R_{i,t} + \lambda_t + u_{it+1}, \quad i = 1, ..., N \text{ and } t = 1, ..., n, \quad (1)$$

where the α_i are station-level effects, β_1 and β_2 are parameters, and u_{it+1} is a disturbance. The time specific quantity λ_t in (1) is specified by the equation

$$\lambda_t = \gamma_0 + \gamma_1 T_t + \gamma_2 R_t + \gamma_3 \ln(CO_{2,t}), \quad (2)$$

which relates the spatial aggregates $(T_t, R_t) = \left(N^{-1}\sum_{i=1}^N T_{it}, N^{-1}\sum_{i=1}^N R_{it}\right)$ and the logarithm of the CO_2 equivalent series, $\ln(CO_{2,t})$. Phillips et al (2018) added the following mechanisms for the generation of local radiation effects R_{it} and global CO_2

$$\begin{aligned} R_{it} &= R_{it}^0 + \delta_{ri}' G_t + P_{it}, \; P_{it} = P_{i0} + U_{it}^P, & (3)\\ \ln(CO_{2,t}) &= \delta_{c0} + \delta_{c1} t + \delta_c' G_t + u_{ct}, & (4)\\ G_t &= \sum_{s=1}^t u_{gs}, \; U_{it}^P = \sum_{k=1}^t u_{ik}^P, & (5) \end{aligned}$$

which provide for both global ($\delta_{ri}' G_t$) and local (P_{it}) stochastic trend determinants of R_{it} and a deterministic drift ($\delta_{c1} t$) in conjunction with global stochastic trend components ($\delta_c' G_t$) as the primary drivers of the logarithm of global CO_2.

Equation (2) may be interpreted as a form of energy balance relationship that captures the global linkage between temperature, radiation and greenhouse gas atmospheric influences, allowing for the presence of stochastic and deterministic trend effects. The balance in these global elements is measured by λ_t and is assumed to be one of the drivers impacting local temperature in the subsequent time period. The dynamic panel regression equation (1) therefore characterizes the dynamic adjustment mechanism of station level temperature T_{it+1} as an autoregression on past temperature T_{it}, radiation R_{it}, and global energy balancing effects λ_t. Equation (2) is specified without error, so that the observed aggregate variables $(T_t, R_t, \ln(CO_{2,t}))$ are assumed to impact station-level temperature in (1) directly without noise. The possibility of including unobserved noise in the specification of λ_t and the impact on the asymptotic theory of this inclusion of measurement error in (2) is considered later in the paper. Readers are referred to Phillips et al (2018) for detailed discussion of the specification of (1)-(5) and their justification in terms of relevant atmospheric considerations and empirical

assessments using observed data. The global variables are shown in Figure 1 over the time period 1964-2005.

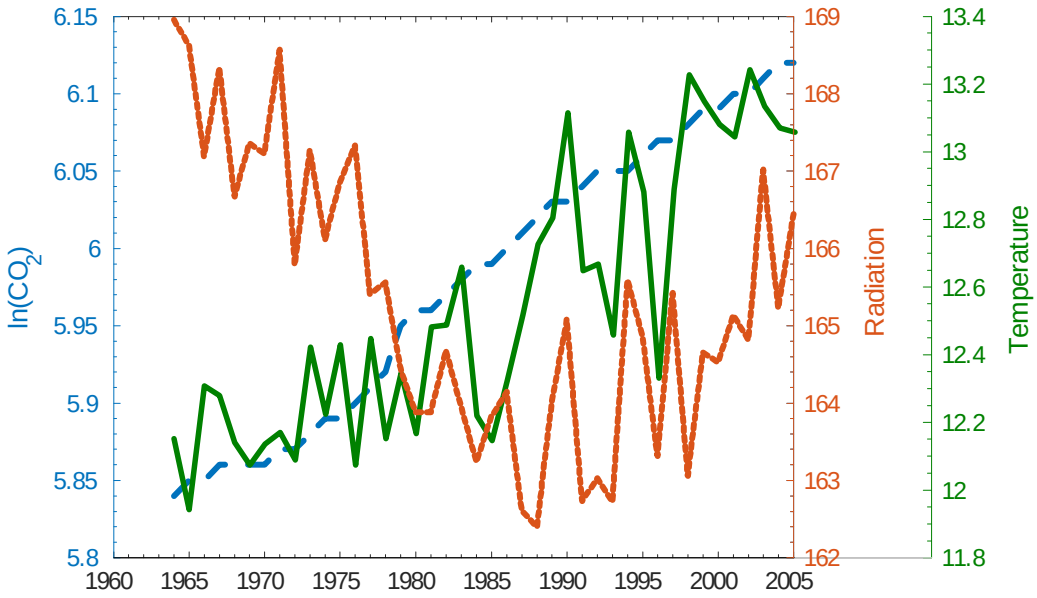


Figure 1 (Phillips et al., 2018): Global temperature ($\bar{T}_t$, green, solid), downwelling radiation ($\overline{R}_t$, orange, dotted), and CO_2 equivalent ($\ln(CO_{2t})$, blue, dashed) over 1964-2005.

The following assumptions about the various components of this system are based on those used in Phillips et al (2018) to which the reader is referred for detailed discussion. Some implications of the conditions on convergence rates and asymptotic bias and efficiency are discussed later in the paper.

Assumption A

(i) *The panel regression errors* $\{u_{it}\} \sim_{iid} \left(0, \sigma_u^2\right)$ *over* i *and* t *and are independent of the random sequences* $\left\{u_{it}^P\right\}, \{\delta_{ri}\}, \{u_{ct}\}$ *for all* (i,t). *The idiosyncratic loading factors* $\{\delta_{ri}\} \sim_{iid} (\delta_r, \Sigma_r)$ *and station-level effects* $\alpha_i \sim_{iid} \left(\alpha, \sigma_\alpha^2\right)$ *are independent and both are independent of* $\left\{u_{it}^P\right\}, \{u_{ct}\}$ *for all* (i,t), *where the* $\left\{u_{it}^P\right\}$ *are defined in A(iii) and the* $\{u_{ct}\}$ *in A(iv).*

(ii) $\bar{R}_t^0 = N^{-1} \sum_{i=1}^N R_{it}^0 \to_{a.s.} R^0 = \lim_{N\to\infty} \left\{N^{-1} \sum_{i=1}^N \mathbb{E}\left(R_{it}^0\right)\right\}$.

(iii) $P_{it} = P_{i0} + \sum_{k=1}^t u_{ik}^P =: P_{i0} + U_{it}^P$ *where* $u_{ik}^P \sim_{iid} \left(0, \sigma_p^2\right)$ *with finite fourth*

moments over i and t,

$$\bar{P}_0 = N^{-1}\sum_{i=1}^{N} P_{i0} \to_{a.s.} P^0 = \lim_{N\to\infty}\left\{N^{-1}\sum_{i=1}^{N}\mathbb{E}\left(P_{it}^0\right)\right\},$$

$N^{-1/2}\sum_{i=1}^{N}\left(P_{i0} - \mathbb{E}\left(P_{it}^0\right)\right) = O_p(1)$ and the partial sums U_{it}^P satisfy the invariance principle $n^{-1/2}U_{it}^P \Rightarrow U_i^P(r) \equiv BM\left(\sigma_P^2\right)$ for all i.

(iv) *$u_{gt} \sim_{iid}(0,\Sigma_g)$ with finite fourth moments has partial sums $U_{gt} = \sum_{k=1}^{t} u_{gk}$ that satisfy the invariance principle $n^{-1/2}U_{g\lfloor nr\rfloor} \Rightarrow U_g(r) \equiv BM(\Sigma_g)$, vector Brownian motion with covariance matrix $\Sigma_g > 0$;* and *$u_{ct} \sim_{iid}\left(0,\sigma_c^2\right)$ with finite fourth moments has partial sums $U_{ct} = \sum_{k=1}^{t} u_{ck}$ which satisfy the invariance principle $n^{-1/2}U_{c\lfloor nr\rfloor} \Rightarrow U_c(r) \equiv BM\left(\sigma_c^2\right)$, with $\sigma_c^2 > 0$.*

(v) *$|\beta_1| < 1$, $|\beta_1 + \gamma_1| < 1$.*

(vi) *$(n, N) \to \infty$ with $\frac{n}{N} + \frac{N}{n^3} \to 0$.*

An important feature of the model (1) and (2) is that it can be used to measure transient climate sensitivity (TCS) to CO_2 emissions. This parameter plays a major role in discussions about the potential impact of greenhouse gas emissions on Earth's climate. TCS is defined as the expected global temperature after a doubling of CO_2 and has the following analytic form (Magnus et al., 2011; and Storelvmo et al., 2016)

$$TCS = \frac{\gamma_3}{1-\beta_1-\gamma_1} \times \ln(2). \tag{6}$$

Phillips et al (2018) developed a simple and direct cointegration regression approach to the estimation of the parameter TCS using the long run relationship among the variables $(T_t, R_t, \ln(CO_{2,t}))$ that is implied by (1) and (2). A different, station-level approach is to estimate the parameters of the dynamic panel regression model (1) combined with the parameters that appear in the aggregate balancing relation (2) and to use these estimates in conjunction with formula (6) to obtain an estimate of TCS and an associated confidence interval.

The present contribution is concerned primarily with studying this station-level approach to estimation. As expected, the limit theory of TCS estimates obtained in this way from estimates of the complete model differ from those obtained by fitting the long-run relationship alone. Full panel regression estimation of the system (1) and (2) can be performed in various ways, for instance, by WG, diff-GMM, and sys-GMM techniques, with many additional variations depending on the precise selection of instrumental variables in the use of diff-GMM and sys-GMM techniques. Intriguingly, the resulting estimates of (6) obtained in this way turn out to be invariant to the method employed in the panel regression estimation of (1) and (2). This invariance holds even though the individual parameter estimates of $(\beta_1, \beta_2, \gamma_1, \gamma_2)$ obtained by WG, diff-GMM,

and sys-GMM differ. In some cases, particularly sys-GMM, the differences are huge – see Table 1 below and the attendant discussion. These differences arise primarily because of the substantial heterogeneity in the fixed effects α_i in the climate panel regression equation (1) which capture the large local variation in station temperature levels.

3 Estimation by Dynamic Panel Regression

3.1 Common Trends and Global Cointegration

The system (1) and (2) involves the station-level panel adjustment mechanism (1) with global effects imparted by the time specific effects λ_t, which in turn depend on global averages over stations. To reconcile these two components, aggregation of (1) gives

$$T_{t+1} = \bar{\alpha} + \beta_1 T_t + \beta_2 R_t + \lambda_t + \frac{1}{N}\sum_{i=1}^{N} u_{it+1} = \bar{\alpha} + \beta_1 T_t + \beta_2 R_t + \lambda_t + \bar{u}_{\cdot t+1},$$

where $\bar{\alpha} = N^{-1}\sum_{i=1}^{N}\alpha_i$ and $\bar{u}_{\cdot t+1} = N^{-1}\sum_{i=1}^{N} u_{it+1}$. Following standard practice for identification purposes in the presence of fixed individual and time effects, we set $\bar{\alpha} = 0$. Substituting (2) gives the global equation

$$T_{t+1} = \gamma_0 + (\beta_1 + \gamma_1) T_t + (\beta_2 + \gamma_2) R_t + \gamma_3 \ln(CO_{2,t}) + \bar{u}_{\cdot t+1}. \tag{7}$$

Setting $\theta_1 = \beta_1 + \gamma_1$ and $\theta_2 = \beta_2 + \gamma_2$, it is convenient to write (7) as

$$T_{t+1} = \gamma_0 + \theta_1 T_t + \theta_2 R_t + \gamma_3 \ln(CO_{2,t}) + \bar{u}_{\cdot t+1}, \tag{8}$$

and solving by back substitution gives the stochastic trend representation of T_t, in conjunction that of with $(R_t, \ln(CO_{2,t}))$, which is given in Phillips et al. (2018, theorem 1), viz.,

$$W_t = \begin{bmatrix} T_t \\ R_t \\ \ln(CO_{2,t}) \end{bmatrix} = \begin{bmatrix} \delta_{T0} + \delta_{T1} t + \delta_T' U_{gt} + u_{Tt}^{+} \\ \delta_{r0} + \delta_r' U_{gt} + O_p\left(\sqrt{\frac{n}{N}}\right) \\ \delta_{c0} + \delta_{c1} t + \delta_c' U_{gt} + u_{ct} \end{bmatrix} =: \delta_{w0} + \delta_{w1} t + D_w U_{gt} + u_{wt}^{+}, \tag{9}$$

where

$$\delta_{T1} = \frac{\gamma_3 \delta_{c1}}{1-\theta_1},\ \delta_T = \frac{\theta_2 \delta_r + \gamma_3 \delta_c}{1-\theta_1}, \tag{10}$$

$$u_{Tt} = \gamma_3 \sum_{j=0}^{\infty} \theta_1^j u_{ct-1-j} - \frac{\theta_1}{1-\theta_1} \sum_{k=0}^{\infty} \theta_1^k \left[\theta_2 \delta_r + \gamma_3 \delta_c\right]' u_{gt-1-k} - \delta_T' u_{gt} \tag{11}$$

$$u_{Tt}^{+} = u_{Tt} + O_p\left(\frac{1}{\sqrt{N}} + \sqrt{\frac{n}{N}} + t\,|\theta_1|^t\right), \tag{12}$$

where $D_w' = [\delta_T, \delta_r, \delta_c]$, $\delta_{w0} = [\delta_{T0}, \delta_{r0}, \delta_{c0}]'$, $\delta_{w1} = [\delta_{T1}, 0, \delta_{c1}]'$, $u_{wt}^{+} = u_{wt} + O_p\left(\frac{1}{\sqrt{N}} + \sqrt{\frac{n}{N}} + t\left|\theta_1\right|^t\right)$ and $u_{wt} = [u_{Tt}, 0, u_{ct}]'$.

From the trend representation (9) the following long run cointegrating relationship among the global variables $(T_t, R_t, \ln(CO_{2,t}))$ is obtained

$$T_t = \frac{\theta_2}{1-\theta_1} R_t + \frac{\gamma_3}{1-\theta_1} \ln(CO_{2,t}) + \frac{\mu}{1-\theta_1} + \frac{1}{1-\theta_1}\zeta_t, \tag{13}$$

using: (i) the fact that $\delta_T = \frac{\theta_2\delta_r + \gamma_3\delta_c}{1-\theta_1}$, which delivers cointegration among the stochastic trend components of $(T_t, R_t, \ln(CO_{2,t}))$; and (ii) the linkage $\delta_{T1} = \frac{\gamma_3\delta_{c1}}{1-\theta_1}$, which ensures deterministic co-movement of the linear trends in T_t and $\ln(CO_{2,t})$. The equation error (or equilibrium error correction) in (13) is

$$\frac{1}{1-\theta_1}\zeta_t = u_{Tt} - \frac{\gamma_3}{1-\theta_1} u_{ct} + O_p\left(\frac{1}{\sqrt{N}} + \sqrt{\frac{n}{N}} + t\left|\theta_1\right|^t\right), \tag{14}$$

which is a stationary, weakly dependent time series up to an asymptotically negligible component.

Importantly, the cointegrating relation (13) is distinct from the time specific effect λ_t. In fact, (13) represents the ultimate global linkage in these variables that results from integrating the time specific effects λ_t with the station-level adjustment mechanism and the global aggregation process that leads to $(T_t, R_t, \ln(CO_{2,t}))$. Moreover, the coefficient of $\ln(CO_{2,t})$ in the relationship (13) gives the transient climate sensitivity parameter (6) upon scaling by ln(2)

$$TCS = \frac{\gamma_3}{1-\theta_1} \times \ln(2) = \frac{\gamma_3}{1-\beta_1-\gamma_1} \times \ln(2), \tag{15}$$

which means that the TCS parameter can be estimated directly from appropriate econometric estimation (such as fully-modified least squares (FM-OLS)) of the long run cointegrating relation (13) without regard to the dynamic adjustment mechanism (1). That approach was followed in Phillips et al (2018) where an asymptotic theory of inference was developed for the methodology. That approach will not be pursued further in the present work although we do discuss later a key difference between the asymptotic theory of the resulting FM-OLS of TCS and the asymptotic theory of estimates of TCS based on panel regression estimates such as WG, diff-GMM, and sys-GMM.

3.2 Dynamic Panel Estimation and Invariance Properties

The alternative station-level approach uses panel regression methods to estimate both (1) and (2). This approach was used in Magnus et al. (2011) and Storelvmo et al. (2016). Specifically, sys-GMM methods were employed by Magnus et al. (2011) and Storelvmo et al. (2016) because their estimates of the panel autoregressive coefficient β_1 exceeded 0.9 and dynamic panel regressions with autoregressive coefficients close to unity are known to lead to weak

instrumentation in diff-GMM methods, thereby reducing efficiency but retaining consistency (Kruiniger, 2009; Phillips, 2018). In the present application, as might be expected given the global coverage of the station-level observations, there is considerable heterogeneity in the fixed effects α_i of the dynamic panel regression (1), a feature that is known to produce sys-GMM estimates of the coefficients in dynamic panel regression that can be substantially biased (Hayakawa, 2007, 2015). For this reason, we might expect some large differences in the coefficient estimates among these three panel regression procedures.

For the observational data used in Storelvmo et al. (2016) and Phillips et al. (2018), the differences are substantial, particularly between sys-GMM and the other two approaches. Table 1 below provides estimates of the parameters of the system (1) and (2). The massive difference between the sys-GMM estimate of the parameter β_1 (0.8665) and the estimates obtained by diff-GMM (0.1125) and WG (0.1346) is striking - the sys-GMM estimate is more than six times greater than the WG estimate and nearly eight times greater than the diff-GMM estimate. The implications of these differences for the station-level dynamic adjustment mechanism of temperature are enormous. Similar major differences occur in the estimation of the parameter γ_1 in the aggregate relation for λ_t. Table 1 also reports the ratio $r_\alpha = \hat{\sigma}_\alpha / \hat{\sigma}_u$ of the estimated standard deviation $\hat{\sigma}_\alpha$ of the fitted fixed effects α_i to the standard deviation of the fitted equation errors $u_{i,t+1}$. For the WG estimates, this ratio is 15.043 which is ten times greater than the corresponding value from sys-GMM, showing the major differences in how the two methods capture and represent the observed variation in the data at the local level.

Table 1: Dynamic Panel Regression and Transient Climate Sensitivity Estimates

	Estimation Method		
	WG	diff-GMM	sys-GMM
Parameter			
β_1	0.1346	0.1125	0.8665
β_2	-0.0001	-0.0048	0.0098
γ_1	-0.0230	-0.0010	-0.7549
γ_2	0.0262	0.0309	0.0162
γ_3	3.6400	3.6400	3.6400
$\beta_1 + \gamma_1$	0.1116	0.1116	0.1116
$\beta_2 + \gamma_2$	0.0261	0.0261	0.0260
$r_\alpha = \frac{\sigma_\alpha}{\sigma_u}$	15.043	12.825	1.4769
TCS	2.8399	2.8399	2.8399

Notes: $r_\alpha = \frac{\sigma_\alpha}{\sigma_u} = \left(\frac{\mathbb{V}(\alpha_i)}{\mathbb{V}(u_{it})} \right)^{1/2}$

Even more striking is that, in spite of the differences in the estimates of the individual coefficients, estimates of the composite parameters $\beta_1 + \gamma_1$, $\beta_2 + \gamma_2$,

and the transient climate sensitivity parameter TCS are all invariant to the method of estimation of the dynamic panel regression equation. This equivalence is established analytically in Theorem 1 below. An important implication of this analytic invariance is that the TCS estimate has the same asymptotic theory for the different panel regression methods and thus the same induced asymptotic confidence interval.

To proceed, it is convenient to write the model (1) and (2) in the form:

$$\begin{aligned} T_{i,t+1} &= \alpha_i + \beta_1 T_{i,t} + \beta_2 R_{i,t} + \lambda_t + u_{it+1} =: \alpha_i + \beta' X_{it} + \lambda_t + u_{it+1} \quad (16) \\ \lambda_t &= \gamma_0 + \gamma_1 T_t + \gamma_2 R_t + \gamma_3 \ln(CO_{2,t}) =: \gamma_0 + \gamma' W_t, \quad (17) \end{aligned}$$

with notation $X_{it} = (T_{i,t}, R_{i,t})'$ and $W_t = (T_t, R_t, \ln(CO_{2,t}))'$. It follows by aggregation and the normalization condition $\bar{\alpha} = 0$ that

$$T_{t+1} = \beta' X_t + \lambda_t + \bar{u}_{\cdot t+1} = \left(\beta', 0\right) W_t + \lambda_t + \bar{u}_{\cdot t+1}.$$

The model (16)-(17) can also be written in the combined factor augmented form

$$T_{i,t+1} = \gamma_0 + \alpha_i + \beta' X_{i,t} + \gamma' W_t + u_{it+1}, \quad (18)$$

which is a dynamic panel model with common factor given by the component $\gamma' W_t$ with observable W_t. The technical complications involved in the analysis of (18) arise because: (i) the common factor aggregate $W_t = \left(\bar{X}'_{\cdot t}, \ln(CO_{2,t})\right)'$ relates to the observable station level variables X_{it} that appear as regressors in (18), as well as the exogenous variable $CO_{2,t}$; (ii) the regressors $(X_{i,t}, W_t)$ have deterministic and stochastic trend components; and (iii) there is cointegration (both deterministic and stochastic) among the elements of the global aggregate W_t. Aggregating (18) and using the identification condition $\bar{\alpha} = 0$ gives (8). Setting $\theta = (\theta_1, \theta_2, \theta_3)' = (\theta_1, \theta_2, \gamma_3)'$, the global dynamic regression is

$$T_{t+1} = \gamma_0 + \theta_1 T_t + \theta_2 R_t + \gamma_3 \ln(CO_{2t}) + \bar{u}_{\cdot t+1} = \gamma_0 + \theta' W_t + \bar{u}_{\cdot t+1}, \quad (19)$$

which it is convenient to write in observation form as

$$T^+ = \gamma_0 \iota_{n-1} + W\theta + \bar{U} = (T_2, T_3, ..., T_n)' \quad (20)$$

where $T^+ = (T_2, T_3, ..., T_n)'$, $\bar{U} = [\bar{u}_{\cdot 2}, ..., \bar{u}_{\cdot n}]'$, and ι_{n-1} is an $(n-1)$ vector of ones.

We now proceed to analyze the estimation of this station-level system and to develop asymptotic theory for the resulting coefficient estimates and the associated TCS parameter. For the purpose of the discussion below it is convenient to work with the WG estimator. But, as will be demonstrated, the results obtained for the TCS parameter estimates (and for certain linear contrasts of the other coefficients, notably $\beta_1 + \gamma_1$ and $\beta_2 + \gamma_2$) apply also to diff-GMM and sys-GMM procedures.

The WG procedure involves the following steps.

Step 1. Estimate the dynamic panel model by least squares, which involves estimating the time specific effect λ_t as the time specific intercept in the regression (1). That is, applying least squares with intercept standardized so that $\bar{\alpha} = 0$, we obtain

$$\hat{\lambda}_t = T_{t+1} - \hat{\beta}' X_t = T_{t+1} - \hat{\beta}_1 T_t - \hat{\beta}_2 R_t = T_{t+1} - \left(\hat{\beta}', 0\right) W_t, \qquad (21)$$

with

$$\hat{\beta} = \left(\sum_{t=1}^{n-1} \sum_{i=1}^{N} \tilde{X}_{i,t} \tilde{X}'_{i,t} \right)^{-1} \left(\sum_{t=1}^{n-1} \sum_{i=1}^{N} \tilde{X}_{i,t} \tilde{T}_{i,t+1} \right), \qquad (22)$$

where we use the notation $\tilde{A}_{it} = A_{it} - \bar{A}_{i\cdot} - \bar{A}_{t\cdot} + \bar{A}_{\cdot\cdot}$ with $\bar{A}_{i\cdot} = (n-1)^{-1} \sum_{t=1}^{n-1} A_{it}$, $\bar{A}_{t\cdot} = N^{-1} \sum_{i=1}^{N} A_{it}$, and $\bar{A}_{\cdot\cdot} = (n-1)^{-1} N^{-1} \sum_{t=1}^{n-1} \sum_{i=1}^{N} A_{it}$. This means that the time specific and station specific effects are estimated by regression elimination and the slope coefficients β are estimated using pooled least squares regression after elimination of these effects.

Step 2. Regress the fitted $\hat{\lambda}_t$ on $(1, T_t, R_t, \ln(CO_{2,t}))$ by least squares giving

$$\hat{\lambda}_t = \hat{\gamma}_0 + \hat{\gamma}_1 T_t + \hat{\gamma}_2 R_t + \hat{\gamma}_3 \ln(CO_{2,t}) = \hat{\gamma}_0 + \hat{\gamma}' W_t, \qquad (23)$$

and the corresponding vector of coefficient estimates $\left(\hat{\gamma}_0, \hat{\gamma}'\right) = \left(\hat{\gamma}_0, \hat{\gamma}_1, \hat{\gamma}_2, \hat{\gamma}_3\right)$.

Step 3. Estimate the TCS parameter using the coefficient estimates $\left(\hat{\beta}_1, \hat{\gamma}_1, \hat{\gamma}_3\right)$ giving

$$\widehat{TCS} = \frac{\hat{\gamma}_3}{1 - \hat{\beta}_1 - \hat{\gamma}_1} \times \ln(2).$$

Steps 1 and 2 may be amalgamated in a combined least squares regression that minimizes the following objective function with respect to $\left(\alpha_i, \beta', \gamma_0, \gamma'\right)$ subject to the identification condition that $\bar{\alpha} = \sum_{i=1}^{N} \alpha_i = 0$

$$\sum_{t=1}^{n-1} \sum_{i=1}^{N} \left\{ T_{i,t+1} - \alpha_i - \beta' X_{i,t} - \gamma_0 - \gamma' W_t \right\}^2, \qquad (24)$$

which leads to the same estimates of the coefficients $\left(\hat{\beta}, \hat{\gamma}\right)$ as those obtained by following Steps 1 and 2 above. Writing the vector of estimated time effects obtained in Step 1 as $\hat{\lambda} = \left(\hat{\lambda}_t\right)$, it is apparent from (23) that the slope coefficients estimates of γ take the form

$$\hat{\gamma} = \left(\tilde{W}'\tilde{W}\right)^{-1} \left(\tilde{W}'\hat{\lambda}\right). \qquad (25)$$

where $\tilde{W} = W - \bar{W}$ is the matrix of deviations from time series means $\tilde{W}_t = W_t - \bar{W}$.

The estimates $\left(\hat{\beta},\hat{\gamma}\right)$ and implied estimate $\widehat{TCS}$ of the TCS parameter above are all obtained using WG estimation of the panel regression system (16)-(17). Somewhat remarkably, as the following result shows, the resulting estimate $\widehat{TCS}$ as well as the corresponding estimates of the linear contrasts $\hat{\theta}_i = \hat{\beta}_i + \hat{\gamma}_i$, $(i = 1, 2)$, are invariant to the method of estimation of the panel regression equation estimates $\left(\hat{\beta},\hat{\gamma}\right)$, whether by WG, diff-GMM or sys-GMM.

Theorem 1 (Estimation Invariance) *Station-level estimation of the dynamic panel regression model (16)-(17) by the methods WG, diff-GMM, and sys-GMM all lead to the common estimate* $\hat{\theta} = \left(\tilde{W}'\tilde{W}\right)^{-1}\tilde{W}'T^{+}$ *of the slope coefficient* θ *in (19) and the common estimate of the* TCS parameter

$$\widehat{TCS} = \frac{\hat{\gamma}_3}{1-\hat{\theta}_1} \times \ln(2).$$

Remarks

1. Whereas the composite estimates $\left(\hat{\beta}_1 + \hat{\gamma}_1\right)$ and $\left(\hat{\beta}_2 + \hat{\gamma}_2\right)$ are invariant to the panel regression method employed, the individual estimates $\left(\hat{\beta}_1, \hat{\beta}_2, \hat{\gamma}_1, \hat{\gamma}_2\right)$ are not invariant but involve compensatory adjustments between the estimated coefficients $\hat{\beta}_i$ and $\hat{\gamma}_i$. As shown in the proof of Theorem 1 these adjustments ensure that the estimation error for the composite estimate $\hat{\theta}$ of θ in (19) satisfy the system

$$\hat{\theta} - \theta = \begin{bmatrix} \left(\hat{\beta}_1 + \hat{\gamma}_1\right) - (\beta_1 + \gamma_1) \\ \left(\hat{\beta}_2 + \hat{\gamma}_2\right) - (\beta_2 + \gamma_2) \\ (\hat{\gamma}_3 - \gamma_3) \end{bmatrix} = \left(\tilde{W}'\tilde{W}\right)^{-1}\tilde{W}'\bar{U}, \qquad (26)$$

   which is determined solely by least squares regression of (the aggregate time series matrices) $\bar{U}$ on $\tilde{W}$, making the estimation error $\hat{\theta} - \theta$ of the composite parameters invariant to the method of estimation of the panel regression equation (16).

2. The intuitive explanation for this invariance is that the individual estimated coefficients $\hat{\beta}_i$ in (16) depend on data $(T_{i,t+1}, X_{i,t})$ that upon aggregation necessarily satisfy the global dynamic relationship (19), which upon time series demeaning is just $\tilde{T}_{t+1} = \theta'\tilde{W}_t + \tilde{\bar{u}}_{.t+1}$ and which, in turn, depends only on the composite vector θ. Thus, the parameter θ may be interpreted as a global composite parameter and this aggregate relationship may be interpreted as a reduced form dynamic equation for the global variables. The estimates $\left(\hat{\beta}_1, \hat{\beta}_2\right)$ of the system parameters (β_1, β_2) are used to estimate the time specific effects λ_t by cross section aggregation

giving $\hat{\lambda}_t$ as shown in (21), giving $\hat{\lambda}_t = T_{t+1} - \left(\hat{\beta}', 0\right) W_t$. Correspondingly, when the parameter γ is estimated in (23) using these specific fitted values $\hat{\lambda}_t$, we have $\hat{\gamma} = \left(\check{W}'\check{W}\right)^{-1}\left(\check{W}'\hat{\lambda}\right)$, so that the resulting estimates satisfy

$$\hat{\gamma} = \left(\check{W}'\check{W}\right)^{-1}\check{W}'\left\{T^{+} - W\begin{bmatrix}\hat{\beta}\\ 0\end{bmatrix}\right\} = \left(\check{W}'\check{W}\right)^{-1}\check{W}'T^{+} - \begin{bmatrix}\hat{\beta}\\ 0\end{bmatrix} \quad (27)$$

or

$$\hat{\theta} = \begin{bmatrix}\hat{\beta}_1 + \hat{\gamma}_1\\ \hat{\beta}_2 + \hat{\gamma}_2\\ \hat{\gamma}_3\end{bmatrix} = \left(\check{W}'\check{W}\right)^{-1}\left(\check{W}'T^{+}\right), \quad (28)$$

showing invariance and the manner in which the compensatory adjustments in the composite estimates are automatically embodied by virtue of the cross section aggregation and the regression (23). In effect, the estimate $\hat{\gamma}$ adjusts to whichever specific fitted values $\hat{\lambda}_t$ are obtained from the particular panel regression method of estimation that produces the estimates $\left(\hat{\beta}_1, \hat{\beta}_2\right)$. Thus, the estimates $\left(\hat{\beta}_1 + \hat{\gamma}_1\right)$, $\left(\hat{\beta}_2 + \hat{\gamma}_2\right)$, and $\hat{\gamma}_3$ of $(\beta_1 + \gamma_1)$, $(\beta_2 + \gamma_2)$, and γ_3 are each invariant to the choice of estimation procedure for the coefficients β in the panel regression (16). In every case, the estimate $\hat{\theta}$ of the composite parameter θ ends up taking the same value and is invariant to the panel regression method.

4 Asymptotic Theory

In view of the invariance properties established in Theorem 1, It is convenient to do the analysis with the (invariant) composite parameter estimate $\hat{\theta}$ and the implied estimate $\widehat{TCS}$. It is also convenient to fix ideas by working with the WG estimates of the parameters (β, γ) and, hence, θ and TCS.

We start by writing the common trend representation (9) as

$$W_t = \delta_{w0} + \delta_w t + V_t + u_{wt}^{+}, \quad (29)$$

where $V_t = D_w U_{gt} =: [V_{Tt}, V_{rt}, V_{ct}]'$, and $\delta_w = (\delta_{T1}, 0, \delta_{c1})'$. Subtracting time series means gives $\check{W}_t = \delta_w \check{t} + \check{V}_t + \check{u}_{wt}^{+}$, with $\check{V}_t = V_t - \bar{V}$. Then

$$\hat{\theta} - \theta = \left(\check{W}'\check{W}\right)^{-1}\left(\check{W}'\bar{U}\right), \quad (30)$$

and the limit theory needs to take account of degeneracy in the asymptotic form of the sample moment matrix $\check{W}'\check{W} = \sum_{t=1}^{n} \check{W}_t \check{W}_t'$ arising from the presence of both linear and stochastic trends in W_t. We remark also that asymptotics for the second component of (30), $\check{W}'\bar{U} = \sum_{t=1}^{n} \check{W}_t \bar{u}_{\cdot t+1}$, depends on the behavior of the cross section averaged elements $\bar{u}_{\cdot t+1} = N^{-1}\sum_{i=1}^{N} u_{i,t+1}$. Under Assumption A(i) and using $\rightsquigarrow$ to denote weak convergence, these elements satisfy a

CLT $\sqrt{N}\bar{u}_{\cdot t+1} \rightsquigarrow \xi_{t+1}$, say, and are therefore of order $O_p\left(N^{-1/2}\right)$. Further, in view of Assumptions A(i)-(iv), we have the functional laws $n^{-1/2}V_{\lfloor nr \rfloor} \rightsquigarrow V(r)$, $n^{-1/2}\tilde{V}_{\lfloor nr \rfloor} \rightsquigarrow \tilde{V}(r) = V(r) - \int_0^1 V(s)\,ds$, and an implied functional law for partial sums of the limit variates ξ_{t+1}, viz., $n^{-1/2}\sum_{t=1}^{\lfloor nr \rfloor} \xi_t \rightsquigarrow B_\xi(r) \equiv BM\left(\sigma_\xi^2\right)$, where $\sigma_\xi^2 = \mathbb{E}\left(\xi_t^2\right) = \sigma_u^2$.

To handle the asymptotic degeneracy of the sample moment matrix, we proceed in the usual fashion by rotation of the coordinate system of the regressors to isolate directions of different magnitudes (Park and Phillips, 1988, 1989). Define the deterministic trend direction $h = \delta_w / \left(\delta_w' \delta_w\right)^{1/2}$ in (29) and let $H_\perp$ be an orthogonal complement of h so that the matrix

$$H = [h, H_\perp] = \frac{1}{\left(\delta_w' \delta_w\right)^{1/2}} \begin{bmatrix} \frac{\delta_{c1}\gamma_3}{1-\theta_1} & 0 & -\delta_{c1} \\ 0 & \left(\delta_w' \delta_w\right)^{1/2} & 0 \\ \delta_{c1} & 0 & \frac{\delta_{c1}\gamma_3}{1-\theta_1} \end{bmatrix}, \tag{31}$$

is orthogonal and $\delta_w' \delta_w = \left(\frac{\delta_{c1}}{1-\theta_1}\right)^2 \left\{(1-\theta_1)^2 + \gamma_3^2\right\}$. Rotating the system by H gives

$$H'\tilde{W}_t = \begin{cases} a\tilde{t} + h'\tilde{V}_t + O_p(1) \\ V_{\perp t} + O_p(1) \end{cases}, \text{ with } a = \left(\delta_w' \delta_w\right)^{1/2}, \tag{32}$$

which isolates the deterministic trend in the leading coordinate and the stochastic trend in the remaining coordinates, which we have written as $V_{\perp t} = H_\perp' V_t$. Corresponding to these coordinates, define the scaling matrix $D_n = \text{diag}\left(n^{3/2}, nI_2\right)$.

With these preliminaries, we are able to state the following asymptotic result concerning the composite parameter estimate $\hat{\theta}$ in (30) and its mixed normal ($\mathcal{MN}$) limit theory corresponding to the different directions of deterministic and stochastic trends in the component variables. This result also enables us to derive the limit theory for the estimate of $TCS = \frac{\theta_3}{1-\theta_1} \times \ln(2) =: g(\theta_a)$, where $\theta_a = (\theta_1, \theta_3)$, based on the panel regression estimate $\hat{\theta}_a$.

Theorem 2 Under Assumption A and as $(n, N) \to \infty$:

(i) $n\sqrt{N}\left(\hat{\theta} - \theta\right) \rightsquigarrow H_\perp \left(\int_0^1 \tilde{V}_{\perp,\tilde{r}} \tilde{V}_{\perp,\tilde{r}}'\right)^{-1} \int_0^1 \tilde{V}_{\perp,\tilde{r}} dB_\xi \equiv \mathcal{MN}\left(0, \sigma_\xi^2 H_\perp \left(\int_0^1 \tilde{V}_{\perp,\tilde{r}} \tilde{V}_{\perp,\tilde{r}}'\right)^{-1} H_\perp'\right)$;

(ii) $n^{3/2}\sqrt{N}h'\left(\hat{\theta} - \theta\right) \rightsquigarrow \left(a\int_0^1 \tilde{r}_{\tilde{V}_\perp}^2\right)^{-1} \left(\int_0^1 \tilde{r}_{\tilde{V}_\perp} dB_\xi\right) \equiv \mathcal{MN}\left(0, \frac{\sigma_\xi^2}{a^2}\left(\int_0^1 \tilde{r}_{\tilde{V}_\perp}^2\right)^{-1}\right)$;

(iii) $n^{3/2}\sqrt{N}\left(\widehat{TCS} - TCS\right) \rightsquigarrow \frac{1}{\delta_{c1}(1-\theta_1)}\left(\int_0^1 \tilde{r}_{\tilde{V}_\perp}^2\right)^{-1} \left(\int_0^1 \tilde{r}_{\tilde{V}_\perp} dB_\xi\right) \equiv \mathcal{MN}\left(0, \frac{\sigma_\xi^2}{\delta_{c1}^2(1-\theta_1)^2}\left(\int_0^1 \tilde{r}_{\tilde{V}_\perp}^2\right)^{-1}\right)$,

where $\tilde{r}(r) = r - \int_0^1 s\,ds$, $\tilde{r}_{\tilde{V}_\perp} = \tilde{r} - \left(\int_0^1 \tilde{r}\tilde{V}_\perp\right)\left(\int_0^1 \tilde{V}_\perp \tilde{V}_\perp'\right)^{-1} \tilde{V}_\perp$, and $\tilde{V}_{\perp,r}(r) = \tilde{V}_\perp(r) - \left(\int_0^1 \tilde{V}_\perp \tilde{r}\right)\left(\int_0^1 \tilde{r}^2\right)^{-1} \tilde{r}$.

Remarks

3. In (i) and (ii), $\tilde{V}_{\perp,\tilde{r}}(r)$ is the L_2 projection residual of $\tilde{V}_\perp = H_\perp' \tilde{V}$ on $\tilde{r}$, and $\tilde{r}_{\tilde{V}_\perp}$ is the L_2 projection residual of $\tilde{r}$ on $\tilde{V}_\perp$.

4. In the deterministic trend direction h, (ii) shows that $h'\hat{\theta}$ has the faster convergence rate $O\left(n^{3/2}\sqrt{N}\right)$ consonant with both a deterministic linear trend and cross section aggregation effects. In the alternate direction $H_\perp$, the stochastic trend dominates and the convergence rate is $O\left(n\sqrt{N}\right)$, combining the influence of the stochastic trend and cross section aggregation, giving

$$\begin{aligned} n\sqrt{N}H_\perp'\left(\hat{\theta}-\theta\right) &\rightsquigarrow \left(\int_0^1 \tilde{V}_{\perp,\tilde{r}}\tilde{V}_{\perp,\tilde{r}}'\right)^{-1}\int_0^1 \tilde{V}_{\perp,\tilde{r}}dB_\xi \\ &\equiv \mathcal{MN}\left(0,\sigma_\xi^2\left(\int_0^1 \tilde{V}_{\perp,\tilde{r}}\tilde{V}_{\perp,\tilde{r}}'\right)^{-1}\right). \end{aligned} \tag{33}$$

   This slower rate of convergence also dominates the limit distribution theory for the full vector $\hat{\theta}$, which is a singular mixed normal distribution with support determined by the range space of $H_\perp$, as given by (i).

5. As shown in the Appendix

$$n^2\left(\tilde{W}'\tilde{W}\right)^{-1} \rightsquigarrow H_\perp\left(\int_0^1 \tilde{V}_{\perp\tilde{r}}\tilde{V}_{\perp\tilde{r}}'\right)^{-1}H_\perp',$$

   so that the usual formula $\hat{\sigma}_u^2\left(\tilde{W}'\tilde{W}\right)^{-1}$, employing a consistent estimate $\hat{\sigma}_u^2$ of panel regression equation error variance σ_u^2, suffices for the asymptotic variance matrix in (i). This formula holds in spite of the degenerate asymptotic rank of the signal matrix $\tilde{W}'\tilde{W}$ and the scaling by $\sqrt{N}$ of the estimation error in (i). The reason for the latter is that the estimation error $\hat{\theta}-\theta = \left(\tilde{W}'\tilde{W}\right)^{-1}\left(\tilde{W}'\bar{U}\right)$ from (30), and the moment matrix $\tilde{W}'\bar{U} = \sum_{t=1}^n \tilde{W}_t\bar{u}_{\cdot t+1}$ involves the cross section sample mean $\bar{u}_{\cdot t+1} = N^{-1}\sum_{i=1}^N u_{i,t+1}$ whose variance is σ_u^2/N, so cross section sample size scaling is already implicitly incorporated in $\frac{\hat{\sigma}_u^2}{N}\left(\tilde{W}'\tilde{W}\right)^{-1}$ and the estimated variance matrix of $n\sqrt{N}\left(\hat{\theta}-\theta\right)$ is then $\hat{\sigma}_u^2 n^2\left(\tilde{W}'\tilde{W}\right)^{-1} \rightsquigarrow \sigma_\xi^2 H_\perp\left(\int_0^1 \tilde{V}_{\perp\tilde{r}}\tilde{V}_{\perp\tilde{r}}'\right)^{-1}H_\perp'$, as required.

6. The $n\sqrt{N}$ convergence rate of $\hat{\theta}$ is explained by the use of cross section averaging in conjunction with time series averaging in the presence of nonstationary data with stochastic trends in the direction $H_\perp$. As discussed

in Remark 2, estimation of θ by panel regression techniques essentially involves, after cross section aggregation, estimation of the global dynamic relationship (19), or $T_{t+1} = \gamma_0 + \theta' W_t + \bar{u}_{\cdot t+1}$. Upon time series demeaning, the global dynamics follow the equation

$$\tilde{T}_{t+1} = \theta' \tilde{W}_t + \tilde{\bar{u}}_{\cdot t+1} \tag{34}$$

which, in turn, depends only on the composite vector θ. Thus, the parameter θ may be interpreted as a global composite parameter and this aggregate relationship may be interpreted as a reduced form dynamic equation for the global variables. The error in (34) is

$$\tilde{\bar{u}}_{\cdot t+1} = \bar{u}_{\cdot t+1} - \frac{1}{n}\sum_{t=1}^{n-1} u_{\cdot t+1} = \frac{1}{N}\sum_{i=1}^{N} u_{i,t+1} - \frac{1}{nN}\sum_{t=1}^{n-1}\sum_{i=1}^{N} u_{i,t+1} = O_p\left(\frac{1}{\sqrt{N}}\right), \tag{35}$$

where the $O_p\left(N^{-1/2}\right)$ order holds under Assumption A(i) in which the dynamic panel regression errors of (1) are assumed to satisfy $\{u_{it}\} \sim_{iid} \left(0, \sigma_u^2\right)$ over i and t. WG, diff-GMM, and sys-GMM estimation of the components of θ all lead, as shown by the invariance result of Theorem 1, to least squares regression on (34), whose error is $O_p\left(N^{-1/2}\right)$, which in turn affects the convergence rate of all the respective coefficient estimates by $\sqrt{N}$ scaling. In consequence, the deterministic and stochastic trends in the global vector variable W_t lead to the dual convergence rates of $n\sqrt{N}$ and $n^{3/2}\sqrt{N}$ for $\hat{\theta}$ in the respective directions $H_\perp$ and h (in (33) and (ii)) where each rate is scaled by the $\sqrt{N}$ factor in view of (35)[1].

7. When $N \to \infty$ and Assumption A holds, the convergence rate $n^{3/2}\sqrt{N}$ of $\widehat{TCS}$ exceeds the convergence rate $n^{3/2}$ of the FM-OLS estimator of TCS studied in Phillips et al. (2018). This divergence is explained as follows. The FM-OLS estimator of TCS is based on a cointegrating regression estimation of equation (13) among the elements of W_t in which the TCS parameter appears directly as the coefficient of the $\ln(CO_{2t})$ variable scaled by $\ln(2)$. Upon time series demeaning this cointegrating equation has the form

$$\tilde{T}_t = \frac{\theta_2}{1-\theta_1}\tilde{R}_t + \frac{\gamma_3}{1-\theta_1}\widetilde{\ln(CO_{2,t})} + \frac{1}{1-\theta_1}\tilde{\zeta}_t, \tag{36}$$

[1] The $n\sqrt{N}$ and $n^{3/2}\sqrt{N}$ rates of convergence apply under (35) and $\bar{u}_{\cdot t+1} = \frac{1}{N}\sum_{i=1}^{N} u_{i,t=1} \to_p 0$. More generally by the ergodic theorem under cross section stationarity, $\bar{u}_{\cdot t+1} = \frac{1}{N}\sum_{i=1}^{N} u_{i,t=1} \to_{a.s.} \mathbb{E}\left(u_{i,t+1}|\mathcal{C}_{t+1}\right) =: \zeta_{t+1}$ where $\mathcal{C}_{t+1}$ is a filtration on the probability space of the aggregate variables that is generated by time series common global shocks. In such cases, the convergence rate is $O(n)$ and $O\left(n^{3/2}\right)$ rather than $O\left(n\sqrt{N}\right)$ and $O\left(n^{3/2}\sqrt{N}\right)$; and the corresponding limit distributions are affected by the time series properties of the global common shock process ζ_t. The FM-OLS estimator used in Phillips et al (2018) is robust to this extension under general weak dependence conditions on ζ_t because endogeneity and serial dependence is accounted for in FM-OLS regression. Panel regression estimators based on WG and GMM methods do not take such effects into account and are generally inconsistent, as would be expected in dynamic models with serially dependent disturbances.

where ζ_t, which is given by (14), is a stationary, weakly dependent equilibrium error term up to an asymptotically negligible residual component. In (36) the panel regression errors u_{it} have been eliminated up to an asymptotically negligible term by cross section averaging. The dominant $O_p(1)$ component of ζ_t in (14) is the composite stationary error

$$\begin{aligned} u_{Tt} - \frac{\gamma_3}{1-\theta_1} u_{ct} &= \gamma_3 \sum_{j=0}^{\infty} \theta_1^j u_{ct-1-j} - \frac{\theta_1}{1-\theta_1} \sum_{k=0}^{\infty} \theta_1^k \left[\theta_2 \delta_r + \gamma_3 \delta_c\right]' u_{gt-1-k} \\ &\quad -\delta_T' u_{gt} - \frac{\gamma_3}{1-\theta_1} u_{ct}, \end{aligned}$$

which is a serially dependent linear process of the innovations (u_{ct}, u_{gt}). Thus, (36) is a cointegrating regression equation with asymptotically stationary errors. The use of FM-OLS regression and other efficient methods of cointegrating equation estimation therefore produces asymptotically unbiased and asymptotically efficient estimates of the coefficients in (36) whose rates of convergence are determined by the trend behavior of the component regressors. Since $\ln(CO_{2,t})$ has a linear deterministic drift, the coefficient of this variable in (36) and hence the implied estimate of the TCS parameter have a convergence rate of $O\left(n^{3/2}\right)$, as shown in Phillips et al. (2018). By contrast, under Assumption A(i) and specifically the requirements that: (a) $\{u_{it}\} \sim_{iid} \left(0, \sigma_u^2\right)$ over i and t; and (b) that the energy balance (time specific effect) variable λ_t is not subject to measurement error, the convergence rate of panel dynamic regression estimation of TCS by WG (or the GMM methods) is $O\left(n^{3/2}\sqrt{N}\right)$. Violations of condition (a) that introduce serial dependence in u_{it} lead to endogeneity in the dynamic panel regression with consequent effects (including inconsistency) on the asymptotics of these panel regression estimates. Violations of (b) induce a time series measurement error ($u_{\lambda t}$, say) into the factor augmented form of the global dynamic regression equation (18). The presence of such time series measurement errors in λ_t mean that the global dynamic regression equation (34) now has a residual $\tilde{u}_{\lambda t}$ of order $O_p(1)$, rather than a residual of order $O_p\left(N^{-1/2}\right)$ as in (35). This affects the rate of convergence, which becomes at most $O\left(n^{3/2}\right)$ – like that of FM-OLS – and introduces the possibility of endogeneity and serial correlation bias induced by the properties of $u_{\lambda t}$. In consequence, Theorem 2 only holds under the strict environment of Assumption A(i) or analogous stationary and ergodic martingale difference assumptions. Accordingly, the use of the long-run cointegrating regression equation (36) to estimate the parameter TCS by methods such as FM-OLS that take weak dependence and possible endogeneity of the composite errors into account provides a more robust approach to the estimation of transient climate sensitivity and, as a result, seems preferable to the use of direct panel regression methods such WG, diff-GMM, and sys-GMM.

8. From (iii) the (conditional) variance of the limit distribution of $\widehat{TCS}$ is

$\frac{\sigma_\xi^2}{\delta_{c1}^2(1-\theta_1)^2}\left(\int_0^1 \tilde{r}_{\tilde{V}_\perp}^2\right)^{-1}$, which is seen to diverge when $\delta_{c1} \to 0$ or $\theta_1 = \beta_1 + \gamma_1 \to 1$. The reason for divergence is that when $\delta_{c1} = 0$ there is no deterministic trend in $\ln(CO_{2t})$ and hence no deterministic trend in T_t or the common trend representation given in Theorem 1. In this case, the rate of convergence is $O\left(n\sqrt{N}\right)$ not $O\left(n^{3/2}\sqrt{N}\right)$, explaining the divergence in the result (iii). When $\theta_1 = 1$, there is a second unit root in the global dynamic regression equation (19), implying that T_t now has a quadratic deterministic trend and does not (deterministically) co-move or cointegrate with $\ln(CO_{2t})$ and R_t. In this case, the joint limit distribution of $\hat{\theta}$ is again singular but is now dominated by the stochastic trend component (which has the lowest order in the signal moment matrix), so the rate of convergence is again $O\left(n\sqrt{N}\right)$ rather than $O\left(n^{3/2}\sqrt{N}\right)$, explaining the divergence of the limit variance in (iii) when $\theta_1 \to 1$.

9. Under Assumption A, it follows from Theorem 2 and is shown in the Appendix that, using (iii), we can construct by dynamic panel regression an asymptotically valid $100(1-\alpha)\%$ confidence interval for the TCS parameter. This interval has the form

$$\widehat{TCS} \pm z_\alpha \left\{\hat{\sigma}_u^2 \hat{g}_a' E_a \left(\tilde{W}'\tilde{W}\right)^{-1} E_a' \hat{g}_a\right\}^{1/2}, \tag{37}$$

where $\hat{\sigma}_u^2$ is a consistent estimate of σ_u^2, E_a is the selector matrix

$$E_a = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and

$$\hat{g}_a' = g_a'\left(\hat{\theta}_1, \hat{\theta}_3\right) = \ln(2)\left(\frac{\hat{\theta}_3}{\left(1-\hat{\theta}_1\right)^2}, \frac{1}{1-\hat{\theta}_1}\right),$$

is the estimated gradient vector of the function $TCS = g(\theta_a)$ evaluated at $\hat{\theta}_a = \left(\hat{\theta}_1, \hat{\theta}_3\right)'$, and z_α is the $100(1-\alpha/2)$ percentile of the standard normal distribution.

5 Simulation Evidence

We report below results of a small simulation exercise with panel WG (within group least squares), diff-GMM (Difference GMM), and (non-optimal) sys-GMM (System GMM) estimation of the parameters in the following panel ARX(1) model (Storelvmo et al, 2015):

$$T_{it} = \alpha_i + \beta_1 T_{it-1} + \beta_2 R_{it-1} + \lambda_{t-1} + u_{it},\ t = 1,..,n;\ i = 1,...,N \tag{38}$$

$$\lambda_t = \gamma_0 + \gamma_1 \bar{T}_{\cdot t} + \gamma_2 \bar{R}_{\cdot t} + \gamma_3 \log(CO_{2t}) \tag{39}$$

with $u_{it} \sim_{iid} N\left(0, \sigma_u^2\right)$ and $\alpha_i \sim_{iid} N\left(0, \sigma_\alpha^2\right)$ and parameter settings based on the WG estimates obtained using the observed climate data with $n = 42$, $N = 963$, viz.,

$$\begin{aligned} \beta_1 &= 0.1346, \beta_2 = -0.0001, && (40) \\ \gamma_0 &= -14.915, \gamma_1 = -0.023, \gamma_2 = 0.0262, \gamma_3 = 3.640 && (41) \\ \sigma_\alpha &= 7.4147, \sigma_u = 0.4929 && (42) \end{aligned}$$

The simulations utilize the observed exogenous data on $(R_{it}, CO_{2t})_{t=1,i=1}^{n,N}$ and use (38) and (39) to generate simulated data for $(T_{it}, \lambda_t)_{t=1,i=1}^{n,N}$ recursively based on the parameter settings (40) - (41). The exercise is designed to shed light on the finite sample properties of various dynamic panel regression procedures in the context of the climate model (38) and (39) with data that relates closely to what was used in the empirical study.

The Figures 1-4 collected in Appendix A show densities of the WG, diff-GMM, and sys-GMM estimates of the first equation (38) of this model based on $R = 1500$ replications with sample sizes $n = 42$, $N = 50$, using only the first 50 cross section observations of $(R_{it}, CO_{2t})_{t=1,i=1}^{n,N}$ and therefore much smaller than the observational cross section sample size $N = 963$. The data were generated as described above with true parameter settings (40) - (41) and observed data for radiation R_{it} and CO_2 equivalent CO_{2t}. Simulation results based on the full cross section sample size $N = 963$ are reported in the subsequent Figures 5-8.

The WG densities show little bias (as might be expected with time series sample size $n = 42$) and seem to conform well with asymptotic normality for both β_1 and β_2. The Diff_GMM estimates show little bias in the estimation of β_2 but show downward bias in the estimation of β_1, and have much greater variance than the WG estimates, for both β_1 and β_2. By contrast the sys-GMM estimates are biased for both parameters. The sys-GMM estimates of β_1 are particularly heavily biased upwards from a true value of $\beta_1 = 0.135$ to a value around unity. The reason is the large ratio

$$r_\alpha = \frac{\sigma_\alpha}{\sigma_u} = \frac{7.4147}{0.4929} = 15.043,$$

of the standard deviation of the individual effects relative to the equation error. System GMM (both optimal and non-optimal versions) is known to be very sensitive to heterogeneity in the fixed effects α_i and, in particular, to the magnitude of r_α^2 (Hayakawa, 2015; Bun and Windmeijer, 2010), which in the present case is $r_\alpha^2 = 226.29$. For a simple panel AR(1) model with fixed effects, for instance, Hayakawa shows that non-efficient system GMM is actually inconsistent when $n/N \to c > 0$ and the probability limit of the system GMM estimate of β_1 tends to unity when $r_\alpha^2 \to \infty$. This analytic finding corresponds closely with the simulation results obtained here for the more complex model (38) - (39) with its multiple sources of nonstationarity.

These simulations confirm the existence of substantial bias in system GMM estimation in the present context. The findings are very similar for the data-

realistic sample size setting $N = 963$, although the distributions are much tighter in view of the larger value of the cross section sample size N. Interestingly, the system GMM estimates of β_1 in this case are centred around 0.8 rather than unity, which corresponds closely to the sys-GMM estimate obtained with the observed data where $\hat{\beta}_1 = 0.864$ (see Table 1). Moreover, since the ratio $\frac{n}{N} = \frac{42}{963} = 0.0436$ is close to zero in this case, Hayakawa's (2015) expression for the bias in his theorem 4(a) indicates that the bias will be smaller for $N = 963$ than when $N = 50$ and this analytic result for the bias matches the simulation findings for the temperature data.

6 Concluding Remarks

Panel data econometric methods seem well suited to assess the impact on global temperature of rising greenhouse gas (GHG) concentrations in Earth's atmosphere. They have the advantage of modeling the aggregate impact of GHG on temperature while also incorporating the effects of changes in downwelling surface radiation at the station level. In this way, panel models may account for some of the observed 'local dimming' that has occurred during the past half century due to rising levels of local pollution. Recent work by Magnus et al. (2011) and Storelvmo et al. (2016) sought to model these effects through system estimation of a dynamic panel regression framework, finding that the dimming influence of aerosols on surface radiation masked more than 30% of the aggregate effect of rising CO_2 levels on Earth's average temperature.

The analytic and simulation results of the present paper show that these local dimming effects are robust to the econometric methodology used to estimate Earth's transient climate sensitivity. Estimates of this aggregate-level parameter are found to be invariant to the dynamic panel regression method employed. However, estimates of some of the individual parameters in the dynamic panel regression system can differ substantially. In particular, system GMM methods are found to be unreliable in estimating the panel autoregressive coefficient and certain aggregate parameters, suffering from considerable bias. Both the simulation and analytic results favor within group methods for time series and cross section sample sizes of the order now available in observed spatio-temporal datasets. Within group panel estimation also gives results that are broadly in line with findings from direct time series cointegrating regressions of the aggregate data. This correspondence between the results of methods that employ disaggregate and aggregate data gives some assurance of the reliability of the estimates of climate sensitivity to CO_2 levels.

7 References

Bun, M. J. G. and F. Windmeijer (2010). The weak instrument problem of the system GMM estimator in dynamic panel data models, *Econometrics Journal*, 13, 95-126.

Hayakawa, K. (2007) Small sample bias properties of the system GMM estimator in dynamic panel data models. *Economics Letters,* 95, 32–38.

Hayakawa, K. (2015). The asymptotic properties of the system GMM estimator in dynamic panel data models when both N and T are large, *Econometric Theory*, 31, 647-667.

Hofmann, D. J., J. H. Butler, E. J. Dlugokencky, J. W. Elkins, K. Masarie, S. A. Montzka & P. Tans (2006). "The role of carbon dioxide in climate forcing from 1979 to 2004: introduction of the Annual Greenhouse Gas Index" *Tellus B,* 58, 614–619.

Kaufmann R. K., Kauppi H., Mann M. L., Stock J. H. (2011) "Reconciling anthropogenic climate change with observed temperature 1998–2008, *Proc Natl Acad Sciences,* 108 (29), 11790–11793.

Kaufmann R. K., Kauppi H., Mann M. L., Stock J. H. (2013). "Does temperature contain a stochastic trend: linking statistical results to physical mechanisms," *Climatic Change,* 118, 729–743.

Kruiniger, H. (2009) "GMM Estimation of Dynamic Panel Data Models with Persistent Data," *Econometric Theory,* 25, 1348–1391.

Magnus, J. R., Melenberg, B. & Muris, C. (2011). "Global Warming and Local Dimming: The Statistical Evidence," *Journal of the American Statistical Association,* 106, 452-464.

Park, J. Y. and P. C. B. Phillips (1988). "Statistical Inference in Regressions With Integrated Processes: Part 1," *Econometric Theory* 4, 468–497.

Park, J. Y. and P. C. B. Phillips (1989). "Statistical Inference in Regressions With Integrated Processes: Part 2," *Econometric Theory* 5, 95-131.

Phillips, P. C. B. (2018) "Dynamic Panel Anderson-Hsiao Estimation With Roots Near Unity," *Econometric Theory* (forthcoming).

Phillips, P. C. B., T. Leirvik & T. Storelvmo (2017). "Econometric Measurement of Earth's Transient Climate Sensitivity" Yale Working Paper.

Storelvmo, T., T. Leirvik, U. Lohmann, P. C. B. Phillips, and M. Wild (2016) "Disentangling Greenhouse Warming and Aerosol Cooling to Reveal Earth's Climate Sensitivity," *Nature Geoscience*, Vol. 9, No. 3, March, 2016, pp. 1-6.

8 Appendix A: Additional Figures

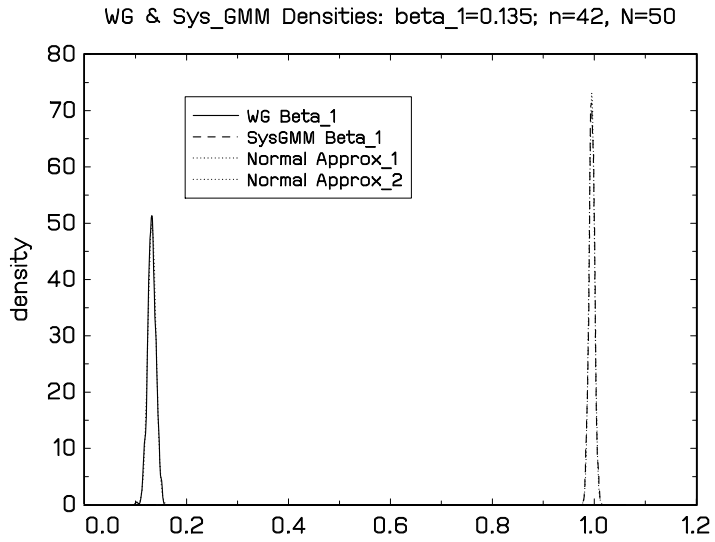


Figure 1: Kernel estimates of the densities of WG and sys-GMM estimates of β_1 based on $R = 1500$ replications with $n = 42, N = 50$, and true value $\beta_1 = 0.135$.

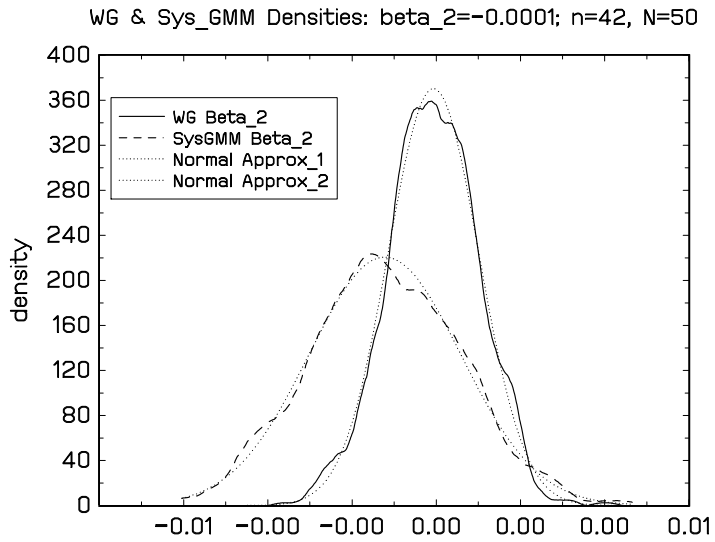


Figure 2: Kernel estimates of the densities of WG and sys-GMM estimates of β_2 based on $R = 1500$ replications with $n = 42, N = 50$, and true value $\beta_2 = -0.0001$.

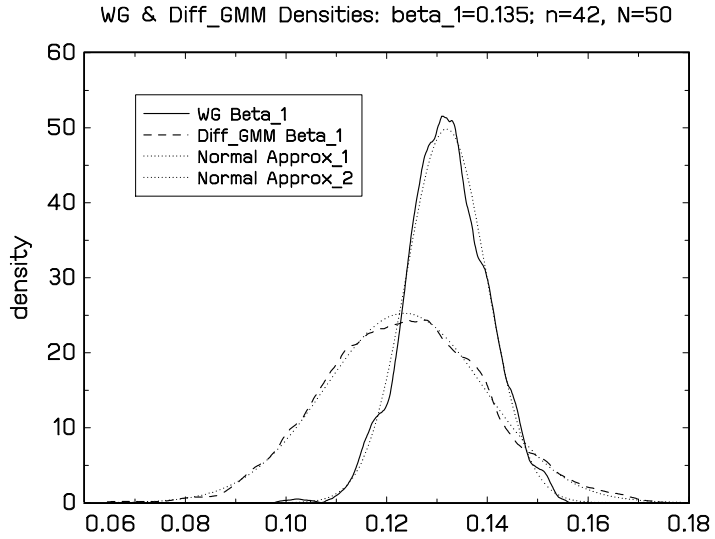


Figure 3: Kernel estimates of the densities of WG and diff-GMM estimates of β_1 based on $R = 1500$ replications with $n = 42, N = 50$, and true value $\beta_1 = 0.135$.

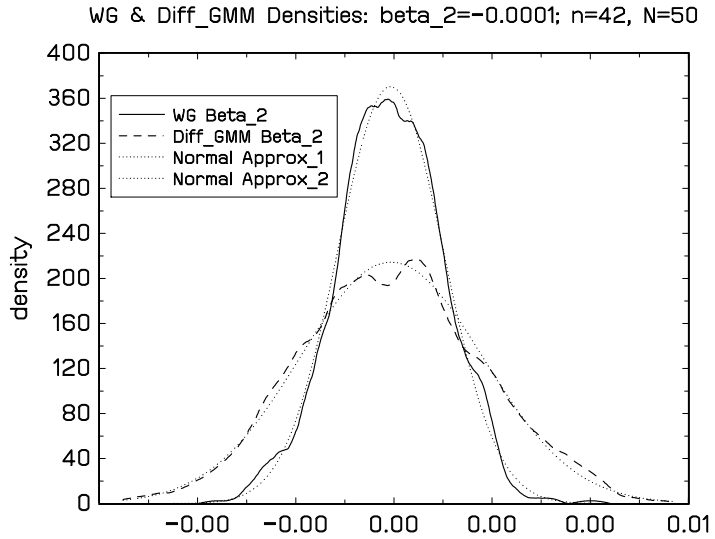


Figure 4: Kernel estimates of the densities of WG and diff-GMM estimates of β_2 based on $R = 1500$ replications with $n = 42, N = 50$, and true value $\beta_2 = -0.0001$.

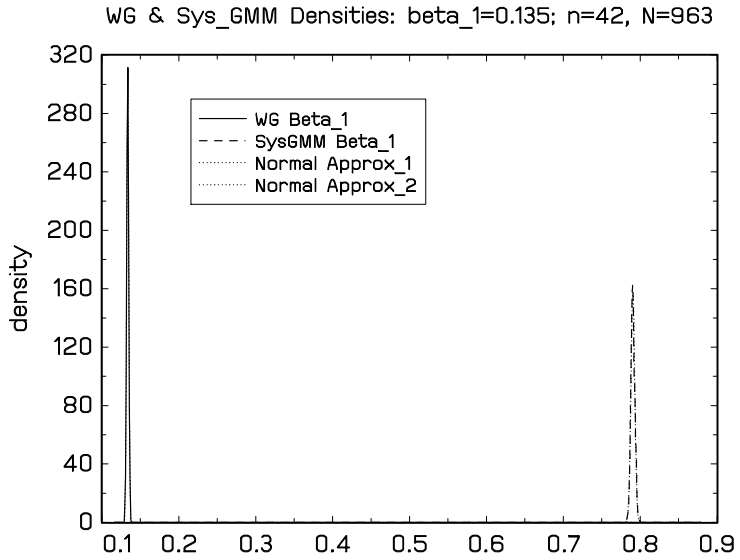


Figure 5: Kernel estimates of the densities of WG and sys-GMM estimates of β_1 based on $R = 1500$ replications with $n = 42, N = 963$, and true value $\beta_1 = 0.135$.

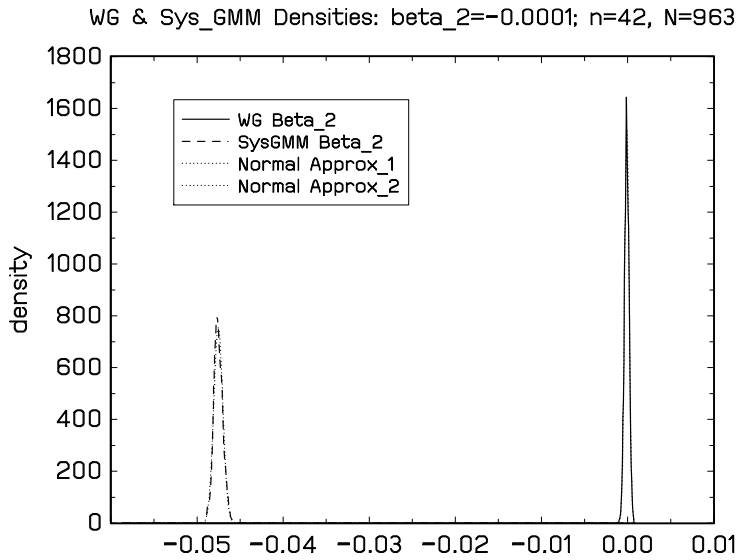


Figure 6: Kernel estimates of the densities of WG and sys-GMM estimates of β_2 based on $R = 1500$ replications with $n = 42, N = 50$, and true value $\beta_2 = -0.001$.

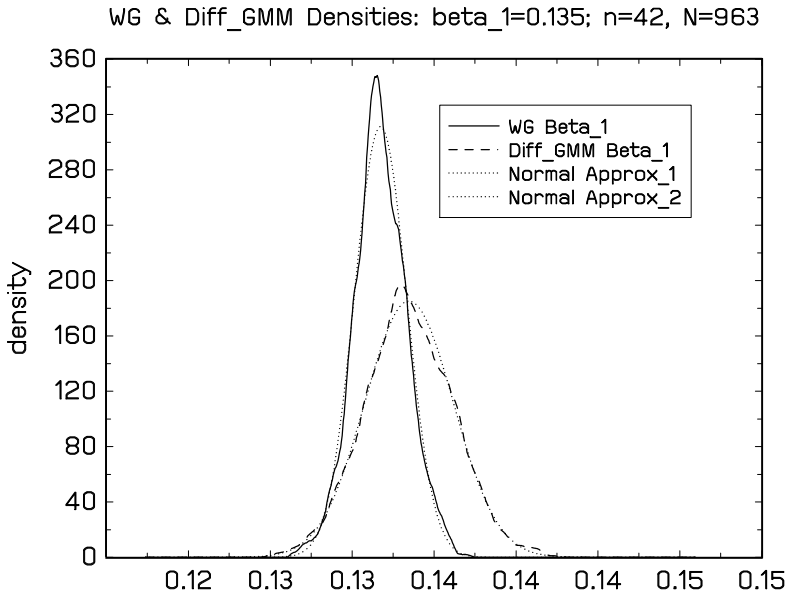


Figure 7: Kernel estimates of the densities of WG and diff-GMM estimates of β_1 based on $R = 1500$ replications with $n = 42, N = 963$, and true value $\beta_1 = 0.135$.

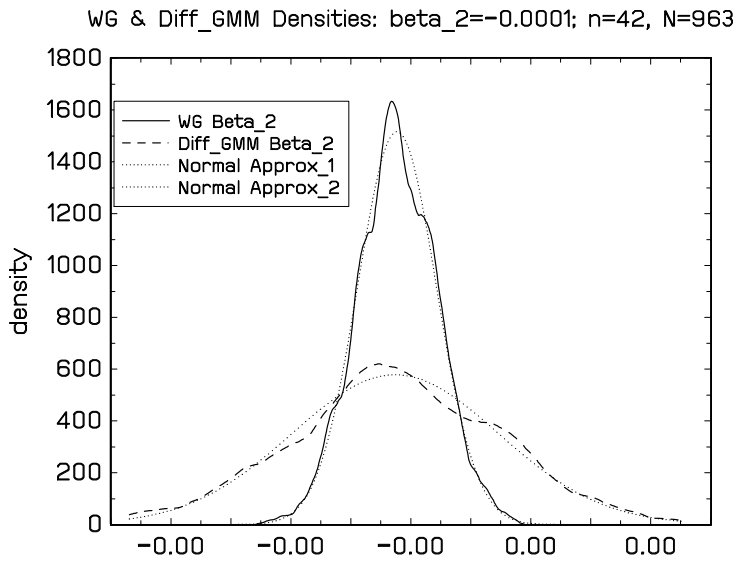


Figure 8: Kernel estimates of the densities of WG and diff-GMM estimates of β_2 based on $R = 1500$ replications with $n = 42, N = 963$, and true value $\beta_2 = -0.0001$.

9 Appendix B: Proofs

9.1 Proof of Theorem 1

The proof is immediate by simple algebraic manipulation, as shown in the remarks leading to (27) and (28). In what follows, we establish the estimation error (26), which is useful in the development of asymptotics. We use the implied form of the aggregate dynamic relation (19), viz.,

$$T_{t+1} = \gamma_0 + \theta' W_t + \bar{u}_{\cdot t+1}, \text{ with } \theta = (\theta_1, \theta_2, \gamma_3)', \text{ and } \theta_i = \beta_i + \gamma_i \text{ for } i = 1, 2 \tag{43}$$

which in matrix observation form is

$$T^+ = \gamma_0 \iota_{n-1} + W\theta + \bar{U} = (T_2, T_3, ..., T_n)' \tag{44}$$

where $T^+ = (T_2, T_3, ..., T_n)'$, $\bar{U} = [\bar{u}_{\cdot 2}, ..., \bar{u}_{\cdot n}]'$, and $\iota_{n-1} = (1, ..., 1)'$ is $(n-1) \times 1$. Using (28) we then have

$$\hat{\theta} = \begin{bmatrix} \hat{\beta}_1 + \hat{\gamma}_1 \\ \hat{\beta}_2 + \hat{\gamma}_2 \\ \hat{\gamma}_3 \end{bmatrix} = \left(\tilde{W}'\tilde{W}\right)^{-1} \left(\tilde{W}' T^+\right) = \theta + \left(\tilde{W}'\tilde{W}\right)^{-1} \left(\tilde{W}' \bar{U}^+\right), \tag{45}$$

which gives (26). We note that the time specific intercept λ_t in the regression is estimated by the regression residuals

$$\hat{\lambda}_t = T_{t+1} - \hat{\beta}' X_t = T_{t+1} - \hat{\beta}_1 T_t - \hat{\beta}_2 R_t = T_{t+1} - \left(\hat{\beta}', 0\right) W_t \tag{46}$$

using the identification condition that $\bar{\alpha} = 0$, as in Step 1 of the WG estimation. However, equation (46) applies not only for the WG estimate $\hat{\beta}$ but also when the panel regression equation (1) is estimated by diff-GMM, sys-GMM or other methods, in which case the residuals $\hat{\lambda}_t$ themselves depend on the method of estimation. Using the vector of these residuals $\hat{\lambda} = \left(\hat{\lambda}_t\right)$, the slope coefficients γ in equation (2) are estimated by least squares regression giving $\hat{\gamma} = \left(\tilde{W}'\tilde{W}\right)^{-1} \left(\tilde{W}'\hat{\lambda}\right)$, just as in the case of WG estimation. These coefficient estimates $\hat{\gamma}$, just as $\hat{\lambda}$, also depend on the method of estimation of the slope coefficients $\hat{\beta}$ in (1). Then, as in (27) we have

$$\hat{\gamma} = \left(\tilde{W}'\tilde{W}\right)^{-1} \tilde{W}' \left\{ T^+ - W \begin{bmatrix} \hat{\beta} \\ 0 \end{bmatrix} \right\} = \left(\tilde{W}'\tilde{W}\right)^{-1} \tilde{W}' T^+ - \begin{bmatrix} \hat{\beta} \\ 0 \end{bmatrix},$$

which reveals the compensatory adjustments between the estimated panel regression coefficients $\hat{\beta}$ and the estimated coefficients $\hat{\gamma}$ in the aggregate relation. Finally, transposing (45) gives

$$\hat{\theta} - \theta = \begin{bmatrix} \left(\hat{\beta}_1 + \hat{\gamma}_1\right) - (\beta_1 + \gamma_1) \\ \left(\hat{\beta}_2 + \hat{\gamma}_2\right) - (\beta_2 + \gamma_2) \\ (\hat{\gamma}_3 - \gamma_3) \end{bmatrix} = \left(\tilde{W}'\tilde{W}\right)^{-1} \tilde{W}' \bar{U}, \tag{47}$$

which shows that the estimates $\left(\hat{\beta}_1+\hat{\gamma}_1\right)$, $\left(\hat{\beta}_2+\hat{\gamma}_2\right)$, and $\hat{\gamma}_3$ of $(\beta_1+\gamma_1)$, $(\beta_2+\gamma_2)$, and γ_3 are each invariant to the choice of estimation procedure for the coefficients β in the panel regression (16). We deduce that the same is true for the implied estimate of the parameter TCS, viz.,

$$\widehat{TCS}=\frac{\hat{\gamma}_3}{1-\left(\hat{\beta}_1+\hat{\gamma}_1\right)}\times\ln(2)=:g\left(\hat{\theta}_1,\hat{\gamma}_3\right),\text{ with }\hat{\theta}_1=\hat{\beta}_1+\hat{\gamma}_1\text{ and }\theta_1=\beta_1+\gamma_1,$$

thereby establishing the stated invariance result. ■

9.2 Proof of Theorem 2

(i) Define the scaling matrix $D_n=\text{diag}\left(n^{3/2},nI_2\right)$ conformably with the rotation matrix $H=[h,H_\perp]$ given by (32). Then, by standard weak convergence methods, we have

$$\begin{aligned}
& D_n^{-1}H'\tilde{W}'\tilde{W}HD_n^{-1}=D_n^{-1}H'\sum_{t=1}^{n}\tilde{W}_t\tilde{W}_t'HD_n^{-1}\\
= & \begin{bmatrix} \frac{1}{n^3}\sum_{t=1}^{n}\left(a\tilde{t}+h'\tilde{V}_t+O_p(1)\right)^2 & \frac{1}{n^{5/2}}\sum_{t=1}^{n}\left(a\tilde{t}+h'\tilde{V}_t+O_p(1)\right)\left(V_{\perp t}'+O_p(1)\right)\\ \frac{1}{n^{5/2}}\sum_{t=1}^{n}\left(\tilde{V}_{\perp t}+O_p(1)\right)\left(a\tilde{t}+h'\tilde{V}_t+O_p(1)\right) & \frac{1}{n^2}\sum_{t=1}^{n}\left(\tilde{V}_{\perp t}+O_p(1)\right)\left(\tilde{V}_{\perp t}'+O_p(1)\right)\end{bmatrix}\\
\rightsquigarrow & \begin{bmatrix} a^2\int_0^1\tilde{r}^2 & a\int_0^1\tilde{r}\tilde{V}_\perp'\\ a\int_0^1\tilde{V}_\perp\tilde{r} & \int_0^1\tilde{V}_\perp\tilde{V}_\perp'\end{bmatrix},
\end{aligned}\tag{48}$$

where $a=\left(\delta_w'\delta_w\right)^{1/2}$. Inverting and by joint convergence and continuous mapping we have

$$D_nH'\left(\tilde{W}'\tilde{W}\right)^{-1}HD_n=D_n\left(H'\tilde{W}'\tilde{W}H\right)^{-1}D_n\rightsquigarrow\begin{bmatrix} a^2\int_0^1\tilde{r}^2 & a\int_0^1\tilde{r}\tilde{V}_\perp'\\ a\int_0^1\tilde{V}_\perp\tilde{r} & \int_0^1\tilde{V}_\perp\tilde{V}_\perp'\end{bmatrix}^{-1}.\tag{49}$$

it follows that

$$\begin{aligned}
& D_nH'\left(\hat{\theta}-\theta\right)=D_nH'\left(\tilde{W}'\tilde{W}\right)^{-1}\tilde{W}'\bar{U}=D_nH'\left(\tilde{W}'\tilde{W}\right)^{-1}HD_nD_n^{-1}H'\tilde{W}'\bar{U}\\
= & \left\{D_n\left(H'\tilde{W}'\tilde{W}H\right)^{-1}D_n\right\}\left\{D_n^{-1}H'\tilde{W}'\bar{U}\right\}\\
= & \left\{D_n\left(H'\tilde{W}'\tilde{W}H\right)^{-1}D_n\right\}\left\{\frac{1}{\sqrt{N}}\left[\sum_{t=1}^{n}D_n^{-1}H'\tilde{W}_t\xi_{t+1}+o_p(1)\right]\right\}.
\end{aligned}$$

Then

$$\sqrt{N}D_nH'\left(\hat{\theta}-\theta\right)\rightsquigarrow\begin{bmatrix} a^2\int_0^1\tilde{r}^2 & a\int_0^1\tilde{r}\tilde{V}_\perp'\\ a\int_0^1\tilde{V}_\perp\tilde{r} & \int_0^1\tilde{V}_\perp\tilde{V}_\perp'\end{bmatrix}^{-1}\begin{bmatrix} a\int_0^1\tilde{r}\tilde{V}hdB_\xi\\ \int_0^1\tilde{V}_\perp dB_\xi\end{bmatrix},\tag{50}$$

using the fact that

$$\sum_{t=1}^{n} D_n^{-1} H' \tilde{W}_t \xi_{t+1} = \begin{bmatrix} \frac{1}{n^{3/2}} \sum_{t=1}^{n} h' \tilde{W}_t \xi_{t+1} \\ \frac{1}{n} \sum_{t=1}^{n} H_\perp' \tilde{W}_t \xi_{t+1} \end{bmatrix} = \begin{bmatrix} \frac{1}{n^{3/2}} \sum_{t=1}^{n} \left(a\tilde{t} + h' \tilde{V}_t + O_p(1) \right) \xi_{t+1} \\ \frac{1}{n} \sum_{t=1}^{n} \left(\tilde{V}_{\perp t} + O_p(1) \right) \xi_{t+1} \end{bmatrix}$$
$$\rightsquigarrow \begin{bmatrix} a \int_0^1 \tilde{r} dB_\xi \\ \int_0^1 \tilde{V}_\perp dB_\xi \end{bmatrix}.$$

Thus

$$\begin{bmatrix} \sqrt{N} n^{3/2} h' \left(\hat{\theta} - \theta \right) \\ \sqrt{N} n H_\perp' \left(\hat{\theta} - \theta \right) \end{bmatrix} \rightsquigarrow \begin{bmatrix} a^2 \int_0^1 \tilde{r}^2 & a \int_0^1 \tilde{r} \tilde{V}_\perp' \\ a \int_0^1 \tilde{V}_\perp \tilde{r} & \int_0^1 \tilde{V}_\perp \tilde{V}_\perp' \end{bmatrix}^{-1} \begin{matrix} a \int_0^1 \tilde{r} dB_\xi \\ \int_0^1 \tilde{V}_\perp dB_\xi \end{matrix}. \quad (51)$$

The partitioned inverse in (50) can be written explicitly as follows. For notational convenience define the projection residuals

$$\tilde{r}_{\tilde{V}_\perp} = \tilde{r} - \left(\int_0^1 \tilde{r} \tilde{V}_\perp \right) \left(\int_0^1 \tilde{V}_\perp \tilde{V}_\perp' \right)^{-1} \tilde{V}_\perp,$$
$$\tilde{V}_{\perp,r}(r) = \tilde{V}_\perp(r) - \left(a \int_0^1 \tilde{V}_\perp \tilde{r} \right) \left(a^2 \int_0^1 \tilde{r}^2 \right)^{-1} a\tilde{r} = \tilde{V}_\perp(r) - \left(\int_0^1 \tilde{V}_\perp \tilde{r} \right) \left(\int_0^1 \tilde{r}^2 \right)^{-1} \tilde{r},$$

and then the inverse limit signal matrix has the following explicit form

$$\begin{bmatrix} a^2 \int_0^1 \tilde{r}^2 & a \int_0^1 \tilde{r} \tilde{V}_\perp' \\ a \int_0^1 \tilde{V}_\perp \tilde{r} & \int_0^1 \tilde{V}_\perp \tilde{V}_\perp' \end{bmatrix}^{-1} \quad (52)$$
$$= \begin{bmatrix} \left(a^2 \int_0^1 \tilde{r}_{\tilde{V}_\perp}^2 \right)^{-1} & -\left(a^2 \int_0^1 \tilde{r}_{\tilde{V}_\perp}^2 \right)^{-1} \left(a \int_0^1 \tilde{r} \tilde{V}_\perp' \right) \left(\int_0^1 \tilde{V}_\perp \tilde{V}_\perp' \right)^{-1} \\ -\left(\int_0^1 \tilde{V}_{\perp,r} \tilde{V}_{\perp,r}' \right)^{-1} \left(a \int_0^1 \tilde{V}_\perp \tilde{r} \right) \left(a^2 \int_0^1 \tilde{r}^2 \right)^{-1} & \left(\int_0^1 \tilde{V}_{\perp,r} \tilde{V}_{\perp,r}' \right)^{-1} \end{bmatrix}.$$

These results lead to the required limit theory for $n\sqrt{N}\left(\hat{\theta} - \theta\right)$. We use (50) and the decomposition

$$n\sqrt{N}\left(\hat{\theta} - \theta\right) = n\sqrt{N}\left[hh' + H_\perp H_\perp'\right]\left(\hat{\psi} - \psi\right) = h\left[\sqrt{N} n h' \left(\hat{\theta} - \theta\right)\right] + H_\perp \left[\sqrt{N} n H_\perp' \left(\hat{\theta} - \theta\right)\right]$$
$$= H_\perp \left[\sqrt{N} n H_\perp' \left(\hat{\theta} - \theta\right)\right] + o_p(1)$$
$$\rightsquigarrow H_\perp [0, I_2] \begin{bmatrix} a^2 \int_0^1 \tilde{r}^2 & a \int_0^1 \tilde{r} \tilde{V}' H_\perp \\ a \int_0^1 H_\perp' \tilde{V} \tilde{r} & \int_0^1 H_\perp' \tilde{V} \tilde{V}' H_\perp \end{bmatrix}^{-1} \begin{bmatrix} a \int_0^1 \tilde{r} dB_\xi \\ \int_0^1 \tilde{V}_\perp dB_\xi \end{bmatrix}$$
$$= H_\perp \begin{bmatrix} -\left(\int_0^1 \tilde{V}_{\perp,\tilde{r}} \tilde{V}_{\perp,\tilde{r}}' \right)^{-1} \left(a \int_0^1 \tilde{V}_\perp \tilde{r} \right) \left(a^2 \int_0^1 \tilde{r}^2 \right)^{-1} & \left(\int_0^1 \tilde{V}_{\perp,\tilde{r}} \tilde{V}_{\perp,\tilde{r}}' \right)^{-1} \end{bmatrix} \begin{bmatrix} a \int_0^1 \tilde{r} dB_\xi \\ \int_0^1 \tilde{V}_\perp dB_\xi \end{bmatrix}$$

$$
\begin{aligned}
&= H_{\perp}\left(\int_0^1 \tilde{V}_{\perp,\tilde{r}}\tilde{V}'_{\perp,\tilde{r}}\right)^{-1}\left\{\int_0^1 \tilde{V}_{\perp}dB_{\xi} - \left(\int_0^1 \tilde{V}_{\perp}\tilde{r}\right)\left(\int_0^1 \tilde{r}^2\right)^{-1}\int_0^1 \tilde{r}dB_{\xi}\right\} \\
&= H_{\perp}\left(\int_0^1 \tilde{V}_{\perp,\tilde{r}}\tilde{V}'_{\perp,\tilde{r}}\right)^{-1}\int_0^1 \tilde{V}_{\perp,\tilde{r}}dB_{\xi} \\
&\equiv \mathcal{MN}\left(0, \sigma_{\xi}^2 H_{\perp}\left(\int_0^1 \tilde{V}_{\perp,\tilde{r}}\tilde{V}'_{\perp,\tilde{r}}\right)^{-1} H'_{\perp}\right), \qquad (53)
\end{aligned}
$$

with $\tilde{V}_{\perp,\tilde{r}}(r) = \tilde{V}_{\perp}(r) - \left(\int_0^1 \tilde{V}_{\perp}\tilde{r}\right)\left(\int_0^1 \tilde{r}^2\right)^{-1}\tilde{r}$, the L_2 projection residual of $\tilde{V}_{\perp}$ on $\tilde{r}$. This result gives the limit theory for the vector $n\sqrt{N}\left(\hat{\theta}-\theta\right)$, and hence its individual elements, showing that the limit distribution is singular because of the presence of a multivariate deterministic time trend in the regressors.

(ii) The explicit inverse given in (52) also enables us to find the limit distribution of the coefficient estimates in the linear trend direction. In particular, we have from (51) and (52) that

$$
\begin{aligned}
&\sqrt{N}n^{3/2}h'\left(\hat{\theta}-\theta\right) \rightsquigarrow [1,0]\begin{bmatrix} a^2\int_0^1 \tilde{r}^2 & a\int_0^1 \tilde{r}\tilde{V}'_{\perp} \\ a\int_0^1 \tilde{V}_{\perp}\tilde{r} & \int_0^1 \tilde{V}_{\perp}\tilde{V}'_{\perp}\end{bmatrix}^{-1}\begin{bmatrix} a\int_0^1 \tilde{r}dB_{\xi} \\ \int_0^1 \tilde{V}_{\perp}dB_{\xi}\end{bmatrix} \\
= &\begin{bmatrix} \left(a^2\int_0^1 \tilde{r}^2_{\tilde{V}_{\perp}}\right)^{-1} & -\left(a^2\int_0^1 \tilde{r}^2_{\tilde{V}_{\perp}}\right)^{-1}\left(a\int_0^1 \tilde{r}\tilde{V}'_{\perp}\right)\left(\int_0^1 \tilde{V}_{\perp}\tilde{V}'_{\perp}\right)^{-1}\end{bmatrix}\begin{bmatrix} a\int_0^1 \tilde{r}dB_{\xi} \\ \int_0^1 \tilde{V}_{\perp}dB_{\xi}\end{bmatrix} \\
= &\left(a^2\int_0^1 \tilde{r}^2_{\tilde{V}_{\perp}}\right)^{-1}\left(a\int_0^1 \tilde{r}dB_{\xi}\right) - \left(a^2\int_0^1 \tilde{r}^2_{\tilde{V}_{\perp}}\right)^{-1}\left(a\int_0^1 \tilde{r}\tilde{V}'_{\perp}\right)\left(\int_0^1 \tilde{V}_{\perp}\tilde{V}'_{\perp}\right)^{-1}\int_0^1 \tilde{V}_{\perp}dB_{\xi} \\
= &\left(a^2\int_0^1 \tilde{r}^2_{\tilde{V}_{\perp}}\right)^{-1}\left(a\int_0^1\left[\tilde{r} - \left(\int_0^1 \tilde{r}\tilde{V}'_{\perp}\right)\left(\int_0^1 \tilde{V}_{\perp}\tilde{V}'_{\perp}\right)^{-1}\tilde{V}_{\perp}\right]dB_{\xi}\right) \\
= &\left(a\int_0^1 \tilde{r}^2_{\tilde{V}_{\perp}}\right)^{-1}\left(\int_0^1 \tilde{r}_{\tilde{V}_{\perp}}dB_{\xi}\right) \equiv \mathcal{MN}\left(0, \frac{\sigma_{\xi}^2}{a^2}\left(\int_0^1 \tilde{r}^2_{\tilde{V}_{\perp}}\right)^{-1}\right), \qquad (54)
\end{aligned}
$$

giving the stated result.

(iii) We next proceed to examine the TCS estimate $\widehat{TCS}$ and develop its asymptotic theory. Some care is needed in application of the usual delta method because of the singularity of the limit theory for $\hat{\theta}$ and its effects on the limit distribution of $\widehat{TCS}$. Set $\theta_a = (\theta_1, \theta_3) = (\theta_1, \gamma_3)$, write TCS and $\widehat{TCS}$ as

$$
TCS = g(\theta_a) = \frac{\theta_3}{1-\theta_1}\times \ln(2), \quad \widehat{TCS} = g\left(\hat{\theta}_a\right) = \frac{\hat{\theta}_3}{1-\hat{\theta}_1}\times \ln(2), \qquad (55)
$$

and define the gradient vector

$$
g_a(\theta_a) = \ln(2)\left(\frac{\theta_3}{(1-\theta_1)^2}, \frac{1}{1-\theta_1}\right)'. \qquad (56)
$$

Observe that the leading column of the orthogonal matrix H in (31) is

$$h = \left(\frac{\delta_{c1}\gamma_3}{1-\theta_1}, 0, \delta_{c1}\right)' \Big/ \left[\left(\frac{\delta_{c1}}{1-\theta_1}\right)^2 \left\{(1-\theta_1)^2 + \gamma_3^2\right\}\right]^{1/2}$$

and so

$$g_a' E_a = \ln(2) \left(\frac{\theta_3}{(1-\theta_1)^2}, 0, \frac{1}{1-\theta_1}\right). \tag{57}$$

Then

$$g_a' E_a H_\perp = g_a' \begin{bmatrix} 0 & -\delta_{c1} \\ 0 & \frac{\delta_{c1}\gamma_3}{1-\theta_1} \end{bmatrix} = \begin{bmatrix} 0 & -\delta_{c1}\frac{\theta_3}{(1-\theta_1)^2} + \frac{\delta_{c1}\gamma_3}{(1-\theta_1)^2} \end{bmatrix} = [0, 0], \tag{58}$$

since $\theta_3 = \gamma_3$. Since $\hat{\theta}_a - \theta_a = O_p\left(n^{-1}N^{-1/2}\right)$ by (53), it follows by application of the delta method and use of (53) and (58) that $n\sqrt{N} g_a' \left(\hat{\theta} - \theta\right) \to_p 0$ and, hence,

$$n\sqrt{N}\left(\widehat{TCS} - TCS\right) = n\sqrt{N} g_a' \left(\hat{\theta} - \theta\right) + o_p(1) = o_p(1).$$

The limit distribution of $\widehat{TCS}$ is then obtained by using the limit distribution of the coefficient estimates $\hat{\theta}$ in the linear trend direction. To do so, we proceed as follows. First note that

$$H = [h, H_\perp] = \frac{1}{\left(\delta_w'\delta_w\right)^{1/2}} \begin{bmatrix} \frac{\delta_{c1}\gamma_3}{1-\theta_1} & 0 & -\delta_{c1} \\ 0 & \left(\delta_w'\delta_w\right)^{1/2} & 0 \\ \delta_{c1} & 0 & \frac{\delta_{c1}\gamma_3}{1-\theta_1} \end{bmatrix},$$

with $\delta_w'\delta_w = \left(\frac{\delta_{c1}}{1-\theta_1}\right)^2 \left\{(1-\theta_1)^2 + \gamma_3^2\right\} =: a^2$. Write the product

$$g_a' E_a = \ln(2) \left(\frac{\theta_3}{(1-\theta_1)^2}, 0, \frac{1}{1-\theta_1}\right) = \frac{\left(\delta_w'\delta_w\right)^{1/2}}{\delta_{c1}(1-\theta_1)} h'. \tag{59}$$

Proceeding in the same way as (54), it follows that

$$\begin{aligned} & n^{3/2}\sqrt{N}\left(\widehat{TCS} - TCS\right) = n\sqrt{N} g_a' \left(\hat{\theta} - \theta\right) + o_p(1) = \frac{\left(\delta_w'\delta_w\right)^{1/2}}{\delta_{c1}(1-\theta_1)} n^{3/2}\sqrt{N} h' \left(\hat{\theta} - \theta\right) \\ \rightsquigarrow\ & \frac{\left(\delta_w'\delta_w\right)^{1/2}}{\delta_{c1}(1-\theta_1)} [1, 0] \begin{bmatrix} a^2 \int_0^1 \tilde{r}^2 & a\int_0^1 \tilde{r}\tilde{V}_\perp' \\ a\int_0^1 \tilde{V}_\perp \tilde{r} & \int_0^1 \tilde{V}_\perp \tilde{V}_\perp' \end{bmatrix}^{-1} \begin{bmatrix} a\int_0^1 \tilde{r} dB_\xi \\ \int_0^1 \tilde{V}_\perp dB_\xi \end{bmatrix} \\ =\ & \frac{\left(\delta_w'\delta_w\right)^{1/2}}{\delta_{c1}(1-\theta_1)} \left(a\int_0^1 \tilde{r}_{\tilde{V}_\perp}^2\right)^{-1} \left(\int_0^1 \tilde{r}_{\tilde{V}_\perp} dB_\xi\right) \equiv \frac{\left(\delta_w'\delta_w\right)^{1/2}}{\delta_{c1}(1-\theta_1)} \mathcal{MN}\left(0, \frac{\sigma_\xi^2}{a^2}\left(\int_0^1 \tilde{r}_{\tilde{V}_\perp}^2\right)^{-1}\right) \\ =\ & \mathcal{MN}\left(0, \frac{\sigma_\xi^2}{\delta_{c1}^2(1-\theta_1)^2}\left(\int_0^1 \tilde{r}_{\tilde{V}_\perp}^2\right)^{-1}\right), \end{aligned} \tag{60}$$

as required.

■

9.3 Estimating the Asymptotic Variance Matrix of $\hat{\theta}$

The asymptotic variance matrix of $\hat{\theta}$ may be estimated in the usual way. To show this, note that standard partitioned matrix inversion gives

$$
\begin{aligned}
& n^2\left(\tilde{W}'\tilde{W}\right)^{-1} = n^2\left(HH'\tilde{W}'\tilde{W}HH'\right)^{-1} = n^2H\left(H'\tilde{W}'\tilde{W}H\right)^{-1}H' \\
= \;& n^2H\begin{bmatrix} a^2\sum_{t=1}^n \tilde{t}^2 & a\sum_{t=1}^n \tilde{t}\tilde{V}_t'H_\perp \\ a\sum_{t=1}^n H_\perp'\tilde{V}_t\tilde{t} & \sum_{t=1}^n H_\perp'\tilde{V}_t\tilde{V}_t'H_\perp \end{bmatrix}^{-1} H'\left\{1+o_p\left(1\right)\right\} \\
= \;& :n^2H\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & A_{22} \end{bmatrix}^{-1} H'\left\{1+o_p\left(1\right)\right\} \qquad (61) \\
= \;& n^2H\begin{bmatrix} a_{11.2}^{-1} & -a_{11.2}^{-1}a_{12}A_{22}^{-1} \\ -A_{22.1}^{-1}a_{21}a_{11}^{-1} & A_{22.1}^{-1} \end{bmatrix} H'\left\{1+o_p\left(1\right)\right\} \\
= \;& H\begin{bmatrix} O_p\left(n^{-1}\right) & O_p\left(n^{-1/2}\right) \\ O_p\left(n^{-1/2}\right) & \left(\frac{A_{22.1}}{n^2}\right)^{-1} \end{bmatrix} H'\left\{1+o_p\left(1\right)\right\} \\
\rightsquigarrow \;& H\begin{bmatrix} 0 & 0 \\ 0 & \left\{\int_0^1 H_\perp'\tilde{V}\tilde{V}'H_\perp - \left(\int_0^1 H_\perp'\tilde{V}\tilde{r}\right)\left(\int_0^1 \tilde{r}^2\right)^{-1}\left(\int_0^1 \tilde{r}\tilde{V}'H_\perp\right)\right\}^{-1} \end{bmatrix} H' \\
= \;& H_\perp\left(\int_0^1 \tilde{V}_{\perp\tilde{r}}\tilde{V}_{\perp\tilde{r}}'\right)^{-1}H_\perp'
\end{aligned}
$$

because

$$
\begin{aligned}
n^2a_{11.2}^{-1} &= \frac{n^2}{n^3}\left\{\frac{a^2\sum_{t=1}^n \tilde{t}^2}{n^3} - \left(\frac{a\sum_{t=1}^n \tilde{t}\tilde{V}_t'H_\perp}{n^{5/2}}\right)\left(\frac{\sum_{t=1}^n H_\perp'\tilde{V}_t\tilde{V}_t'H_\perp}{n^2}\right)^{-1}\left(\frac{a\sum_{t=1}^n H_\perp'\tilde{V}_t\tilde{t}}{n^{5/2}}\right)\right\}^{-1} \\
&= O_p\left(n^{-1}\right), \\
n^2A_{22.1}^{-1} &= \left\{\frac{1}{n^2}\sum_{t=1}^n H_\perp'\tilde{V}_t\tilde{V}_t'H_\perp - \left(\frac{a\sum_{t=1}^n H_\perp'\tilde{V}_t\tilde{t}}{n^{5/2}}\right)\left(\frac{a^2\sum_{t=1}^n \tilde{t}^2}{n^3}\right)^{-1}\left(\frac{a\sum_{t=1}^n \tilde{t}\tilde{V}_t'H_\perp}{n^{5/2}}\right)\right\}^{-1}, \\
&\rightsquigarrow \left\{\int_0^1 H_\perp'\tilde{V}\tilde{V}'H_\perp - \left(\int_0^1 H_\perp'\tilde{V}\tilde{r}\right)\left(\int_0^1 \tilde{r}^2\right)^{-1}\left(\int_0^1 \tilde{r}\tilde{V}'H_\perp\right)\right\}^{-1} = \left(\int_0^1 \tilde{V}_{\perp\tilde{r}}\tilde{V}_{\perp\tilde{r}}'\right)^{-1}, \\
n^2A_{22.1}^{-1}a_{21}a_{11}^{-1} &= \frac{n^2}{n^{5/2}}\left\{\left(\frac{A_{22.1}}{n^2}\right)^{-1}\frac{a_{21}}{n^{5/2}}\left(\frac{a_{11}}{n^3}\right)^{-1}\right\} = O_p\left(n^{-1'2}\right).
\end{aligned}
$$

Next the (cross section asymptotic) panel regression error variance σ_ξ^2 is to be estimated. Under Assumption A(i) $\sigma_\xi^2 = \sigma_u^2$ and σ_u^2 may be estimated from the residual of the combined panel regression (24), viz

$$
\hat{\sigma}_u^2 = \frac{1}{Nn}\sum_{i=1}^n\sum_{t=1}^{n-1}\hat{u}_{it+1}^2 = \frac{1}{Nn}\sum_{i=1}^n\sum_{t=1}^{n-1}\left(\tilde{T}_{i,t+1} - \hat{\beta}_1\tilde{T}_{i,t} - \hat{\beta}_2\tilde{R}_{i,t} - \hat{\gamma}'\tilde{W}_t\right)^2,
$$

which is consistent for σ_u^2 under Assumption A, where

$$\hat{u}_{it+1} = T_{i,t+1} - \hat{\alpha}_i - \hat{\beta}_1 T_{i,t} - \hat{\beta}_2 R_{i,t} - \hat{\gamma}_0 - \hat{\gamma}' W_t = \tilde{T}_{i,t+1} - \hat{\beta}_1 \tilde{T}_{i,t} - \hat{\beta}_2 \tilde{R}_{i,t} - \hat{\gamma}' \tilde{W}_t.$$

With these results in hand, we can construct the following consistent estimate of the conditional variance matrix of $n\sqrt{N}\left(\hat{\theta}-\theta\right)$ in Theorem 2, viz.,

$$n^2 N \left(\frac{\hat{\sigma}_u^2}{N}\right)\left(\tilde{W}'\tilde{W}\right)^{-1} = \hat{\sigma}_u^2 n^2 \left(\tilde{W}'\tilde{W}\right)^{-1} \rightsquigarrow \sigma_\xi^2 H_\perp \left(\int_0^1 \tilde{V}_{\perp,\tilde{r}} \tilde{V}'_{\perp,\tilde{r}}\right)^{-1} H'_\perp, \tag{62}$$

So, the asymptotic variance is given by the usual formula $\hat{\sigma}_u^2\left(\tilde{W}'\tilde{W}\right)^{-1}$. Note that the effective sample size scaling involved in (62) is n^2, corresponding to the presence of stochastic trends in the signal matrix $\tilde{W}'\tilde{W}$. The scaling by $\sqrt{N}$ in the standardized estimation error $n\sqrt{N}\left(\hat{\theta}-\theta\right)$ arises because of the estimation error $\hat{\theta}-\theta = \left(\tilde{W}'\tilde{W}\right)^{-1}\left(\tilde{W}'\bar{U}\right)$ from (30), and the moment matrix $\tilde{W}'\bar{U} = \sum_{t=1}^n \tilde{W}_t \bar{u}_{.t+1}$ involves the cross section sample mean $\bar{u}_{.t+1} = N^{-1}\sum_{i=1}^N u_{i,t+1}$ whose variance is σ_u^2/N, so cross section sample size scaling is already implicitly incorporated in $\frac{\hat{\sigma}_u^2}{N}\left(\tilde{W}'\tilde{W}\right)^{-1}$ and the estimated variance matrix of $n\sqrt{N}\left(\hat{\theta}-\theta\right)$ is then $\hat{\sigma}_u^2 n^2\left(\tilde{W}'\tilde{W}\right)^{-1} \rightsquigarrow \sigma_\xi^2 H_\perp \left(\int_0^1 \tilde{V}_{\perp\tilde{r}}\tilde{V}'_{\perp\tilde{r}}\right)^{-1} H'_\perp$, as required.

Proceeding in a related way we can estimate the conditional variance of the estimate $\widehat{TCS}$ of the TCS parameter and, using this, a $100\,(1-\alpha)\,\%$ confidence interval for TCS. Using the same notation as in (61) and the definitions

$$\begin{aligned} g'_a &= g'_a(\theta_1,\theta_3) = \ln(2)\left(\frac{\theta_3}{(1-\theta_1)^2}, \frac{1}{1-\theta_1}\right), \\ E_a &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \end{aligned}$$

so that $g'_a E_a = \frac{\left(\delta'_w \delta_w\right)^{1/2}}{\delta_{c1}(1-\theta_1)} h'$, as in (59), we obtain

$$\begin{aligned} & n^3 g'_a E_a \left(\tilde{W}'\tilde{W}\right)^{-1} E'_a g_a = g'_a E_a H \left(H'\tilde{W}'\tilde{W}H\right)^{-1} H' E'_a g_a \\ = \;& \frac{\delta'_w \delta_w}{\delta_{c1}^2 (1-\theta_1)^2} n^3 [1,0] \begin{bmatrix} a_{11.2}^{-1} & -a_{11.2}^{-1} a_{12} A_{22}^{-1} \\ -A_{22.1}^{-1} a_{21} a_{11}^{-1} & A_{22.1}^{-1} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \{1 + o_p(1)\} \\ = \;& \frac{\delta'_w \delta_w}{\delta_{c1}^2 (1-\theta_1)^2} \left(\frac{a_{11.2}}{n^3}\right)^{-1} \{1 + o_p(1)\} \\ \rightsquigarrow \;& \frac{\delta'_w \delta_w}{\delta_{c1}^2 (1-\theta_1)^2} \left(a^2 \int_0^1 \tilde{r}^2_{\tilde{V}_\perp}\right)^{-1} = \frac{1}{\delta_{c1}^2 (1-\theta_1)^2}\left(\int_0^1 \tilde{r}^2_{\tilde{V}_\perp}\right)^{-1}, \end{aligned}$$

since

$$\frac{a_{11.2}}{n^3} = \frac{a^2 \sum_{t=1}^n \tilde{t}^2}{n^3} - \left(\frac{a \sum_{t=1}^n \tilde{t} \tilde{V}_t' H_\perp}{n^{5/2}} \right) \left(\frac{\sum_{t=1}^n H_\perp' \tilde{V}_t \tilde{V}_t' H_\perp}{n^2} \right)^{-1} \left(\frac{a \sum_{t=1}^n H_\perp' \tilde{V}_t \tilde{t}}{n^{5/2}} \right) \rightsquigarrow a^2 \int_0^1 \tilde{r}_{\tilde{V}_\perp}^2,$$

and $a^2 = \delta_w' \delta_w$. Thus,

$$n^3 \hat{\sigma}_u^2 g_a' E_a \left(\tilde{W}' \tilde{W} \right)^{-1} E_a' g_a \rightsquigarrow \frac{\sigma_u^2}{\delta_{c1}^2 (1-\theta_1)^2} \left(\int_0^1 \tilde{r}_{\tilde{V}_\perp}^2 \right)^{-1}.$$

Next, since $\hat{\theta}$ and $\hat{\sigma}_u^2$ are consistent for θ and σ_u^2, we have

$$\hat{g}_a' = g_a' \left(\hat{\theta}_1, \hat{\theta}_3 \right) = \ln(2) \left(\frac{\hat{\theta}_3}{\left(1 - \hat{\theta}_1\right)^2}, \frac{1}{1 - \hat{\theta}_1} \right) \rightarrow_p g_a(\theta_1, \theta_2),$$

and

$$n^3 \hat{\sigma}_u^2 \hat{g}_a' E_a \left(\tilde{W}' \tilde{W} \right)^{-1} E_a' \hat{g}_a \rightsquigarrow \frac{\sigma_u^2}{\delta_{c1}^2 (1-\theta_1)^2} \left(\int_0^1 \tilde{r}_{\tilde{V}_\perp}^2 \right)^{-1},$$

giving a consistent estimate of the asymptotic conditional covariance matrix (60) of the limit distribution of $n^{3/2}\sqrt{N}\left(\widehat{TCS} - TCS\right)$. It follows that a $100(1-\alpha)\,\%$ confidence interval for TCS may be constructed as

$$\widehat{TCS} \pm z_\alpha \left\{ \hat{\sigma}_u^2 \hat{g}_a' E_a \left(\tilde{W}' \tilde{W} \right)^{-1} E_a' \hat{g}_a \right\}^{1/2}.$$

■