Published in Journal of Econometrics, 2019 January
https://doi.org/10.1016/j.jeconom.2019.05.002

Econometric Measurement of Earth's Transient Climate Sensitivity*

Peter C. B. Phillips†, Thomas Leirvik‡, Trude Storelvmo§

February 15, 2017

Abstract

How sensitive is Earth's climate to a given increase in atmospheric greenhouse gas (GHG) concentrations? This long-standing and fundamental question in climate science was recently analyzed by dynamic panel data methods using extensive spatio-temporal data of global surface temperatures, solar radiation, and GHG concentrations over the last half century to 2010 (Storelvmo et al, 2016). These methods revealed that atmospheric aerosol effects masked approximately one-third of the continental warming due to increasing GHG concentrations over this period, thereby implying greater climate sensitivity to GHGs than previously thought. The present study provides asymptotic theory justifying the use of these methods when there are stochastic process trends in both the global forcing variables, such as GHGs, and station-level trend effects from such sources as local aerosol pollutants. These asymptotics validate confidence interval construction for econometric measures of Earth's transient climate sensitivity. The methods are applied to observational data and to data generated from three leading global climate models (GCMs) that are sampled spatio-temporally in the same way as the empirical observations. The findings indicate that estimates of transient climate sensitivity produced by these GCMs lie within empirically determined confidence limits but that the GCMs uniformly underestimate the effects of aerosol induced dimming effects. The analysis shows the potential of econometric methods to calibrate GCM performance against observational data and to reveal the respective sensitivity parameters (GHG and non-GHG related) governing GCM temperature trends.

Keywords: Climate sensitivity, Cointegration, Common stochastic trend, Idiosyncratic trend, Spatio-temporal model, Unit root.

JEL Classification: C32, C33

---

*Phillips acknowledges support from the NSF under Grant SES 1258258.
†Yale University, University of Auckland, University of Southampton, Singapore Management University.
‡Nord University
§Yale University

1 Introduction

Global warming is one of the defining issues of our time, and currently affects lives, communities and countries worldwide. Its well-established root cause is the steady climb of atmospheric CO_2, which is now at 50% above pre-industrial levels. Understanding exactly how sensitive Earth's climate is to CO_2 emissions is critically important for efforts to mitigate and adapt to future climate change. Despite this, Earth's climate sensitivity, i.e. the global mean surface temperature increase for a given atmospheric CO_2 increase, remains an elusive quantity, and arguably has come to represent the "holy grail" of climate science. The lack of progress on this issue can partly be attributed to the difficulty of measuring the sensitivity of climate to CO_2 from observational data. Such efforts have been hampered by the fact that aerosol particles, which have a cooling effect on climate, have been increasing along with CO_2, and are therefore "masking" some unknown proportion of CO_2-induced warming to date (e.g., Andreae et al., 2005). Representing the cooling effect of aerosol particles in global climate models (GCMs) has proven notoriously challenging, and GCM estimates of aerosol cooling continue to diverge. Novel and alternative approaches that can assist in meeting this challenge are long overdue.

Realizing that insights from econometrics could be of value in resolving this problem and following earlier modeling work by Magnus et al. (2011), Storelvmo et al. (2016) applied dynamic panel data methods to a rich observational data set of climate variables, and found that $\sim$1/3 of the CO_2 warming of continents to date has likely been masked by aerosol cooling. Studies not accounting for this cooling would falsely conclude that climate is less sensitive to CO_2 than it really is. By taking aerosol cooling into account the Storelvmo et al. study supported climate sensitivities at the upper end of the range already published, for example in the last report from the Intergovernmental Panel on Climate Change (IPCC, Flato et al., 2013).

The Magnus et al. and Storelvmo et al. studies pioneered in applying dynamic panel data methods with observational data to the problem of constrained climate sensitivity. While we are confident that this econometric approach holds great promise for climate studies and is well worth pursuing, we acknowledge that in order to arrive at inferences concerning climate sensitivity using these econometric methods a number of new assumptions and model specification enhancements are needed to adequately account for features in the observed data. The reliability of the climate sensitivity estimate depends on the validity of these assumptions and the suitability of the inferential methodology. Given the complexity of the dynamic panel generating mechanism and the presence of potentially multiple sources of stochastic trends, econometric analysis requires a full development of asymptotic theory of estimation and inference in the presence of such trends whilst allowing for variable co-movement governed by energy balance considerations.

The present paper contributes by addressing these issues. Specifically, we build on our previous study in the following ways: (i) the model in Storelvmo et al. (2016) is extended by provision of an explicit generating mechanism that accommodates stochastic nonstationarity in the data; (ii) asymptotic theory is developed for estimation and inference in the context of this expanded model that refines the method by which we calculate climate sensitivity and its associated confidence interval; and (iii) the refined methodology is applied to both the observational data and the numerical data simulated by three leading GCMs. The developments in (ii) are novel in econometrics because they allow treatment of nonstationarity with cointegrated regressors (with associated signal matrix degeneracies) at both the individual station level data and the global aggregate level. The application in (iii) innovates not only by analyzing GCM simulated data by econometric methods but also by carefully matching GCM-simulated data at times and spatial locations for which observational data are available. This matching serves as a powerful test of the fidelity of the method because the calculated climate sensitivity manifested in the GCMs (as opposed to the real climate system) can be compared to reported values available in the latest IPCC report (Flato et al., 2013). In addition to these contributions, provision of this new econometric analysis of GCM output enables us to identify GCM model shortcomings which have not become apparent in standard GCM validation exercises.

Section 2 provides a brief introduction to the panel econometric framework for modeling key climate variables observable over time at specific station locations. The model is extended in Section 3 to accommodate stochastic driver variables that include both global forcing variables and station-specific aerosol pollution trends. Some econometric implications of the expanded model are explored in Section 4, including the cointegrating structures that arise from energy balance considerations at the individual station and global levels. Asymptotic theory for the panel regression parameter estimates and energy balance parameter estimates is developed in Section 5. These developments enable us to arrive ultimately at econometric estimates for global climate sensitivity and an asymptotically valid confidence interval for this composite parameter based on the parameter estimates emerging from the dynamic panel data analysis (Sections 5 - 6). Finally, Section 7 reports an empirical application of the new methodology to the same observational data used previously in Storelvmo et al. (2016) and to matching simulation output from three leading GCM models. Summary conclusions are given in Section 8. Relevant technical material, proofs of results, and some further discussion are in the Appendix.

2 A Climate Econometric Model

The econometric model used in Storelvmo et al (2016) relates local temperature (T_i) at time $t+1$ to local temperature and surface radiation (R_i), as well as global factors (λ_t, see below), all at time t. The base model was developed and used in Magnus et al. (2011) and has the following two equations

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

where

$$\lambda_t = \gamma_0 + \gamma_1 \overline{T}_t + \gamma_2 \overline{R}_t + \gamma_3 \ln(CO_{2,t}), \tag{2}$$

relates the spatial aggregate variables $\left(\overline{T}_t, \overline{R}_t\right) = \left(N^{-1}\sum_{i=1}^N T_{it}, N^{-1}\sum_{i=1}^N R_{it}\right)$ and the logarithm of the CO_2 equivalent series, $\ln(CO_{2,t})$. In what follows we extend this model to accommodate stochastic forcing variables at both the station-specific and global levels.

A major focal point of our analysis is the ultimate measurement of transient climate sensitivity (TCS). TCS is defined as the expected global temperature after a doubling of CO_2. In the context of the above model, TCS is computed by the following expression, which is derived in the Appendix (see equation (48))

$$TCS = \frac{\gamma_3}{1-\beta_1-\gamma_1} \times \ln(2) =: f(\beta_1, \gamma_1, \gamma_3). \tag{3}$$

Let $\beta = (\beta_1, \beta_2)'$ and $\gamma = (\gamma_1, \gamma_2, \gamma_3)'$. To find an asymptotically valid confidence interval for the function φ using parameter estimates $\hat{\theta} = \left(\hat{\beta}', \hat{\gamma}'\right)'$ of $\theta = (\beta', \gamma')'$ obtained by regression we can use the asymptotic distribution of $\hat{\theta}$.

A complication in the analysis and the asymptotic development is that the variables in the model (1) and (2) have different stochastic properties and orders of magnitude, thereby complicating the asymptotic theory of $\hat{\theta}$. In particular, in deriving the limit distribution of $\hat{\theta}$ and using delta method derivations to analyze meaningful parameters such as TCS , we need to take into account the fact that (2) is a cointegrating relation among stochastically nonstationary time series, whereas $\hat{\beta}$ are estimated coefficients in a panel regression (1) that is transient if $|\beta_1| < 1$ but which nonetheless involves stochastically trending data in T_{it} and R_{it}. The analysis therefore requires account of the fact that the covariates in (1) are not all stationary and that persistent local and global shocks affect these covariates. The asymptotic development presents corresponding challenges and, as will become clear, these do not fall neatly within existing results in econometrics for time series and panel regressions involving stochastically nonstationary variables. The paper therefore develops the model to accommodate these features and provides asymptotic theory for coefficient estimates that allows for inference about the quantity TCS within this expanded framework.

To proceed we complete the model in a way that clarifies the relationship between the transient equation variables (T_{it}, R_{it}) and the global variables $\left(\overline{T}_t, \overline{R}_t, \ln(CO_{2,t})\right)$. Magnus et al. (2011) use station level data, aggregating and averaging the station data to obtain $\left(\overline{T}_t, \overline{R}_t\right)$. Neither that paper nor Storelvmo et al. (2016) provided a complete model capturing the linkages of the station level data to the equilibrium energy balance in a way that accommodates potential stochastic nonstationarity in the variables and additional forcing variables at both station and global levels. In what follows, therefore, we develop the model so that the linkages are explicit, clarifying the stochastic orders of the various components at the station level and the aggregate level. The limit distribution theory for the panel regression estimates can then be established. This limit theory enables us to obtain an asymptotically valid confidence interval for TCS which in turn facilitates inference about climate sensitivity to GHG emissions.

3 Extensions for Local and Global Forcing

To complete the specification of (1) we prescribe the generating mechanism of local radiation effects R_{it}. We can reasonably assume that R_{it} has both stationary and nonstationary components, which combine linearly to produce the total station level downwelling radiation as

$$R_{it} = R_{it}^0 + \delta_{ri}' G_t + P_{it}. \tag{4}$$

In this specification R_{it}^0 is a stationary component of local radiation that characterizes stationary fluctuations about some fixed mean level $\mathbb{E}\left(R_{it}^0\right)$. The component

$$G_t = G_0 + \sum_{s=1}^{t} u_{gs} =: G_0 + U_{gt} \tag{5}$$

in (4) is an m_g- vector of global forcing variables that are stochastically nonstationary with global shocks u_{gt} and δ_{ri} is an idioscyncratic factor loading parameter vector that captures the station level effect of the common global shock G_t, measured as the idiosyncratic proportion (δ_{ir}) of the full global effect, giving the term $\delta_{ri}' G_t$ (or simply $\delta_{ri} G_t$ if G_t is a scalar variable of global effects) in (4). This formulation means that there are nonstationary latent global forcing variables that affect the local system radiation variable R_{it}. The term P_{it} represents any remaining local idiosyncratic trend effects (such as those caused by station-specific aerosol pollution trends) that may be present in R_{it} which differ in source and character from the global common shock G_t.

The specification (4) therefore encapsulates stationary fluctuations (R_{it}^0), station effects of common global shocks ($\delta_{ri}' G_t$), and station-specific trends (P_{it}) that may be present in downwelling observed radiation. Both G_t and P_{it} can be considered latent variables within R_{it} and, consequently, T_{it+1} in (1). To fix ideas and proceed with an asymptotic development, we make the following assumption about the components in (1), (2) and (4)

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* $\{\delta_{ri}\} \sim_{iid} (\delta_r, \Sigma_r)$ *are independent of* $\left\{u_{it}^P\right\}, \{u_{ct}\}$ *for all* (i,t), *and the* $\left\{u_{it}^P\right\}$ *are defined in A(iii) and the* $\{u_{ct}\}$ *in Assumption C(ii).*

(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* $u_{ik}^{P} \sim_{iid} \left(0, \sigma_p^2\right)$ *over* i, $\bar{P}_0 = N^{-1} \sum_{i=1}^{N} P_{i0} \to_{a.s.} P^0 = \lim_{N\to\infty} \left\{ N^{-1} \sum_{i=1}^{N} \mathbb{E}\left(P_{it}^{0}\right) \right\}$, *and the partial sums* U_{it}^{P} *satisfy the invariance principle* $n^{-1/2} U_{it}^{P} \Rightarrow U_{i}^{P}(r) \equiv BM\left(\omega_{uP}^2\right)$ *for all* i *and with* $\omega_{uP}^2 = \sum_{h=-\infty}^{\infty} \gamma_{uP}(h) > 0$ *where* $\mathbb{E}\left(u_{it}^{P} u_{it+h}^{P}\right) = \gamma_{uP}(h)$ *for all* i *and* $K^{\varepsilon} \sum_{h=K}^{\infty} |\gamma_{uP}(h)| = o(1)$ *for some* $\varepsilon > 0$ *as* $K \to \infty$.

(iv) $\frac{1}{n} + \frac{1}{N} + \frac{n}{N} \to 0$.

Assumption B

(i) $|\beta_1| < 1$, $|\beta_1 + \gamma_1| < 1$, *and* λ_t *is an asymptotically stationary equilibrium error.*

(ii) $\{u_{gt}\}$ *has partial sums* $U_{gt} = \sum_{k=1}^{t} u_{gk}$ *that satisfy the invariance principle* $n^{-1/2} U_{g\lfloor nr \rfloor} \Rightarrow U_g(r) \equiv BM(\Omega_g)$, *Brownian motion with covariance matrix* $\Omega_g > 0$.

Conditions A(i) and B(i) imply that station level temperature effects involve transient adjustments to local radiation $R_{i,t}$, global influences imported via λ_t, and the panel system errors u_{it}. Upon station averaging of (4), we obtain

$$\overline{R}_t = \bar{R}_t^0 + \bar{\delta}_{rN}' G_t + \bar{P}_t,$$

with $\bar{R}_t^0 = N^{-1} \sum_{i=1}^{N} R_{it}^0$, $\bar{P}_t = N^{-1} \sum_{i=1}^{N} P_{it}$ and $\bar{\delta}_{rN} = N^{-1} \sum_{i=1}^{N} \delta_{ri}$. Under A(i) the loading factors δ_{ri} obey a strong law so that $\bar{\delta}_{rN} = N^{-1} \sum_{i=1}^{N} \delta_{ri} \to_{a.s.} \delta_r$, and similarly by A(ii) and A(iii) $\left(\bar{R}_t^0, \bar{P}_0\right) \to_{a.s.} \left(R^0, P^0\right) = \lim_{N\to\infty} \left\{ N^{-1} \sum_{i=1}^{N} \left(\mathbb{E}\left(R_{it}^0\right), \mathbb{E}\left(P_{i0}\right)\right) \right\}$ as $N \to \infty$. Thus, the global radiation effect is measured (asymptotically as $N \to \infty$) by

$$\overline{R}_t = R^0 + \delta_r' G_t + \bar{P}_t + o_{a.s.}(1) = \left(R^0 + \delta_r' G_0\right) + \delta_r' U_{gt} + \bar{P}_t + o_{a.s.}(1), \tag{6}$$

which evidently imports the nonstationarity of the partial sum process U_{gt} from G_t but with a small average coefficient effect, measured by the parameter δ_r.

According to A (iii) the local idiosyncratic trend component P_{it} has the stochastic trend representation $P_{it} = P_{i0} + U_{it}^P$. The partial sum component $U_{it}^P = \sum_{k=1}^{t} u_{ik}^P$ is assumed to satisfy a functional law and this implies that station-specific stochatic trends play a role in the limit theory, as will become apparent. However, at the global level these station specific trends are subject to cross section averaging, so that

$$\begin{aligned} \bar{P}_t &= \bar{P}_0 + N^{-1} \sum_{i=1}^{N} U_{it}^P = \bar{P}_0 + \sum_{k=1}^{t} \left(N^{-1} \sum_{i=1}^{N} u_{ik}^P \right) = \bar{P}_0 + \frac{\sqrt{n}}{\sqrt{N}} \times \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\frac{1}{\sqrt{n}} \sum_{k=1}^{t} u_{ik}^P \right) \\ &= \bar{P}_0 + O_p\left(\sqrt{\frac{n}{N}}\right) \to_p P^0 = \lim_{N\to\infty} N^{-1} \sum_{i=1}^{N} \mathbb{E}\left(P_{i0}\right), \end{aligned}$$

provided $\frac{n}{N} \to 0$ as $N \to \infty$, which is assumed in A(iv) and requires that the time series sample is small relative to the number of stations (spatial locations). Thus, $\bar{P}_t = P^0 + O_p\left(\sqrt{n/N}\right)$. It follows that when $\frac{n}{N} \to 0$ station-specific stochastic trends such as aerosol pollution average out through global averaging to some mean global level $P_0 = \bar{P}_0$. In effect, global averaging of the local pollution trends P_{it} to some mean level means that some areas may be cleaning up while others are deteriorating over time, leading to a net average effect that is negligible or constant. If there is any general global trend in pollution (say) then it can be considered part of the global effect G_t. That is, if any common aerosol pollution trends are present in local radiation these will be absorbed in the latent common global shock G_t and via the individual factor loading δ_i. Unlike the local trend effects in $\bar{P}_t$ that average out asymptotically, common trends that are embodied in G_t do have persistent effects in the model. Thus, any common world-wide aerosol pollution trends that may be present are manifested through G_t or as a separate component of a latent vector of common global shocks G_t.

It follows that the extended model (4) for local radiation impacts global radiation effects in a form that can be represented under the above assumptions as

$$\overline{R}_t = \delta_{r0} + \delta_r' U_{gt} + O_p\left(\sqrt{\frac{n}{N}}\right), \quad \text{where } \delta_{r0} = R^0 + \delta_r' G_0 + P_0. \tag{7}$$

Under A(iv) (7) implies that $\overline{R}_t = \delta_{r0} + \delta_r' U_{gt} + o_p(1)$. These conditions mean that global downwelling radiation is modeled as a unit root stochastic trend driven by the common global stochastic trend U_{gt} with average local loading factor δ, and initial conditions determined by a linear combination of mean local radiation (R^0), aerosol pollution (P^0), and initial global trend (δG_0) effects.

Assumption C

(i) $\ln(CO_{2,t}) = \delta_{c0} + \delta_c' U_{gt} + u_{ct}$, *where* (δ_{c0}, δ_c') *are fixed parameters,* $U_{gt} = \sum_{k=1}^{t} u_{gk}$ *as in B(ii), and*

(ii) $\{u_{ct}\}$ *is a zero mean short memory process whose partial sums* $U_{ct} = \sum_{k=1}^{t} u_{ck}$ *satisfy the invariance principle* $n^{-1/2} U_{c\lfloor nr \rfloor} \Rightarrow U_c(r) \equiv BM\left(\omega_c^2\right)$, *with* $\omega_c^2 > 0$.

C(i) and C(ii) imply that the GHG forcing variable $\ln(CO_{2,t})$ follows a stochastic trend driven by U_{gt}. This assumption means that the latent trend process U_{gt} affects both $\ln(CO_{2,t})$ and $\overline{R}_t$. If U_{gt} is a scalar process, then $\ln(CO_{2,t})$ and $\overline{R}_t$ share a single common stochastic trend driver U_{gt}, whereas if U_{gt} is a vector process, then more than one component of U_{gt} may combine to produce a common trend driver of $\ln(CO_{2,t})$ and $\overline{R}_t$. This formulation allows for some flexibility in the latent forcing variables that underlie GHG and radiation effects.

4 Econometric Implications

Aggregating (1) over stations gives

$$\overline{T}_{t+1} = \beta_1 \overline{T}_t + \beta_2 \bar{R}_t + \lambda_t + \frac{1}{N}\sum_{i=1}^{N} u_{it+1} = \beta_1 \overline{T}_t + \beta_2 \bar{R}_t + \lambda_t + O_p\left(N^{-1/2}\right), \quad (8)$$

so that

$$\lambda_t = \overline{T}_{t+1} - \beta_1 \overline{T}_t - \beta_2 \bar{R}_t + O_p\left(N^{-1/2}\right). \quad (9)$$

Since the energy balance variable $\lambda_t = \gamma_0 + \gamma_1 \overline{T}_t + \gamma_2 \overline{R}_t + \gamma_3 \ln(CO_{2,t})$ is assumed to be stationary and $\bar{R}_t$ has a stochastic trend, (8) implies that $\bar{R}_t$ cointegrates with the quasi-difference $\overline{T}_{t+1} - \beta_1 \overline{T}_t$. Thus, equations (2) and (8) together produce two cointegrating relationships among the three aggregate variables $\left(\overline{T}_t, \overline{R}_t, \ln(CO_{2,t})\right)$. These relationships can be expressed in terms of the common stochastic trend U_{gt} that acts as a forcing variable on the aggregate time series. The common trend expressions are defined in the following theorem.

Theorem 1 (Common trend drivers) *Under Assumptions A-C,* $\bar{W}_t = \left(\overline{T}_t, \overline{R}_t, \ln(CO_{2,t})\right)'$ *is a vector of stochastic trends driven by* U_{gt} *of the form* $\bar{W}_t = \delta_w + \Delta_w U_{gt} + u_{wt}^{+}$ *where* $u_{wt}^{+} = (u_{Tt}, 0, u_{ct})' + o_p(1)$ *is asymptotically stationary and*

$$\bar{W}_t := \begin{bmatrix} \overline{T}_t \\ \overline{R}_t \\ \ln(CO_{2,t}) \end{bmatrix} = \begin{bmatrix} \delta_{T0} + \delta_T' U_{gt} + u_{Tt}^{+} \\ \delta_{r0} + \delta_r' U_{gt} + O_p\left(\sqrt{\frac{n}{N}}\right) \\ \delta_{c0} + \delta_c' U_{gt} + u_{ct} \end{bmatrix} =: \delta_w + \Delta_w U_{gt} + u_{wt}^{+}, \quad (10)$$

where

$$\delta_{T0} = \frac{\gamma_0 + (\beta_2 + \gamma_2)\,\delta_{r0} + \gamma_3 \delta_{c0}}{1 - \beta_1 - \gamma_1}, \qquad \delta_T = \frac{\delta_r + \delta_c}{1 - \beta_1 - \gamma_1}, \quad (11)$$

$$u_{Tt} = \gamma_3 \sum_{j=0}^{t} (\beta_1 + \gamma_1)^j u_{ct-1-j} - \frac{(\beta_1 + \gamma_1)^2}{1 - \beta_1 - \gamma_1} \sum_{k=0}^{\infty} (\beta_1 + \gamma_1)^k u_{gt-1-k} - \delta_T' u_{gt}, \quad (12)$$

$$u_{Tt}^{+} = u_{Tt} + O_p\left(\frac{1}{\sqrt{N}} + \sqrt{\frac{n}{N}} + |\beta_1 + \gamma_1|^t\right), \quad (13)$$

with $\Delta_w' = [\delta_T, \delta_r, \delta_c]$, $\delta_w = [\delta_{T0}, \delta_{r0}, \delta_{c0}]'$, $u_{wt}^{+} = u_{wt} + O_p\left(\frac{1}{\sqrt{N}} + \sqrt{\frac{n}{N}} + |\beta_1 + \gamma_1|^t\right)$ *and* $u_{wt} = [u_{Tt}, 0, u_{ct}]'$.

Remarks

1. Since $\delta_T = \frac{\delta_r+\delta_c}{1-\beta_1-\gamma_1}$, it is apparent from (10) that $\overline{T}_t$ is cointegrated with $\left(\overline{R}_t, \ln(CO_{2,t})\right)$. In particular

$$\begin{aligned} & (1-\beta_1-\gamma_1)\,\overline{T}_t = (\delta_r+\delta_c)'\,U_{gt} + \delta_{T0}\,(1-\beta_1-\gamma_1) + u_{Tt}^{+}\,(1-\beta_1-\gamma_1) \\ = \; & \overline{R}_t + \ln(CO_{2,t}) + \mu + \zeta_t, \end{aligned} \tag{14}$$

with $\mu = \delta_{T0}\,(1-\beta_1-\gamma_1) - (\delta_{r0}+\delta_{c0})$ and $\zeta_t = u_{Tt}\,(1-\beta_1-\gamma_1) - u_{ct} + o_p\,(1)$.

2. From (2) and B(i) we have the energy balance cointegrating relationship $\lambda_t = \gamma_0 + \gamma_1\overline{T}_t + \gamma_2\overline{R}_t + \gamma_3\ln(CO_{2,t})$. Combining the latter with (10) implies that

$$\begin{aligned} \lambda_t &= (\gamma_0+\gamma_1\delta_{T0}+\gamma_2\delta_{r0}+\gamma_3\delta_{c0}) + \left(\gamma_1\delta_T'+\gamma_2\delta_r'+\gamma_3\delta_c'\right)U_{gt} + \gamma_1 u_{Tt} + \gamma_3 u_{ct} + o_p\,(1) \\ &=: \bar{\gamma}_0 + \bar{\gamma}_g' U_{gt} + u_{\lambda t} = \bar{\gamma}_0 + u_{\lambda t} \end{aligned} \tag{15}$$

which implies that

$$\bar{\gamma}_g' := \gamma_1\delta_T' + \gamma_2\delta_r' + \gamma_3\delta_c' = 0. \tag{16}$$

Hence, asymptotic stationarity of the energy balance error λ_t implies that the coefficients of the stochastic trend inputs satisfy (16) and then λ_t has the explicit formulation $\lambda_t = \bar{\gamma}_0 + u_{\lambda t}$ in terms of the stationary inputs (u_{Tt}, u_{ct}), where u_{Tt} is defined in (12).

3. From (9) we have

$$\begin{aligned} \lambda_t &= \overline{T}_{t+1} - \beta_1\overline{T}_t - \beta_2\bar{R}_t + O_p\left(N^{-1/2}\right) \\ &= \left(\delta_{T0} + \delta_T' U_{gt+1} + u_{Tt+1}\right) - \beta_1\overline{T}_t - \beta_2\bar{R}_t + O_p\left(\sqrt{\frac{n}{N}}\right) \\ &= (1-\beta_1)\,\overline{T}_t - \beta_2\bar{R}_t + \delta_T' u_{gt+1} + u_{Tt+1} - u_{Tt} + O_p\left(\sqrt{\frac{n}{N}}\right) \end{aligned}$$

which shows that the following linear combination of $\left(\overline{T}_t, \bar{R}_t\right)$

$$(1-\beta_1)\,\overline{T}_t - \beta_2\bar{R}_t = \lambda_t - \left\{\delta_T' u_{gt+1} + u_{Tt+1} - u_{Tt}\right\} + O_p\left(\sqrt{\frac{n}{N}}\right) \tag{17}$$

is asymptotically integrated of order zero (written as $\simeq_a I\,(0)$) as $N \to \infty$.

4. Thus, (2) and (17) deliver the (asymptotic) cointegrating relations

$$\gamma_1\overline{T}_t + \gamma_2\overline{R}_t + \gamma_3\ln(CO_{2,t}) \;\simeq_a\; I\,(0)\,, \tag{18}$$

$$(\beta_1 - 1)\,\overline{T}_t + \beta_2\overline{R}_t \;\simeq_a\; I\,(0)\,, \tag{19}$$

which require the following two conditions on the coefficients

$$\begin{aligned} \gamma_1 \delta_T' + \gamma_2 \delta_r' + \gamma_3 \delta_c' &= 0, \quad (20) \\ (\beta_1 - 1)\, \delta_T' + \beta_2 \delta_r' &= 0, \quad (21) \end{aligned}$$

Importantly, (17) implies the long run relationship $\overline{T}_t \simeq_a \frac{\beta_2}{1-\beta_1} \overline{R}_t + I(0)$ between aggregate temperature and downwelling radiation. Using observational data over the period 1964-2010, Storelvmo et al. (2016) obtained the empirical estimates $\beta_1 = 0.9212$ and $\beta_2 = 0.0127$, which lead to $\overline{T}_t \simeq_a 0.16 \times \overline{R}_t$.

Define the cointegrating matrix

$$\beta_\gamma' = \begin{bmatrix} \gamma_1 & \gamma_2 & \gamma_3 \\ \beta_1 - 1 & \beta_2 & 0 \end{bmatrix} \quad (22)$$

for which $\beta_\gamma' \Delta_w = 0$, so that Δ_w has unit rank. Let the m_g- vector β_w be an orthonormal vector complement of β_γ, and write the $3 \times m_g$ matrix Δ_w in the outer product form $\Delta_w = \beta_w a'$ for some m_g- vector a. Then,

$$\bar{W}_t = \delta_w + \beta_w a' U_{gt} + u_{wt}^+, \quad (23)$$

from which it follows that $\bar{W}_t$ has a one-dimensional forcing variable $U_{wt} = a' U_{gt}$ formed from the components of U_{gt}. Each of the time series $\left(\overline{T}_t, \overline{R}_t, \ln(CO_{2,t})\right)$ is therefore influenced by the composite effects of U_{wt} and we may write $\bar{W}_t$ in simplified form as

$$\bar{W}_t = \delta_w + \beta_w U_{wt} + u_{wt}^+. \quad (24)$$

It is convenient in what follows to define a subvector of the variables in (1) and (24) as follows. Define $X_{it} = (T_{i,t}, R_{i,t})'$ and then the station average $\bar{X}_t = \left(\overline{T}_t, \overline{R}_t\right)'$ has the following subvector form from (10)

$$\bar{X}_t = \begin{bmatrix} \overline{T}_t \\ \overline{R}_t \end{bmatrix} = \begin{bmatrix} \delta_{T0} + \delta_T' U_{gt} + u_{Tt} \\ \delta_{r0} + \delta_r' U_{gt} + O_p\left(\sqrt{\frac{n}{N}}\right) \end{bmatrix} =: \delta_{x0} + \beta_x U_{wt} + u_{xt}, \quad (25)$$

with $\beta_x' = [\beta_{wT}, \beta_{wr}]$, $\delta_{x0} = [\delta_{T0}, \delta_{r0}]'$, $u_{xt} = (u_{Tt}, 0) + o_p(1)$. In addition to the aggregate variables, it is useful to write the panel elements in terms of the global shock U_{gt} and the station-level trend effects U_{it}^P. Using (4), (5), and A(iii) we find that

$$R_{it} = \left(R_{it}^0 + \delta_{ri}' G_0 + P_{i0}\right) + \delta_{ri}' U_{gt} + U_{it}^P, \quad (26)$$

revealing the presence of persistent shock effects from $\left(U_{gt}, U_{it}^P\right)$ at the station level on R_{it} and on T_{it} via the panel equation (1). These stochastic trends both play a role in the asymptotic theory of the coefficient estimates from the panel regression.

To proceed in deriving the asymptotic theory for the coefficient estimates of (1) and (2), we first clarify the nature of the simple panel regression estimation procedure used here. The procedure has two steps as follows.

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, if we write the model (1) in regression form as

$$T_{i,t+1} = \beta_1 T_{i,t} + \beta_2 R_{i,t} + \lambda_t + u_{i,t+1} = \beta' X_{i,t} + \lambda_t + u_{i,t+1}, \tag{27}$$

then by spatial averaging

$$\hat{\lambda}_t = \bar{T}_{t+1} - \hat{\beta}' \bar{X}_t = \bar{T}_{t+1} - \hat{\beta}_1 \bar{T}_t - \hat{\beta}_2 \overline{R}_t, \tag{28}$$

with

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

where as above $\tilde{A}_{it} = A_{it} - \bar{A}_t$. This means that the time specific effects are estimated by (between, over i) regression and the coefficients β are estimated using pooled regression after elimination of the time specific effects.

Step 2. Regress $\hat{\lambda}_t$ on $\left(1, \overline{T}_t, \overline{R}_t, \ln(CO_{2,t})\right)$ giving the global cointegrating regression equation

$$\hat{\lambda}_t = \hat{\gamma}_0 + \hat{\gamma}_1 \overline{T}_t + \hat{\gamma}_2 \overline{R}_t + \hat{\gamma}_3 \ln(CO_{2,t}), \tag{29}$$

and the corresponding vector of coefficient estimates $(\hat{\gamma}_0, \hat{\gamma}_1, \hat{\gamma}_2, \hat{\gamma}_3)$. This regression can be performed by several methods, including ordinary least squares (OLS), dynamic OLS (DOLS) (Saikonnen, 1991; Phillips and Loretan, 1991; Stock and Watson, 1993), or fully modified regression (Phillips and Hansen, 1990). Both OLS and DOLS were used in Storelvmo et al. (2016). As shown below, in the present context it will be sufficient to use OLS regression.

5 Asymptotic theory

With this model framework in hand we may obtain a limit theory for the estimates $\left(\hat{\beta}, \hat{\gamma}\right)$ of the panel and cointegrating regression equations (1) and (2). This limit theory provides asymptotics for the estimates $\left(\hat{\beta}_1, \hat{\gamma}_1, \hat{\gamma}_3\right)$ of the relevant parameters $(\beta_1, \gamma_1, \gamma_3)$ that appear in the formula (3) for total climate sensitivity. We concentrate on $\left(\hat{\beta}_1, \hat{\gamma}_1, \hat{\gamma}_3\right)$ in what follows in order to develop methodology for inference about the key parameter TCS in (3).

First, the dynamic panel estimator $\hat{\beta}_1$ is obtained by linear least squares regression on (1). We use the notation $\tilde{A}_{it} = A_{it} - \bar{A}_t$ where $\bar{A}_t = N^{-1} \sum_{i=1}^{N} A_{it}$, so that $\tilde{A}_{it}$ is the cross section de-meaned A_{it}. Next, let

$$\tilde{T}_{i,t,R} = \tilde{T}_{i,t} - \frac{\sum_{s=1}^{T} \sum_{j=1}^{N} \tilde{T}_{j,s} \tilde{R}_{j,s}}{\sum_{s=1}^{n} \sum_{j=1}^{N} \tilde{R}_{j,s}^2} \tilde{R}_{i,t} \text{ and } \tilde{R}_{i,t,T} = \tilde{R}_{i,t} - \frac{\sum_{s=1}^{T} \sum_{j=1}^{N} \tilde{T}_{j,s} \tilde{R}_{j,s}}{\sum_{s=1}^{n} \sum_{j=1}^{N} \tilde{T}_{j,s}^2} \tilde{T}_{j,t} \tag{30}$$

be the residuals from the regressions of $\tilde{T}_{i,t}$ on $\tilde{R}_{i,t}$ and of $\tilde{R}_{i,t}$ on $\tilde{T}_{i,t}$, respectively. The partitioned least squares regression estimates $\left(\hat{\beta}_1, \hat{\beta}_2\right)$ from the panel regression (1) satisfy

$$\hat{\beta}_1 - \beta_1 = \frac{\sum_{t=1}^{n} \sum_{i=1}^{N} \tilde{T}_{i,t,R} u_{it+1}}{\sum_{t=1}^{T} \sum_{i=1}^{N} \tilde{T}_{i,t,R}^2}, \ \hat{\beta}_2 - \beta_2 = \frac{\sum_{t=1}^{n} \sum_{i=1}^{N} \tilde{R}_{i,t,T} u_{it+1}}{\sum_{t=1}^{T} \sum_{i=1}^{N} \tilde{R}_{i,t,T}^2}. \tag{31}$$

The following result gives the asymptotic distributions of $\hat{\beta}_1$ and $\hat{\beta}_2$.

Theorem 2 (Dynamic panel regression asymptotics) *Under Assumptions A, B, and C, as* $(n, N) \to \infty$ *with* $\frac{n}{N} \to 0$, *the following hold:*

(a) $\sqrt{n^2 N}\left(\hat{\beta}_1 - \beta_1\right) \Rightarrow \mathcal{MN}\left(0, \sigma_u^2 \frac{\operatorname{tr}\left\{\Sigma_r \int_0^1 B_g(s) B_g(s)' ds\right\} + \frac{1}{2}\omega_{uP}^2}{\left(\frac{1}{2}\omega_{uP}^2\right)\left(\frac{\beta_2}{1-\beta_1}\right)^2 \operatorname{tr}\left\{\Sigma_r \int_0^1 B_g(s) B_g(s)' ds\right\}}\right)$, *a mixed normal limit distribution with variance mixing variate that depends on the random quantity* $\operatorname{tr}\left\{\Sigma_r \int_0^1 B_g(s) B_g(s)' ds\right\}$ *involving the vector Brownian motion* B_g *associated with the global forcing variable* U_{gt}; and

(b) $\sqrt{n^2 N}\left(\hat{\beta}_2 - \beta_2\right) \Rightarrow \mathcal{N}\left(0, 2\frac{\sigma_u^2}{\omega_{uP}^2}\right)$.

In the proof of Theorem 2 (equation (57) and Lemma A1(iv)) it is shown that

$$\frac{1}{n^2 N} \sum_{t=1}^{n} \sum_{i=1}^{N} \tilde{T}_{i,t,R}^2 \Rightarrow \frac{\left(\frac{1}{2}\omega_{uP}^2\right)\left(\frac{\beta_2}{1-\beta_1}\right)^2 \left(\operatorname{tr}\left\{\Sigma_r \int_0^1 B_g(s) B_g(s)' ds\right\}\right)^2}{\operatorname{tr}\left\{\Sigma_r \int_0^1 B_g(s) B_g(s)' ds\right\} + \frac{1}{2}\omega_{uP}^2}.$$

Limit theory then follows for the self normalized estimation error, giving

$$\left(\sum_{t=1}^{n} \sum_{i=1}^{N} \tilde{T}_{i,t,R}^2\right)^{1/2} \left(\frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}_u}\right) \Rightarrow \mathcal{N}(0,1),$$

where $\hat{\sigma}_u^2$ is the usual least squares residual variance estimate of σ_u^2. Thus, confidence intervals for β_1 can be constructed in the standard way, with a $100(1-\alpha)\%$ interval taking the form

$$\hat{\beta}_1 \pm \frac{\hat{\sigma}_u}{\left(\sum_{t=1}^{n} \sum_{i=1}^{N} \tilde{T}_{i,t,R}^2\right)^{1/2}} z_\alpha, \tag{32}$$

where z_α is the $100(1-\alpha)$ percentile of the standard normal distribution. Similarly, Lemma A1(v) shows that $\frac{1}{n^2 N} \sum_{t=1}^{n} \sum_{i=1}^{N} \tilde{R}_{i,t,T}^2 \to_p \frac{1}{2}\sigma_u^2 \omega_{uP}^2$, and the corresponding self normalized estimation error limit theory is

$$\left(\sum_{t=1}^{n} \sum_{i=1}^{N} \tilde{R}_{i,t,T}^2\right)^{1/2} \left(\frac{\hat{\beta}_2 - \beta_2}{\hat{\sigma}_u}\right) \Rightarrow \mathcal{N}(0,1).$$

Some further comments on this limit theory are in order. First, parts (a) and (b) hold when there are station-level trends P_{it} present in R_{it} and $\omega_{uP}^2 > 0$ as in (4) and A(iii). If $\omega_{uP}^2 \to 0$, then the limit variances in (a) and (b) tend to infinity and the rate of convergence is lower than $\sqrt{n^2 N}$. If there are no station-level trends in R_{it}, the convergence rate is $\sqrt{nN}$ rather than $\sqrt{n^2 N}$. Second, as $\beta_1 \to 1$ it is evident that the asymptotic variance in the limit distribution of $\sqrt{n^2 N}\left(\hat{\beta}_1 - \beta_1\right)$ in part (a) tends to zero, which is indicative of a higher rate of convergence applying than $\sqrt{n^2 N}$, precisely as would be expected because of the additional signal induced by unit root persistence rather than transient adjustment in (1). Third, $\hat{\beta}_2$ is asymptotically normal, rather than mixed normal, because the standardized signal $\frac{1}{n^2 N}\sum_{t=1}^n \sum_{i=1}^N \tilde{R}_{i,t,T}^2 \to_p \frac{1}{2}\sigma_u^2 \omega_{uP}^2$ and is asymptotically constant. This is explained by the fact that global trend effects (that produce a variance normal mixture in the limit theory for $\hat{\beta}_1$) are eliminated from $\tilde{R}_{it}$ in the partitioned regression because of their dominating effect on the other regressor $\tilde{T}_{it}$ which ensures that these stochastic trends are projected out. On the other hand, the station-level trends are subjected to spatial averaging, leading to the presence of the constant factor $\frac{1}{2}\omega_{uP}^2$ in the limiting variance. Finally, combining (a) and (b) we have $\hat{\beta} = \beta + O_p\left(1/\sqrt{n^2 N}\right)$, a property that is useful in what follows later.

First, we proceed to obtain the asymptotic distribution of the parameter estimates of the equilibrium energy balance equation (29). It will be sufficient for our purpose to consider the OLS estimate $\hat{\gamma} = (\hat{\gamma}_1, \hat{\gamma}_2, \hat{\gamma}_3)'$. Before stating the asymptotic theory we provide some useful preliminaries concerning the regression. First, from the panel regression model (1) we obtain estimates of the global energy balance time effects $\{\lambda_t\}$ as in (28) by regression giving

$$\hat{\lambda}_t = \bar{T}_{t+1} - \hat{\beta}'\bar{X}_t = \bar{T}_{t+1} - \hat{\beta}_1 \bar{T}_t - \hat{\beta}_2 \overline{R}_t, \tag{33}$$

where $\hat{\beta} = \left(\hat{\beta}_1, \hat{\beta}_1\right)' = \left(\sum_{t=1}^n \sum_{i=1}^N \tilde{X}_{i,t}\tilde{X}_{i,t}'\right)^{-1}\left(\sum_{t=1}^n \sum_{i=1}^N \tilde{X}_{i,t}\tilde{T}_{i,t+1}\right)$. Since $\hat{\beta} = \beta + O_p\left(1/\sqrt{n^2 N}\right)$ and $\bar{X}_t = \delta_{x0} + \beta_x U_{wt} + u_{xt} = O_p\left(\sqrt{n}\right)$ from (25), we deduce that

$$\begin{aligned}
&\hat{\lambda}_t = \bar{T}_{t+1} - \hat{\beta}'\bar{X}_t = \bar{T}_{t+1} - \beta'\bar{X}_t + \left(\beta - \hat{\beta}\right)'\bar{X}_t \\
= \;& \lambda_t - \left(\hat{\beta} - \beta\right)'\bar{X}_t + N^{-1}\sum_{i=1}^N u_{i,t+1} = \lambda_t + \frac{1}{\sqrt{N}}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N u_{i,t+1}\right) - \sqrt{Nn^2}\left(\hat{\beta} - \beta\right)'\frac{\bar{X}_t}{\sqrt{n}}\frac{\sqrt{n}}{\sqrt{Nn^2}} \\
= \;& \lambda_t + \frac{1}{\sqrt{N}}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N u_{i,t+1}\right) - O_p\left(\frac{1}{\sqrt{Nn}}\right) = \lambda_t + \frac{1}{\sqrt{N}}\xi_{N,t+1} - O_p\left(\frac{1}{\sqrt{Nn}}\right)
\end{aligned} \tag{34}$$

where $\xi_{N,t+1} := N^{-1/2}\sum_{i=1}^N u_{i,t+1} \underset{N\to\infty}{\Rightarrow} \xi_{t+1} \equiv \mathcal{N}\left(0, \sigma_u^2\right)$. Since $\{u_{i,t}\}$ is $iid\left(0, \sigma_u^2\right)$ over t, the same property holds for ξ_t. In a suitably expanded probability space we may replace the weak convergence $N^{-1/2}\sum_{i=1}^{N} u_{i,t+1} \underset{N\to\infty}{\Rightarrow} \xi_{t+1}$ by

$$N^{-1/2}\sum_{i=1}^{N} u_{i,t+1} \to_{a.s.} \xi_{t+1}, \tag{35}$$

and accordingly write (34) as

$$\hat{\lambda}_t = \lambda_t + \frac{1}{\sqrt{N}}\xi_{t+1} + o_{a.s.}\left(\frac{1}{\sqrt{N}}\right), \tag{36}$$

while retaining weak convergence for the original variates and in the final limit theory.

Next the global cointegrating regression equation (29) is fitted using observations $\hat{\lambda}_t$ that come from the panel regression (33). It is convenient to write the equation in the following form

$$\hat{\lambda}_t = \hat{\gamma}_0 + \hat{\gamma}_1\overline{T}_t + \hat{\gamma}_2\overline{R}_t + \hat{\gamma}_3\ln(CO_{2,t}) + \hat{u}_{\lambda t} = \hat{\gamma}_0 + \hat{\gamma}'\bar{W}_t = \left(\hat{\gamma}_0 + \hat{\gamma}'\delta_w\right) + \left(\hat{\gamma}'\beta_w\right)U_{wt} + \hat{\gamma}'u_{wt}^{+} \tag{37}$$

where $\bar{W}_t = \delta_w + \beta_w U_{wt} + u_{wt}^{+}$ from (24), noting that U_{wt} is a scalar $I\left(1\right)$ process and β_w is a vector. Let $\widetilde{\bar{W}}_t = \bar{W}_t - n^{-1}\sum_{t=1}^{n}\bar{W}_t$ and then OLS regression gives

$$\hat{\gamma} = \left(\sum_{t=1}^{n}\widetilde{\bar{W}}_t\widetilde{\bar{W}}_t'\right)^{-1}\left(\sum_{t=1}^{n}\widetilde{\bar{W}}_t\widetilde{\hat{\lambda}}_t\right) = \left(\sum_{t=1}^{n}\widetilde{\bar{W}}_t\widetilde{\bar{W}}_t'\right)^{-1}\left(\sum_{t=1}^{n}\widetilde{\bar{W}}_t\left(\widetilde{\lambda}_t + \frac{1}{\sqrt{N}}\widetilde{\xi}_{N,t+1} + O_p\left(\frac{1}{\sqrt{Nn}}\right)\right)\right) \tag{38}$$

From (2), $\lambda_t = \gamma_0 + \gamma_1\overline{T}_t + \gamma_2\overline{R}_t + \gamma_3\ln(CO_{2,t}) = \gamma_0 + \gamma'\bar{W}_t$ so that $\widetilde{\lambda}_t = \gamma'\widetilde{\bar{W}}_t$, whence

$$\sqrt{N}\left(\hat{\gamma} - \gamma\right) = \left(\sum_{t=1}^{n}\widetilde{\bar{W}}_t\widetilde{\bar{W}}_t'\right)^{-1}\left(\sum_{t=1}^{n}\widetilde{\bar{W}}_t\left(\widetilde{\xi}_{N,t+1} + O_p\left(\frac{1}{\sqrt{n}}\right)\right)\right) \tag{39}$$

Since the regressors in the equation (37) are cointegrated, we rotate coordinates in order to obtain the limit theory (c.f. Park and Phillips, 1988, 1989; Phillips, 1988). In the present case, the rotation is achieved using the orthogonal matrix $H = \left[\beta_w, \beta_\perp\right]$, where $\beta_\perp$ is an orthogonal complement matrix. Then the matrix $\beta_\perp$ provides directions of cointegration because $\beta_\perp'\beta_w = 0$ and therefore $\beta_\perp'$ annihilates the unit root stochastic trend component $\beta_w U_{wt}$ of the vector of variables $\bar{W}_t$. Thus, $\beta_\perp'\bar{W}_t = \beta_\perp'\delta_w + \beta_\perp' u_{wt} \sim_a I\left(0\right)$, i.e., is asymptotically $I\left(0\right)$ and so the vectors of $\beta_\perp$ are cointegrating vectors of $\bar{W}_t$. Hence, $\beta_\perp \in \mathcal{R}\left(\beta_\gamma\right)$, the range space of the cointegrating matrix β_γ defined earlier in (22). The asymptotic distribution of $\hat{\gamma}$ is given in the following result.

Theorem 3 (Energy balance regression asymptotics) *Under Assumptions A, B, and C, as* $(n, N) \to \infty$ *with* $\frac{n}{N} \to 0$, *the following hold:*

(a) $\sqrt{nN}\left(\hat{\gamma} - \gamma\right) \Rightarrow \mathcal{N}\left(0, \sigma_u^2 \beta_\perp \left(\beta_\perp' \left\{\mathbb{E}\left(u_{wt} u_{wt}'\right)\right\} \beta_\perp\right)^{-1} \beta_\perp'\right)$,

(b) $\sqrt{n^2 N}\beta_w'\left(\hat{\gamma} - \gamma\right) \Rightarrow \mathcal{MN}\left(0, \sigma_u^2 \left(\int_0^1 a' \tilde{B}_g(s) \tilde{B}_g(s)' a\right)^{-1}\right)$,

where $\tilde{B}_g(s) = B_g(s) - \int_0^1 B_g(r)\, dr$.

Remarks

5. Part (a) shows that $\hat{\gamma}$ has a limit normal distribution with singular asymptotic covariance matrix $\sigma_u^2 \beta_\perp \left(\beta_\perp' \left\{\mathbb{E}\left(u_{wt} u_{wt}'\right)\right\} \beta_\perp\right)^{-1} \beta_\perp'$, reflecting the cointegration of the regressors in the energy balance equation (37). The matrix $\beta_\perp$ has rank 2, in accord with the number of independent cointegrating vectors in (37), viz., (18) and (19). Thus, the dominating component of the asymptotic theory of $\hat{\gamma}$ is normal and is delivered from the stationary components determined by the cointegration space of $\bar{W}_t$.

6. As shown in the proof of Theorem 3,

$$\left(\frac{1}{n}\sum_{t=1}^{n} H' \widetilde{\bar{W}}_t \widetilde{\bar{W}}_t' H\right)^{-1} \to_p \begin{bmatrix} \left(\beta_\perp' \left\{\mathbb{E}\left(u_{wt} u_{wt}'\right)\right\} \beta_\perp\right)^{-1} & O_{2\times 1} \\ O_{1\times 2} & 0 \end{bmatrix},$$

so that

$$n\left(\sum_{t=1}^{n} \widetilde{\bar{W}}_t \widetilde{\bar{W}}_t'\right)^{-1} \to_p \beta_\perp \left(\beta_\perp' \left\{\mathbb{E}\left(u_{wt} u_{wt}'\right)\right\} \beta_\perp\right)^{-1} \beta_\perp'. \tag{40}$$

and then $n\left(\sum_{t=1}^{n} \widetilde{\bar{W}}_t \widetilde{\bar{W}}_t'\right)^{-1}$ provides a consistent estimate of the signal matrix component of the limiting variance matrix of $\sqrt{nN}\left(\hat{\gamma} - \gamma\right)$. It follows that confidence intervals for linear combinations of γ such as $b'\gamma$ can be constructed using the Gaussian limit theory $\mathcal{N}\left(0, \sigma_u^2 \beta_\perp \left(\beta_\perp' \left\{\mathbb{E}\left(u_{wt} u_{wt}'\right)\right\} \beta_\perp\right)^{-1} \beta_\perp'\right)$ of Part (a) of Theorem 3, which for practical purposes means

$$\sqrt{nN} b'\left(\hat{\gamma} - \gamma\right) \sim_a \mathcal{N}\left(0, n\sigma_u^2 b' \left(\sum_{t=1}^{n} \widetilde{\bar{W}}_t \widetilde{\bar{W}}_t'\right)^{-1} b\right), \tag{41}$$

in view of (40). Thus, an asymptotic $100(1-\alpha)\%$ confidence region for $b'\gamma$ is

$$b'\hat{\gamma} \pm \frac{1}{N}\hat{\sigma}_u^2 b' \left(\sum_{t=1}^{n} \widetilde{\bar{W}}_t \widetilde{\bar{W}}_t'\right)^{-1} b z_\alpha, \tag{42}$$

where $\hat{\sigma}_u^2 = \frac{1}{n}\sum_{t=1}^{n} \hat{u}_{\lambda t}^2$, $\hat{u}_{\lambda t} = \hat{\gamma}_0 + \hat{\gamma}'\bar{W}_t$ are the cointegrating regression residuals, and z_a is the $100(1-\alpha)\%$ percentile of the standard normal distribution. The interval (42) is asymptotically validated by the above argument provided $b'\beta_\perp \neq 0$, that is provided $b \notin \ker(\beta_\perp')$. In other words, the interval is valid provided b is not proportional to β_w, the vector that spans the unit root space of $\bar{W}_t$. Notwithstanding this apparent limitation, as we now show the interval remains valid even in the case where $b = \beta_w$.

7. Part (b) shows that $\beta_w'\hat{\gamma}$ has a mixed normal limit distribution with convergence rate $\sqrt{n^2N}$. The faster rate is due to the fact that the direction β_w isolates the stochastic trend component of the regressor $\bar{W}_t$, thereby producing a stronger signal that leads to faster convergence in this direction. As shown in the proof of part (b), the limit theory can alternately be represented in stochastic integral form as

$$\sqrt{n^2N}\beta_w'(\hat{\gamma}-\gamma) \Rightarrow \left(\int_0^1 a'\tilde{B}_g(s)\tilde{B}_g(s)'a\right)^{-1}\int_0^1 a'\tilde{B}_g(s)\,dB_\xi(s).$$

Moreover, if $b_w = \mu\beta_w$ for some scalar $\mu \neq 0$, then from (84) in the proof of Theorem 3 we have

$$n^2 b_w' \left(\sum_{t=1}^{n} \widetilde{\bar{W}}_t \widetilde{\bar{W}}_t'\right)^{-1} b_w \Rightarrow \mu^2\sigma_u^2\left(\int_0^1 a'\tilde{B}_g(s)\tilde{B}_g(s)'a\right)^{-1}$$

and thus

$$\begin{aligned} &\Rightarrow \left\{b_w'\left(\sum_{t=1}^{n} \widetilde{\bar{W}}_t \widetilde{\bar{W}}_t'\right)^{-1} b_w\right\}^{-1/2}\sqrt{N}b_w'(\hat{\gamma}-\gamma) = \left\{n^2 b_w'\left(\sum_{t=1}^{n} \widetilde{\bar{W}}_t \widetilde{\bar{W}}_t'\right)^{-1} b_w\right\}^{-1/2}\sqrt{n^2N}b_w'(\hat{\gamma}-\gamma) \\ &\Rightarrow \left(\mu^2\sigma_u^2\right)^{-1/2}\left(\int_0^1 a'\tilde{B}_g(s)\tilde{B}_g(s)'a\right)^{1/2} \times \mathcal{MN}\left(0, \mu^2\sigma_u^2\left(\int_0^1 a'\tilde{B}_g(s)\tilde{B}_g(s)'a\right)^{-1}\right) \\ &= \mathcal{N}(0,1). \end{aligned}$$

Hence, a $100(1-\alpha)\%$ confidence region for $b_w'\gamma$ is

$$b_w'\hat{\gamma} \pm \frac{1}{N}\hat{\sigma}_u^2 b_w' \left(\sum_{t=1}^{n} \widetilde{\bar{W}}_t \widetilde{\bar{W}}_t'\right)^{-1} b_w z_\alpha, \tag{43}$$

and so the confidence interval (42) for $b'\gamma$ remains valid for all linear combinations b irrespective of the convergence rate and the limit theory. This uniformity in the confidence interval turns out to be useful in what follows.

6 Confidence Interval for Transient Climate Sensitivity

As in (3), it is convenient to write the formula for TCS in parametric function form as $TCS = f(\theta) = \frac{\gamma_3}{1-\beta_1-\gamma_1}\ln(2)$ with $\theta = (\beta_1, \gamma_1, \gamma_3)'$. Since the panel regression estimate $\hat{\theta}$ is consistent for θ, the delta method in conjunction with the limit distribution theory for $\hat{\theta}$ can be used to provide the asymptotic distribution of $\widehat{TCS} = f\left(\hat{\theta}\right)$. The result is complicated by the degeneracy in the limit distribution theory of the coefficient estimates $\hat{\theta}$ that arises from cointegration among the regressors in (1) and (2). This degeneracy is accommodated in the following result and in the subsequent remarks leading to the construction of a uniformly valid confidence interval for TCS.

Theorem 4 (TCS asymptotics) *Under Assumptions A, B, and C, as* $(n, N) \to \infty$ *with* $\frac{n}{N} \to 0$, the following results hold.

(a) $\sqrt{nN}\left(\widehat{TCS} - TCS\right) \Rightarrow \mathcal{N}\left(0, \sigma_u^2 b'\beta_\perp \left[\beta_\perp' \mathbb{E}\left(u_{wt}u_{wt}'\right)\beta_\perp\right]^{-1}\beta_\perp' b\right)$, *where* $b' = \frac{\ln(2)}{1-\beta_1-\gamma_1} \times \left(\frac{\gamma_3}{1-\beta_1-\gamma_1}, 0, 1\right)$ and $\gamma_3 \neq 0$.

(b) $\sqrt{n^2N}\left(\widehat{TCS} - TCS\right) \Rightarrow \mathcal{MN}\left(0, \sigma_u^2\left(\frac{\ln(2)}{1-\beta_1-\gamma_1}\right)^2\left(\int_0^1 a'\tilde{B}_g(s)\,\tilde{B}_g(s)'\,ads\right)^{-1}\right)$ *when* $\gamma_3 = 0$.

Remarks

8. The limit theory in part (a) applies except in the special case where $\gamma_3 = 0$. In that case there is no global CO_2 impact in the energy balance equation (2) and in that event, $TCS = \frac{\gamma_3}{1-\beta_1-\gamma_1}\ln(2) = 0$ and there is no climate sensitivity to CO_2. Thus part (a) is the case of primary interest.

9. The limit distribution in part (a) is normal and this leads to standard large sample inference. In particular, the variance $\sigma_{TCS}^2 = \sigma_u^2 b'\beta_\perp\left[\beta_\perp'\mathbb{E}\left(u_{wt}u_{wt}'\right)\beta_\perp\right]^{-1}\beta_\perp' b$ can be estimated using (40), which leads to the following consistent estimator

$$\hat{\sigma}_{TCS}^2 = \frac{\hat{\sigma}_u^2}{N}\hat{b}'\left(\sum_{t=1}^{n}\widetilde{\overline{W}}_t\widetilde{\overline{W}}_t'\right)^{-1}\hat{b},$$

where $\hat{b}' = \frac{\ln(2)}{1-\hat{\beta}_1-\hat{\gamma}_1} \times \left(\frac{\hat{\gamma}_3}{1-\hat{\beta}_1-\hat{\gamma}_1}, 0, 1\right)$, $\hat{\sigma}_u^2 = \frac{1}{n}\sum_{t=1}^{n}\hat{u}_{\lambda t}^2$ and $\hat{u}_{\lambda t} = \hat{\gamma}_0 + \hat{\gamma}'\bar{W}_t = \hat{\lambda}_t = \hat{\gamma}_0 + \hat{\gamma}_1\overline{T}_t + \hat{\gamma}_2\overline{R}_t + \hat{\gamma}_3\ln(CO_{2,t})$ are the cointegrating regression residuals. A $100(1-\alpha)\%$ confidence interval for TCS is now obtained as

$$\widehat{TCS} \pm \hat{\sigma}_{TCS}^2 z_\alpha. \tag{44}$$

10. Just as in (43) above, the confidence interval (44) remains valid even in the special case where $\gamma_3 = 0$, $TCS = 0$, and there is no true climate sensitivity to CO_2. Thus, (44) is a uniformly valid confidence interval for TCS and is robust to the presence or absence of climate sensitivity to CO_2.

11. The range of the matrix $\beta_\perp$ defines the cointegration space and this is the same as the range of the matrix β_γ obtained earlier in (22). Hence, we can write

$$\beta_\perp = \beta_\gamma \left(\beta_\gamma' \beta_\gamma\right)^{-1/2}$$

As shown in (89) in the proof of Corollary 4, for a linear combination of θ based on $b' = \frac{\ln(2)}{1-\beta_1-\gamma_1} \times \left(\frac{\gamma_3}{1-\beta_1-\gamma_1}, 0, 1\right)$ we have

$$b'\beta_\gamma = \frac{\ln(2)\gamma_3 \left(1-\beta_1\right)}{\left(1-\beta_1-\gamma_1\right)^2} \left[1, -1\right],$$

so that $b'\beta_\gamma = b'\beta_\perp = 0$ iff $\gamma_3 = 0$ when $|\beta_1| < 1$. This is the special case discussed above where $TCS = 0$.

7 Inference on Earth's Climate Sensitivity

This section reports applications of the above methods to the study of Earth's transient climate sensitivity. The following sections briefly describe these applications which use two sources of information: (i) spatio-temporal empirical observations; and (ii) climate model data computed using the same spatio-temporal coordinates with outputs from several leading climate models.

7.1 Observational Data

We use three observational data sets, each of which records time series at multiple surface stations for one of the three aforementioned variables in Equations (1) and (2): temperature, surface radiation and equivalent CO_2. Due to data availability we limit the study to the 42-year time period from 1964 to 2005. In the following we briefly describe each of the data sets, and refer readers to Storelvmo et al. (2016) and the references therein for further details on the observational data.

7.1.1 Solar radiation data

Surface measurements of monthly mean incoming (i.e. downward) solar radiation (measured in watts per meter squared) are available from the Global Energy Budget Archive (GEBA, Gilgen and Ohmura, 1999) for more than 2,500 surface stations worldwide. The stations are unevenly distributed over Earth's land surface (see Figure 1 of Storelvmo et al., 2016), and are often not continuous in time. For the present study we only included stations that passed our data quality control and that met our time series length requirements, leaving us with approximately 1,300 stations. Since the time increment in Equations (1) and (2) is one year, we created annual means based on the monthly mean GEBA data for each station, and only included years in which data for all 12 months were available. We thus obtained a matrix of $\sim 1,300 \times 42$ surface radiation observations, albeit with occasional data gaps, i.e. an unbalanced panel.

7.1.2 Temperature data

Our surface temperature observations are obtained from the Climate Research Unit (CRU, Harris et al. 2014), available for download from the British Atmospheric Data Center (BADC, https://badc.nerc.ac.uk). Specifically, we use their gridded surface air temperature data set (version 3.10), available at a spatial resolution of 0.5°. Each of the ~1,300 stations for which adequate radiation data existed were then assigned corresponding temperature time series taken from the $0.5° \times 0.5°$ grid in which they were located, creating another $\sim 1,300 \times 42$ data matrix with the same missing data points as for the radiation data.

7.1.3 GHG data

Global and annual mean GHG concentrations are available from the National Oceanic and Atmospheric Administration (NOAA) Annual Greenhouse Gas Index (AGGI, http://www.esrl.noaa.gov/gmd/aggi) data set (Hofmann et al, 2006). The AGGI data set provides time series of equivalent CO_2 concentrations in the atmosphere, which is calculated by taking the radiative forcings associated with changes in all non-CO_2 GHGs (mainly methane and nitrous oxide) and converting them into equivalent changes in atmospheric CO_2 (in other words, the CO_2 increase required to produce the same forcing). Carbon dioxide, nitrous oxide and methane all have long atmospheric lifetimes (from hundreds to tens of years) and are therefore considered well-mixed, meaning that their atmospheric concentrations show little spatial variability. All surface stations are therefore assigned the same 42-yr equivalent CO_2 time series.

7.2 Climate Model Data

We use data from three of the GCMs that participated in the Coupled Model Intercomparison Project - Phase 5 (CMIP5, Taylor et al., 2012), namely the BCC-CSM1.1 (hereafter BCC), the HadCM3 (hereafter CM3) and the BNU models (see Table 1 for the salient features of these models). We use data from their historical simulations, run from 1850 to 2005, forced with changing GHG and aerosol concentrations (Lamarque et al., 2010). While only one such simulation is available for the BNU model, a 3-member ensemble of simulations (r1, r2 and r3) are available for BCC and CM3. The ensemble members

Table 1: Overview of salient GCM features

Short Name	Institution	Horizontal resolution	Reported TCS
BCC	Beijing Climate Center	∼250km	1.7K
CM3	UK Met Office Hadley Center	∼300km	2.0K
BNU	Beijing Normal University	∼250km	2.6K

differ only in their initializations, which are selected from different times in a steady-state pre-industrial simulation by the same model. While the different ensemble members are forced with the same data, their different initial conditions yield slightly different climate trajectories, each considered to be equally likely outcomes.

Generally, each model's realism is judged based on the extent to which the observed climate trajectory lies within the ensemble envelope of trajectories. The purpose of running ensemble simulations is to allow for an assessment of the statistical significance of any apparent differences between different models or between model paths and observations. The CMIP5 data archive contains output from a total of more than 30 different GCMs. The 3 models that we included in the present analysis were not randomly selected, but chosen because they differed in their reported TCS values: BCC has a relatively low reported TCS of 1.7K; CM3 has an intermediate TCS of 2.0K; and BNU has a TCS of 2.6K which lies in the upper end of the TCSs that were reported from GCMs. The reported TCSs for all CMIP5 models are available in Flato et al. (2013). The values are calculated for each model by running a simulation in which atmospheric CO_2 is increased by 1% per year until doubling is reached. The TCS is then calculated as the global mean temperature difference between the last and the first decade of simulation.

7.3 Econometric Analysis of Observational and Climate Model Data

A primary empirical motivation for the present study was to determine whether the econometric analysis applied to observations in Storelvmo et al. (2016) could successfully determine the TCSs of GCMs if the same analysis was applied to GCMs, and if GCM output was only included where observational data is available in both space and time. The degree of success in this exercise is here defined as the extent to which the TCS emerging from the econometric analysis agrees with the reported value for each model – more specifically the extent to which the reported TCS values lie within the calculated 95% confidence interval calculated from observational data as indicated in Theorem 4. Furthermore, differences between observed and modeled sensitivities to radiation and equivalent CO_2, as measured by differences in the estimates of the parameters of Equations (1) and (2), can reveal model shortcomings that are not easily revealed with standard model validation procedures.

Table 2 reports summary statistics for mean annual changes in temperature and radiation. The mean annual change in observed temperature is 0.021°C, with an estimated standard error of 0.038°C. The observed mean change in temperature is similar to the mean change in the GCM model simulations in most cases, which range from 0.009 (HadCM3 r_1) to 0.030 (BNU). The standard deviations for the GCM simulations all exceed the observed standard deviation. These descriptive figures indicate that the GCMs fit observed global average temperature reasonably well but with greater variation over time. This finding is corroborated by the curves shown in Figure 1, which trace the simulated and observed evolution in temperatures for the time period 1964 - 2000.

The mean annual change in the observed radiation is -0.147, with a standard deviation of 0.31. As the radiation time series in Figure 2 shows, there is a strong negative trend in radiation until the early nineties, when the trend shifts positive. This observed pattern impacts the sample mean, as the annual change moves from being mostly negative each year to mostly positive each year. The effects even out upon averaging, but because the period of the negative trend is longer and more persistent than the period of positive trend, the overall mean is negative.

For the GCM simulated radiation data, all but ensembles 1 and 3 for the model HadCM3, have a positive mean. This failure by the GCMs to reproduce observed radiation trends is confirmed by Figure 2, which shows that the GCMs generally show little or no radiation trend for the time period in question. The observed overall negative radiation trend, which has been attributed to changes in atmospheric aerosol loading, was found in Storelvmo et al. (2016) to have caused a cooling that "masked" $\sim$1/3 of the GHG warming for the time period in question. The lack of radiation trend found in the GCMs therefore suggests that these models severely underestimate the cooling effect. Without this bias, the GCMs would require a higher sensitivity to equivalent CO_2 in order to maintain a temperature trend in their simulations consistent with observations. This is a finding of profound importance for the global climate modeling community, which we intend to revisit in more detail in a future publication.

We next turn to the estimated parameters of Equations (1) and (2), and the resulting TCSs calculated as in Equation (3), all provided in Table 3 for both the observations and the GCMs. The observations and GCMs are in reasonable agreement for β_1, while γ_1, which relates the station-averaged temperature in a given year to the station-specific temperature in the previous year, is generally lower in magnitude in the GCMs (most notably BCC) than in the observations. The GCMs also consistently yield lower estimates for γ_3, which relates local temperatures to global equivalent CO_2 concentrations. Ultimately, because the TCS is a function of all the above three parameters, these biases end up partly canceling to produce less biased TCS values (relative to individual parameters) from the GCMs compared to the observations. However, this potentially points to an incorrect climate response function in the GCMs, in which these models respond too slowly to increasing GHG concentrations.

These empirical results are remarkably consistent with findings by Hansen et al. (2011), who argued that GCMs can nevertheless reasonably reproduce 20$^{\text{th}}$ century temperature records because they significantly underestimate aerosol cooling. Our results largely support this argument. Specifically, our observational analysis suggests that local (i.e. station-specific) aerosol-mediated downward radiation trends produce a cooling, as shown by the positive and significant β_2-value. This is consistent with expectations. In contrast, the GCMs produce either negative or insignificant β_2 values, which implies that less solar radiation reaching the surface locally results in a counterintuitive warming, or else no effect at all. The observational analysis further suggests that the station-averaged radiation trend has an insignificant impact on station-specific temperatures (as evident from the γ_2 estimate and associated standard error). The BCC and BNU models support an insignificant γ_2 value, while the CM3 ensemble members produce positive and significant (with the exception of ensemble $r3$) values. In other words, CM3 does produce the expected relationship between radiation and temperature, but mediated through station-averaged rather than local radiation. However, given the lack of radiation trends in the GCMs, the radiation impact on the 1964-2005 temperature simulated by CM3 is nevertheless weak.

Finally, the plots shown in Figure 3 compare the estimated TCS and associated confidence interval based on the observations with those from the GCMs. The TCS resulting from the observational analysis is somewhat higher than those resulting from the GCM analysis, which is not surprising given the lack of aerosol cooling in the GCMs. Notwithstanding this difference, the observational TCS confidence interval does include all GCM-based estimates. Furthermore, the agreement between the estimated and reported TCS values is remarkably good. This is extremely encouraging, because it suggests that the econometric methodology used to arrive at the estimated TCS values is sound in the sense that there is close matching of the empirical model-implied and GCM-implied estimates. In future work we will extend the analysis to the entire CMIP5 GCM archive, in order to examine whether the findings reported here for three leading GCMs hold more generally.

8 Conclusion

The research reported here had three goals: (i) construction of an econometric framework and inferential tools for studying Earth's climate sensitivity to atmospheric greenhouse gases, allowing for empirically acknowledged local aerosol pollution and global forcing variables that embody stochastic trends; (ii) development of asymptotic theory required to validate the use of these econometric tools in practical work on climate; and (iii) application of this modeling and inferential machinery to both observational and global climate model simulated data. The empirical findings reveal that three leading global climate models provide reasonable reproductions of actual temperature trajectories over nearly half a century to 2005 but that these models uniformly underestimate the aerosol cooling induced by negative trends in local downwelling radiation. The application also provides an observational-data-based confidence interval for Earth's transient climate sensitivity to greenhouse gas emissions. The TCS estimates reported by the GCMs all lie within this wide confidence interval but they are all lower than the observation-based estimate of TCS, most likely because of the GCM underestimation of aerosol cooling effects. In these respects, there appears to be a mechanism within the GCMs that compensates in some way(s) for their perceived bias in measuring aerosol effects. A more extensive investigation of this matter will be undertaken in a further study that extends this analysis to the full archive of climate models.

In modeling climate data, theory restrictions that balance global energy forces play a key role in empirical modeling. As shown in the present study, these balancing forces among trending data affect the asymptotic theory of estimation and rates of convergence, but standard methods of inference may still be validated under certain regularity conditions. Extensions of the results given here are possible to more general multivariate panels or temporal-spatial systems that involve transient responses to nonstationary data in conjunction with cointegrating relations that prevail among spatial aggregates. The asymptotic results indicate that, with appropriate methods and regularity conditions, inference is possible even in the case of signal degeneracies that may be induced by co-movement in the data at both the transient and aggregate levels.

	Temperature			
	Mean	Std.dev.	Min	Max.
Observations	0.021	0.248	-0.475	0.605
BCC r_1	0.022	0.273	-0.606	0.495
BCC r_2	0.038	0.336	-0.863	0.734
BCC r_3	0.028	0.299	-0.705	0.674
HadCM3 r_1	0.009	0.341	-0.569	0.666
HadCM3 r_2	0.015	0.272	-0.757	0.708
HadCM3 r_3	0.013	0.333	-0.583	0.629
BNU	0.030	0.281	-0.490	0.668
	Radiation			
	Mean	Std.dev.	Min	Max.
Observations	-0.15	2.00	-5.45	3.90
BCC r_1	0.06	3.79	-8.73	8.12
BCC r_2	0.11	2.46	-5.70	5.30
BCC r_3	0.01	3.45	-8.40	6.67
HadCM3 r_1	-0.09	2.31	-6.87	5.90
HadCM3 r_2	0.00	2.15	-5.68	4.16
HadCM3 r_3	-0.07	2.80	-7.37	5.42
BNU	0.07	2.78	-4.98	7.04

Table 2: Mean annual change in temperature and radiation for the observational data and the three GCMs.

	Parameter estimates						
	β_1	β_2	γ_0	γ_1	γ_2	γ_3	TCS
Observ.	0.919	0.013	0.114	-0.900	0.010	4.571	2.37
Std.Err.	0.005	0.003	0.059	0.168	0.008	1.005	1.07
BCC r_1	0.985	-0.011	-0.049	-0.766	0.009	2.732	1.81
Std. Err.	0.007	0.001	0.080	0.189	0.013	1.281	0.77
BCC r_2	0.916	-0.016	0.010	-0.552	-0.004	2.554	1.92
Std. Err.	0.009	0.001	0.092	0.195	0.020	1.024	0.72
BCC r_3	0.985	-0.011	0.017	-0.766	0.009	1.992	1.69
Std. Err.	0.007	0.001	-0.049	0.189	0.013	0.881	0.41
HadCM3 r_1	0.938	-0.007	-0.160	-0.800	0.032	3.826	2.25
Std. Err.	0.007	0.001	0.082	0.167	0.015	1.203	0.94
HadCM3 r_2	0.925	-0.009	-0.116	-0.894	0.040	3.647	1.91
Std. Err.	0.008	0.001	0.062	0.159	0.012	0.867	1.11
HadCM3 r_3	0.934	-0.010	-0.019	-1.091	0.017	4.295	1.88
Std. Err.	0.008	0.001	0.069	0.166	0.011	0.960	0.71
BNU	0.937	-0.002	-0.118	-0.708	-0.004	2.860	2.09
Std. Err.	0.005	0.001	0.078	0.192	0.016	1.477	0.35

Table 3: Parameter estimates of the econometric model (1)-(2) obtained using observational data and global climate model (GCM) data. Transient Climate Sensitivity (TCS) is found from (3) and its standard error is computed according to Theorem 4.

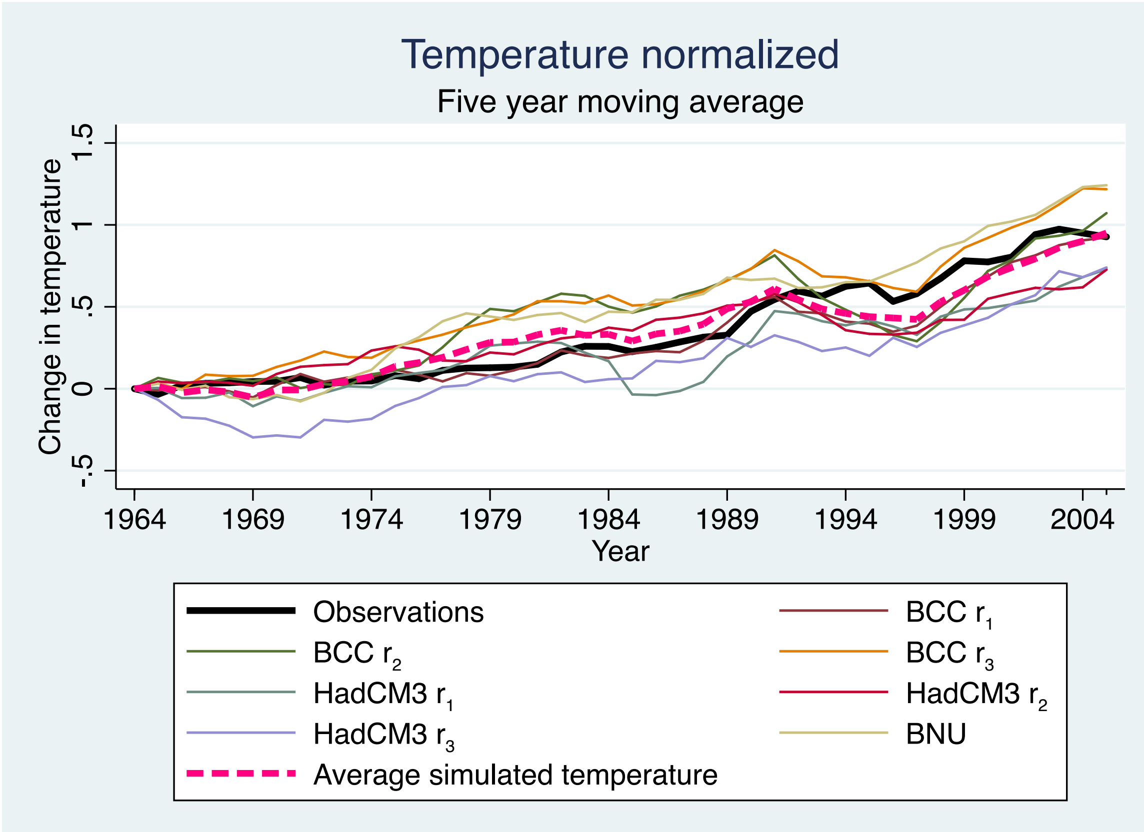

Figure 1: Station-averaged temperature change (in Kelvin) for observations and GCMs.

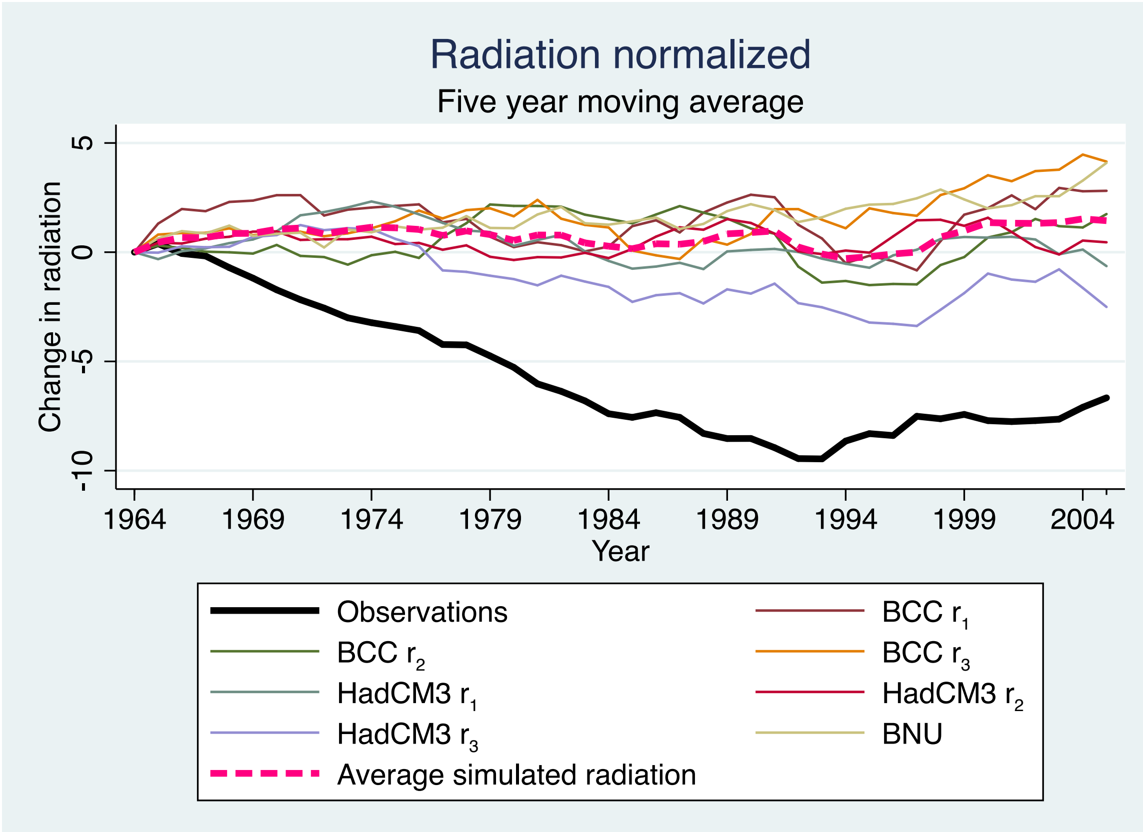

Figure 2: Station-averaged radiation change (in Wm^{-2}) for observations and GCMs.

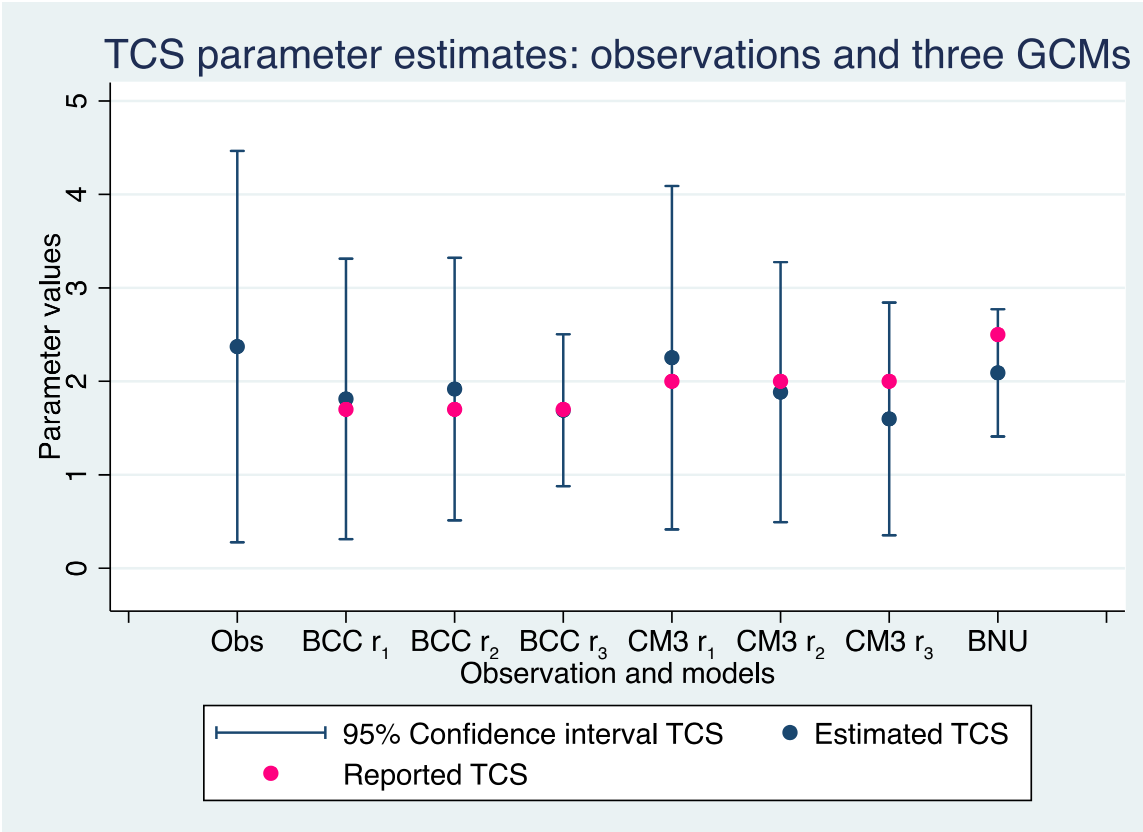

Figure 3: Estimated TCSs (blue symbols) and associated confidence intervals resulting from the econometric analysis, and reported TCS values for the GCMs (red symbols).

9 Appendix

9.1 The TCS Formula

The long-run equilibrium temperature is assumed to be such that $T_{i,t} = T_{i,t-1}$ and in global equilibrium $\overline{T}_t = \overline{T}_{t-1}$. Aggregating the transient relation (1) and using the energy balance equation (2) gives

$$\begin{aligned}\overline{T}_{t+1} &= \beta_1 \overline{T}_t + \beta_2 \bar{R}_t + \lambda_t + o_p(1) \\ &= \beta_1 \overline{T}_t + \beta_2 \bar{R}_t + \gamma_0 + \gamma_1 \overline{T}_t + \gamma_2 \overline{R}_t + \gamma_3 \ln(CO_{2,t}) + o_p(1), \end{aligned} \tag{45}$$

since $N^{-1} \sum_{i=1}^{N} u_{it+1} \rightarrow_p 0$ under Assumption A(i). Solving (45) leads to the equilibrium solution

$$\overline{T}_t = \frac{(\beta_2 + \gamma_2)\, \bar{R}_t + \gamma_3 \ln(CO_{2,t}) + \gamma_0}{1 - \beta_1 - \gamma_1} + o_p(1),$$

and taking differentials gives up to an $o_p(1)$ error we have

$$dT = \frac{\beta_2 + \gamma_2}{1 - \beta_1 - \gamma_1} dR + \frac{\gamma_3 d \ln(CO_2)}{1 - \beta_1 - \gamma_1}, \tag{46}$$

which measures a shift in global steady state temperatures, corresponding to equation (9) in Magnus et al. (2011).

Transient climate sensitivity, TCS, is defined as the expected global temperature after a doubling of CO_2, and is therefore computed using (46) by

$$TCS = \frac{\beta_2 + \gamma_2}{1 - \beta_1 - \gamma_1} \Delta R + \frac{\gamma_3 \Delta \ln(CO_2)}{1 - \beta_1 - \gamma_1} \tag{47}$$

where $\Delta \ln(CO_2) = \ln CO_{2,t+k} - \ln CO_{2,t} = \ln \frac{CO_{2,t+k}}{CO_{2,t}}$, and $\Delta R = R_{t+k} - R_t$. Year $t+k$ is when a doubling of CO_2 happens. If radiation is held constant, then $\Delta R = 0$, and $\Delta \ln(CO_2) = \ln \frac{CO_{2,t+k}}{CO_{2,t}} = \ln \frac{2 \times CO_{2,t}}{CO_{2,t}} = \ln 2$, so that

$$TCS = \frac{\gamma_3}{1 - \beta_1 - \gamma_1} \times \ln(2), \tag{48}$$

giving (3).

9.2 Proof of Theorem 1

The stochastic trend representation of $\overline{R}_t$ and $\ln(CO_{2,t})$ follow directly from (7) and Assumption C(i), giving

$$\begin{bmatrix} \overline{R}_t \\ \ln(CO_{2,t}) \end{bmatrix} = \begin{bmatrix} \delta_{r0} + \delta_r' U_{gt} + O_p\left(\sqrt{\frac{n}{N}}\right) \\ \delta_{c0} + \delta_c' U_{gt} + u_{ct} \end{bmatrix}. \tag{49}$$

To establish the representation of $\overline{T}_t$, we proceed as follows. First, aggregating (1) gives

$$\overline{T}_{t+1} = \beta_1 \overline{T}_t + \beta_2 \bar{R}_t + \lambda_t + \frac{1}{N}\sum_{i=1}^{N} u_{it+1} = \beta_1 \overline{T}_t + \beta_2 \bar{R}_t + \lambda_t + O_p\left(N^{-1/2}\right) \tag{50}$$

and substituting the energy balance relation (2) for λ_t into (50) we have

$$\overline{T}_{t+1} = (\beta_1 + \gamma_1)\overline{T}_t + (\beta_2 + \gamma_2)\bar{R}_t + \gamma_3 \ln(CO_{2,t}) + O_p\left(N^{-1/2}\right). \tag{51}$$

Solving (51) for $\overline{T}_{t+1}$ in terms of the past history of the inputs $\left(\bar{R}_t, \ln(CO_{2,t})\right)$ leads to the representation

$$\begin{aligned}
& \overline{T}_{t+1} = \gamma_0 \sum_{j=0}^{t} (\beta_1+\gamma_1)^j + (\beta_2+\gamma_2)\sum_{j=0}^{t}(\beta_1+\gamma_1)^j \bar{R}_{t-j} \\
& +\gamma_3 \sum_{j=0}^{t} (\beta_1+\gamma_1)^j \ln(CO_{2,t-j}) + O_p\left(N^{-1/2}\right) \\
= \;& \frac{\gamma_0}{1-\beta_1-\gamma_1} + (\beta_2+\gamma_2)\sum_{j=0}^{t}(\beta_1+\gamma_1)^j \bar{R}_{t-j} \\
& +\gamma_3 \sum_{j=0}^{t} (\beta_1+\gamma_1)^j \ln(CO_{2,t-j}) + O_p\left(\frac{1}{\sqrt{N}} + |\beta_1+\gamma_1|^t\right).
\end{aligned} \tag{52}$$

Next substitute the stochastic trend representations of $\bar{R}_t$ and $\ln(CO_{2,t})$ in (52), giving

$$\begin{aligned}
\overline{T}_{t+1} &= \frac{\gamma_0}{1-\beta_1-\gamma_1} + (\beta_2+\gamma_2)\sum_{j=0}^{t}(\beta_1+\gamma_1)^j \left\{\delta_{r0} + \delta_r' U_{gt-j} + +O_p\left(\sqrt{\frac{n}{N}}\right)\right\} \\
&\quad +\gamma_3 \sum_{j=0}^{t}(\beta_1+\gamma_1)^j \left\{\delta_{c0} + \delta_c' U_{gt-j} + u_{ct-j}\right\} + O_p\left(\frac{1}{\sqrt{N}} + |\beta_1+\gamma_1|^t\right) \\
&= \frac{\gamma_0 + (\beta_2+\gamma_2)\delta_{r0} + \gamma_3\delta_{c0}}{1-\beta_1-\gamma_1} + \sum_{j=0}^{t}(\beta_1+\gamma_1)^j (\delta_r+\delta_c)' U_{gt-j} \\
&\quad +\gamma_3 \sum_{j=0}^{t}(\beta_1+\gamma_1)^j u_{ct-j} + O_p\left(\frac{1}{\sqrt{N}} + \sqrt{\frac{n}{N}} + |\beta_1+\gamma_1|^t\right) \\
&= \frac{\gamma_0 + (\beta_2+\gamma_2)\delta_{r0} + \gamma_3\delta_{c0}}{1-\beta_1-\gamma_1} + \frac{(\delta_r+\delta_c)' U_{gt}}{1-\beta_1-\gamma_1} + u_{Tt} + O_p\left(\frac{1}{\sqrt{N}} + \sqrt{\frac{n}{N}} + |\beta_1+\gamma_1|^t\right)
\end{aligned}$$

To demonstrate the final line observe that

$$\begin{aligned}
\sum_{j=0}^{t}(\beta_1+\gamma_1)^j U_{gt-j} &= \sum_{j=0}^{t}(\beta_1+\gamma_1)^j \sum_{s=1}^{t-j} u_{gs} = \sum_{s=1}^{t} u_{gs} \sum_{j=0}^{t-s}(\beta_1+\gamma_1)^j = \sum_{s=1}^{t} u_{gs} \frac{1-(\beta_1+\gamma_1)^{t-s+1}}{1-\beta_1-\gamma_1} \\
&= \frac{1}{1-\beta_1-\gamma_1}\sum_{s=1}^{t} u_{gs} - \frac{1}{1-\beta_1-\gamma_1}\sum_{s=1}^{t} u_{gs}(\beta_1+\gamma_1)^{t-s+1} \\
&= \frac{1}{1-\beta_1-\gamma_1}\sum_{s=1}^{t} u_{gs} - \frac{\beta_1+\gamma_1}{1-\beta_1-\gamma_1}\sum_{k=0}^{t-1} u_{gt-k}(\beta_1+\gamma_1)^k \\
&= \frac{1}{1-\beta_1-\gamma_1}U_{gt} - \frac{\beta_1+\gamma_1}{1-\beta_1-\gamma_1}\sum_{k=0}^{\infty} u_{gt-k}(\beta_1+\gamma_1)^k + O_p\left(|\beta_1+\gamma_1|^t\right) \\
&= \frac{1}{1-\beta_1-\gamma_1}U_{gt} + v_t + O_p\left(|\beta_1+\gamma_1|^t\right), \qquad (53)
\end{aligned}$$

where $v_t = -\frac{\beta_1+\gamma_1}{1-\beta_1-\gamma_1}\sum_{k=0}^{\infty} u_{gt-k}(\beta_1+\gamma_1)^{k+1}$ is stationary and

$$\sum_{k=t}^{\infty} u_{gt-k}(\beta_1+\gamma_1)^{k+1} = (\beta_1+\gamma_1)^{t+1}\sum_{j=0}^{\infty} u_{-j}(\beta_1+\gamma_1)^j = O_p\left(|\beta_1+\gamma_1|^t\right).$$

It follows that

$$\sum_{j=0}^{t}(\beta_1+\gamma_1)^j(\delta_r+\delta_c)' U_{gt-j} = \frac{(\delta_r+\delta_c)' U_{gt}}{1-\beta_1-\gamma_1} + v_t^+,$$

where $v_t^+ = v_t + O_p\left(|\beta_1+\gamma_1|^t\right)$, thereby demonstrating that

$$\begin{aligned}
\overline{T}_{t+1} &= \frac{\gamma_0+(\beta_2+\gamma_2)\delta_{r0}+\gamma_3\delta_{c0}}{1-\beta_1-\gamma_1} + \frac{(\delta_r+\delta_c)' U_{gt}}{1-\beta_1-\gamma_1} + v_t + \gamma_3\sum_{j=0}^{t}(\beta_1+\gamma_1)^j u_{ct-j} + O_p\left(|\beta_1+\gamma_1|^t\right) \\
&= \delta_{T0} + \delta_T' U_{gt} + u_{Tt}^{\#}, \qquad (54)
\end{aligned}$$

with

$$\begin{aligned}
\delta_{T0} &= \frac{\gamma_0+(\beta_2+\gamma_2)\delta_{r0}+\gamma_3\delta_{c0}}{1-\beta_1-\gamma_1}, \quad \delta_T = \frac{\delta_r+\delta_c}{1-\beta_1-\gamma_1}, \\
u_{Tt}^{\#} &= v_t + \gamma_3\sum_{j=0}^{\infty}(\beta_1+\gamma_1)^j u_{ct-j} + O_p\left(\frac{1}{\sqrt{N}} + \sqrt{\frac{n}{N}} + |\beta_1+\gamma_1|^t\right) \\
&= -\frac{\beta_1+\gamma_1}{1-\beta_1-\gamma_1}\sum_{k=0}^{\infty} u_{gt-k}(\beta_1+\gamma_1)^{k+1} + \gamma_3\sum_{j=0}^{\infty}(\beta_1+\gamma_1)^j u_{ct-j} + O_p\left(\frac{1}{\sqrt{N}} + \sqrt{\frac{n}{N}} + |\beta_1+\gamma_1|^t\right)
\end{aligned}$$

so that $u_{Tt}^{\#}$ is asymptotically stationary. Define the stationary process

$$u_{Tt}=-\frac{\beta_1+\gamma_1}{1-\beta_1-\gamma_1}\sum_{k=0}^{\infty}u_{gt-k}\left(\beta_1+\gamma_1\right)^{k+1}+\gamma_3\sum_{j=0}^{\infty}\left(\beta_1+\gamma_1\right)^{j}u_{ct-j}-\delta_T'u_{gt}. \tag{55}$$

Then, combining (49) and (54) and writing

$$\overline{T}_t=\delta_{T0}+\delta_T'U_{gt}+u_{Tt-1}^{\#}-\delta_T'u_{gt}=\delta_{T0}+\delta_T'U_{gt}+u_{Tt}^{+}$$

with

$$u_{Tt}^{+}=u_{Tt-1}^{\#}-\delta_T'u_{gt}=u_{Tt}+O_p\left(\frac{1}{\sqrt{N}}+\sqrt{\frac{n}{N}}+\left|\beta_1+\gamma_1\right|^{t}\right), \tag{56}$$

gives the stated result. ■

9.3 Proof of Theorem 2

We start the proof with the following useful lemma concerning the limit behavior of the components of the panel regression (31).

9.3.1 Lemma A1

Let $\tilde{T}_{it}=T_{it}-\bar{T}_t,\ \tilde{R}_{it}=R_{it}-\bar{R}_t,\ \tilde{\delta}_i=\delta_i-\bar{\delta}$, and $\tilde{T}_{i,t,R}=\tilde{T}_{i,t}-\frac{\sum_{s=1}^{T}\sum_{j=1}^{N}\tilde{T}_{j,s}\tilde{R}_{j,s}}{\sum_{s=1}^{n}\sum_{j=1}^{N}\tilde{R}_{j,s}^2}\tilde{R}_{i,t}$ as in (30). Then as $(n,N)\to\infty$ the following limits hold.

(i) $\frac{1}{n^2}\sum_{t=1}^{n}\frac{1}{N}\sum_{i=1}^{N}\tilde{T}_{i,t}^2\Rightarrow\left(\frac{\beta_2}{1-\beta_1}\right)^2\operatorname{tr}\left\{\Sigma_r\int_0^1B_g\left(a\right)B_g\left(a\right)'da\right\}.$

(ii) $\frac{1}{n^2}\sum_{t=1}^{n}\frac{1}{N}\sum_{i=1}^{N}\tilde{R}_{i,t}^2\Rightarrow\operatorname{tr}\left\{\Sigma_r\int_0^1B_g\left(a\right)B_g\left(a\right)'da\right\}+\frac{1}{2}\omega_{uP}^2.$

(iii) $\frac{1}{n^2}\sum_{t=1}^{n}\frac{1}{N}\sum_{i=1}^{N}\tilde{T}_{i,t}\tilde{R}_{i,t}\Rightarrow\frac{\beta_2}{1-\beta_1}\operatorname{tr}\left\{\Sigma_r\int_0^1B_g\left(a\right)B_g\left(a\right)'da\right\}.$

(iv) $\frac{1}{n^2N}\sum_{t=1}^{n}\sum_{i=1}^{N}\tilde{T}_{i,t,R}^2\Rightarrow\frac{\left(\frac{1}{2}\omega_{uP}^2\right)\left(\frac{\beta_2}{1-\beta_1}\right)^2\operatorname{tr}\left\{\Sigma_r\int_0^1B_g(a)B_g(a)'da\right\}}{\operatorname{tr}\left\{\Sigma_r\int_0^1B_g(a)B_g(a)'da\right\}+\frac{1}{2}\omega_{uP}^2}.$

(v) $\frac{1}{n^2N}\sum_{t=1}^{n}\sum_{i=1}^{N}\tilde{R}_{i,t,T}^2\to_p\frac{1}{2}\sigma_u^2\omega_{uP}^2.$

9.3.2 Proof of Lemma A1

The components in (i) - (iv) arise in a least squares panel regression on the equation $T_{i,t+1}=\beta_1T_{i,t}+\beta_2R_{i,t}+\lambda_t+u_{it+1}$, where the component variables (T_{it},R_{it}) have the following explicit form from (58) and (59)

$$\begin{aligned}T_{it}&=\frac{\bar{\gamma}_0}{1-\beta_1}+\frac{\beta_2}{1-\beta_1}\delta_{ri}'U_{gt}+u_{Tit},\\R_{it}&=\left(R_{it}^0+\delta_{ri}'G_0+P_{i0}\right)+\delta_{ri}'U_{gt}+U_{it}^P,\end{aligned}$$

which lead to the representations

$$\begin{aligned}\tilde{T}_{it} &= \frac{\beta_2}{1-\beta_1}\tilde{\delta}_{ri}' U_{gt} + \tilde{u}_{Tit}, \\ \tilde{R}_{it} &= \left(\tilde{R}_{it}^0 + \tilde{\delta}_i G_0 + \tilde{P}_{i0}\right) + \tilde{\delta}_{ri}' U_{gt} + \tilde{U}_{it}^P.\end{aligned}$$

Primary interest in the development of regression regression asymptotics is in the components of the partitioned regression residuals

$$\tilde{T}_{j,t,R} = \tilde{T}_{j,t} - \frac{\sum_{t=1}^n \sum_{i=1}^N \tilde{T}_{i,t}\tilde{R}_{i,t}}{\sum_{t=1}^n \sum_{i=1}^N \tilde{R}_{i,t}^2}\tilde{R}_{j,t},$$

and the sample moment

$$\sum_{t=1}^n \sum_{i=1}^N \tilde{T}_{i,t,R}^2 = \sum_{t=1}^n \sum_{i=1}^N \tilde{T}_{i,t}^2 - \left(\sum_{t=1}^n \sum_{i=1}^N \tilde{T}_{i,t}\tilde{R}_{i,t}\right)^2 / \sum_{t=1}^n \sum_{i=1}^N \tilde{R}_{i,t}^2,$$

which we consider in turn.

Part (i): Using $\frac{1}{N}\sum_{i=1}^N \tilde{\delta}_{ri}\tilde{\delta}_{ri}' = \frac{1}{N}\sum_{i=1}^N \left(\delta_{ri} - \bar{\delta}_r\right)\left(\delta_{ri} - \bar{\delta}_r\right) \to_{a.s.} \Sigma_r$, the weak convergence $n^{-1/2}U_{g\lfloor n\cdot\rfloor} \Rightarrow B_g(\cdot)$, and continuous mapping we find that

$$\begin{aligned} & \frac{1}{n^2}\sum_{t=1}^n \frac{1}{N}\sum_{i=1}^N \tilde{T}_{i,t}^2 = \frac{1}{n^2}\sum_{t=1}^n \frac{1}{N}\sum_{i=1}^N \left(\frac{\bar{\gamma}_0}{1-\beta_1} + \frac{\beta_2}{1-\beta_1}\tilde{\delta}_{ri}' U_{gt} + u_{Tit}\right)^2 \\ = \; & \frac{1}{n^2}\sum_{t=1}^n \frac{1}{N}\sum_{i=1}^N \left(\frac{\beta_2}{1-\beta_1}\right)^2 \tilde{\delta}_{ri}' U_{gt} U_{gt}' \tilde{\delta}_{ri} + o_p(1) \\ = \; & \left(\frac{\beta_2}{1-\beta_1}\right)^2 \operatorname{tr}\left\{\Sigma_r \frac{1}{n^2}\sum_{t=1}^n \left(U_{gt}U_{gt}'\right)\right\} + \operatorname{tr}\left\{\frac{1}{N}\sum_{i=1}^N \left(\tilde{\delta}_{ri}\tilde{\delta}_{ri} - \Sigma_r\right)\frac{1}{n^2}\sum_{t=1}^n \left(U_{gt}U_{gt}'\right)\right\} + o_p(1) \\ \Rightarrow \; & \left(\frac{\beta_2}{1-\beta_1}\right)^2 \operatorname{tr}\left\{\Sigma_r \int_0^1 B_g(s) B_g(s)' ds\right\}. \end{aligned}$$

Part (ii): Proceeding in the same way we get

$$\begin{aligned} & \frac{1}{N}\sum_{i=1}^N \tilde{R}_{i,t}^2 = \frac{1}{N}\sum_{i=1}^N \left\{\tilde{\delta}_i U_{gt} + \tilde{U}_{it}^P + \left(\tilde{\delta}_i G_0 + \tilde{P}_{i0}\right)\right\}^2 \\ = \; & \operatorname{tr}\left\{\Sigma_r U_{gt} U_{gt}'\right\} + \frac{1}{N}\sum_{i=1}^N \left(U_{it}^P\right)^2 + O_p\left(\sqrt{n}\right) = \operatorname{tr}\left\{\Sigma_r U_{gt} U_{gt}'\right\} + \frac{1}{N}\sum_{i=1}^N \left(\sum_{k=1}^t u_{ik}^P\right)^2 + O_p\left(\sqrt{n}\right) \end{aligned}$$

$$= \quad \operatorname{tr}\left\{\Sigma_r U_{gt} U_{gt}'\right\} + \frac{1}{N}\sum_{i=1}^{N}\left(\sigma_{uP}^2 t + \sum_{k=1}^{t}\left[\left(u_{ik}^P\right)^2 - \sigma_{uP}^2\right] + 2\sum_{j>k}^{t} u_{ik}^P u_{ij}^P\right) + O_p\left(\sqrt{n}\right)$$

$$= \quad \operatorname{tr}\left\{\Sigma_r U_{gt} U_{gt}'\right\} + \frac{1}{N}\sum_{i=1}^{N}\left(\sigma_{uP}^2 t + \sum_{k=1}^{t}\left[\left(u_{ik}^P\right)^2 - \sigma_{uP}^2\right] + 2t\sum_{h=1}^{t-1}\sigma_{uP}(h) + 2\sum_{h=1}^{t-1}\sum_{k=1}^{t}\left(u_{ik}^P u_{ik+h}^P - \sigma_{uP}(h)\right)\right) +$$

$$= \quad \operatorname{tr}\left\{\Sigma_r U_{gt} U_{gt}'\right\} + t\left\{\sigma_{uP}^2 + 2\sum_{h=1}^{t-1}\sigma_{uP}(h)\right\} + O_p\left(\sqrt{n} + \frac{n}{\sqrt{N}}\right), \text{ uniformly in } t \le n.$$

Then

$$\frac{1}{n^2}\sum_{t=1}^{n}\frac{1}{N}\sum_{i=1}^{N}\tilde{R}_{i,t}^2 = \operatorname{tr}\left\{\Sigma_r \frac{1}{n^2}\sum_{t=1}^{n} U_{gt} U_{gt}'\right\} + \frac{1}{n^2}\sum_{t=1}^{n} t\left\{\sigma_{uP}^2 + 2\sum_{h=1}^{t-1}\sigma_{uP}(h)\right\} + O_p\left(\frac{1}{\sqrt{n}}\right)$$

$$= \quad \operatorname{tr}\left\{\Sigma_r \frac{1}{n^2}\sum_{t=1}^{n} U_{gt} U_{gt}'\right\} + \frac{1}{n^2}\sum_{t=1}^{n} t\left\{\sigma_{uP}^2 + 2\sum_{h=1}^{\infty}\sigma_{uP}(h)\right\} - \frac{2}{n^2}\sum_{t=1}^{n} t\sum_{h=t}^{\infty}\sigma_{uP}(h) + O_p\left(\frac{1}{\sqrt{n}}\right)$$

$$= \quad \operatorname{tr}\left\{\Sigma_r \frac{1}{n^2}\sum_{t=1}^{n} U_{gt} U_{gt}'\right\} + \frac{1}{2}\omega_{uP}^2 + o_p(1)$$

$$\Rightarrow \quad \operatorname{tr}\left\{\Sigma_r \int_0^1 B_g(s) B_g(s)' ds\right\} + \frac{1}{2}\omega_{uP}^2,$$

since, by A(iii), $t^{\varepsilon}\sum_{h=t}^{\infty}|\sigma_{uP}(h)| \to 0$ for some $\varepsilon > 0$ as $t \to \infty$ and thus

$$\left|n^{-2}\sum_{t=1}^{n} t\sum_{h=t}^{\infty}\sigma_{uP}(h)\right| \le n^{-2}\sum_{t=1}^{n} t^{1-\varepsilon}\left(t^{\varepsilon}\sum_{h=t}^{\infty}|\sigma_{uP}(h)|\right) = o\left(n^{-\varepsilon}\right).$$

Part (iii):

$$\frac{1}{n^2}\sum_{t=1}^{n}\frac{1}{N}\sum_{i=1}^{N}\tilde{T}_{i,t}\tilde{R}_{i,t} = \frac{1}{n^2}\sum_{t=1}^{n}\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\beta_2}{1-\beta_1}\tilde{\delta}_{ri}' U_{gt} + \tilde{u}_{Tit}\right)\left(\tilde{\delta}_{ri}' U_{gt} + U_{it}^P + \left(\tilde{\delta}_i G_0 + \tilde{P}_{i0}\right)\right)$$

$$= \quad \frac{1}{n^2}\sum_{t=1}^{n}\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\beta_2}{1-\beta_1}\tilde{\delta}_{ri}' U_{gt} U_{gt}'\tilde{\delta}_{ri} + \frac{\beta_2}{1-\beta_1}\tilde{\delta}_{ri}' U_{gt} U_{it}^P\right) + o_p(1)$$

$$= \quad \frac{\beta_2}{1-\beta_1}\operatorname{tr}\left\{\Sigma_r\left(\frac{1}{n^2}\sum_{t=1}^{n} U_{gt} U_{gt}'\right)\right\} + \frac{\beta_2}{1-\beta_1}\frac{1}{n^2}\sum_{t=1}^{n} U_{gt}\frac{1}{N}\sum_{i=1}^{N}\tilde{\delta}_{ir} U_{it}^P + o_p(1)$$

$$\Rightarrow \quad \frac{\beta_2}{1-\beta_1}\operatorname{tr}\left\{\Sigma_r \int_0^1 B_g(s) B_g(s)' ds\right\},$$

using $N^{-1}\sum_{i=1}^{N}\tilde{\delta}_{ri}\tilde{\delta}_{ri}' \rightarrow_{a.s.} \Sigma_r$, $n^{-2}\sum_{t=1}^{n}U_{gt}U_{gt}' \Rightarrow \int_0^1 B_g(s) B_g(s)' ds$, and

$$n^{-1}\sum_{t=1}^{n}\frac{U_{gt}}{\sqrt{n}}\frac{1}{\sqrt{n}}\sum_{s=1}^{t}\left(\frac{1}{N}\sum_{i=1}^{N}\left(\tilde{\delta}_{ir}u_{is}^{P}\right)\right) \rightarrow_p 0.$$

Part (iv): Combining parts (i)-(iii), which evidently also apply jointly, gives

$$\begin{aligned}
\frac{1}{n^2N}\sum_{t=1}^{n}\sum_{i=1}^{N}\tilde{T}_{i,t,R}^2 &= \frac{1}{n^2N}\sum_{t=1}^{n}\sum_{i=1}^{N}\tilde{T}_{i,t}^2 - \left(\frac{1}{n^2N}\sum_{t=1}^{n}\sum_{i=1}^{N}\tilde{T}_{i,t}\tilde{R}_{i,t}\right)^2 / \frac{1}{n^2N}\sum_{t=1}^{n}\sum_{i=1}^{N}\tilde{R}_{i,t}^2 \\
= & \left\{\left(\frac{\beta_2}{1-\beta_1}\right)^2 \operatorname{tr}\left\{\Sigma_r\left(\frac{1}{n^2}\sum_{t=1}^{n}U_{gt}U_{gt}'\right)\right\} + o_p(1)\right\} - \frac{\left\{\frac{\beta_2}{1-\beta_1}\operatorname{tr}\left\{\Sigma_r\left(\frac{1}{n^2}\sum_{t=1}^{n}U_{gt}U_{gt}'\right)\right\} + o_p(1)\right\}^2}{\operatorname{tr}\left\{\Sigma_r\left(\frac{1}{n^2}\sum_{t=1}^{n}U_{gt}U_{gt}'\right)\right\} + \frac{1}{2}\omega_{uP}^2 + o_p(1)} \\
= & \frac{\frac{1}{2}\omega_{uP}^2\left(\frac{\beta_2}{1-\beta_1}\right)^2 \operatorname{tr}\left\{\Sigma_r\left(\frac{1}{n^2}\sum_{t=1}^{n}U_{gt}U_{gt}'\right)\right\}}{\operatorname{tr}\left\{\Sigma_r\left(\frac{1}{n^2}\sum_{t=1}^{n}U_{gt}U_{gt}'\right)\right\} + \frac{1}{2}\omega_{uP}^2} + o_p(1) \\
& \Rightarrow \frac{\left(\frac{1}{2}\omega_{uP}^2\right)\left(\frac{\beta_2}{1-\beta_1}\right)^2 \operatorname{tr}\left\{\Sigma_r\int_0^1 B_g(s)B_g(s)'ds\right\}}{\operatorname{tr}\left\{\Sigma_r\int_0^1 B_g(s)B_g(s)'ds\right\} + \frac{1}{2}\omega_{uP}^2},
\end{aligned} \tag{57}$$

as required.

Part (v): Combining parts (i)-(iii) again, gives

$$\begin{aligned}
\frac{1}{n^2N}\sum_{t=1}^{n}\sum_{i=1}^{N}\tilde{R}_{i,t,T}^2 &= \frac{1}{n^2N}\sum_{t=1}^{n}\sum_{i=1}^{N}\tilde{R}_{i,t}^2 - \left(\frac{1}{n^2N}\sum_{t=1}^{n}\sum_{i=1}^{N}\tilde{T}_{i,t}\tilde{R}_{i,t}\right)^2 / \frac{1}{n^2N}\sum_{t=1}^{n}\sum_{i=1}^{N}\tilde{T}_{i,t}^2 \\
= & \left\{\operatorname{tr}\left\{\Sigma_r\left(\frac{1}{n^2}\sum_{t=1}^{n}U_{gt}U_{gt}'\right)\right\} + \frac{1}{2}\omega_{uP}^2 + o_p(1)\right\} - \frac{\left\{\frac{\beta_2}{1-\beta_1}\operatorname{tr}\left\{\Sigma_r\left(\frac{1}{n^2}\sum_{t=1}^{n}U_{gt}U_{gt}'\right)\right\} + o_p(1)\right\}^2}{\left\{\left(\frac{\beta_2}{1-\beta_1}\right)^2 \operatorname{tr}\left\{\Sigma_r\left(\frac{1}{n^2}\sum_{t=1}^{n}U_{gt}U_{gt}'\right)\right\} + o_p(1)\right\}} \\
= & \frac{\frac{1}{2}\omega_{uP}^2\left(\frac{\beta_2}{1-\beta_1}\right)^2 \operatorname{tr}\left\{\Sigma_r\left(\frac{1}{n^2}\sum_{t=1}^{n}U_{gt}U_{gt}'\right)\right\}}{\left(\frac{\beta_2}{1-\beta_1}\right)^2 \operatorname{tr}\left\{\Sigma_r\left(\frac{1}{n^2}\sum_{t=1}^{n}U_{gt}U_{gt}'\right)\right\}} + o_p(1) \\
\rightarrow_p & \frac{1}{2}\omega_{uP}^2,
\end{aligned}$$

which is constant. ■

With these results in hand, we continue with the proof of Theorem 2. We start by considering the covariate inputs into the transient dynamic equation (1). From (15) the energy balance equation yields

$$\lambda_t = \gamma_0 + \gamma_1 \overline{T}_t + \gamma_2 \overline{R}_t + \gamma_3 \ln(CO_{2,t}) = \bar{\gamma}_0 + u_{\lambda t},$$

where $u_{\lambda t} = \gamma_1 u_{Tt} + \gamma_3 u_{ct} + o_p(1)$ is asymptotically stationary. From (4), (5) and Assumption B(iii) we have

$$R_{it} = R_{it}^0 + \delta_{ri}' G_t + P_{it} = R_{it}^0 + \left(\delta_{ri}' G_0 + P_{i0}\right) + \delta_{ri}' U_{gt} + U_{it}^P =: \delta_{ri}^0 + \delta_{ri}' U_{gt} + U_{it}^P. \quad (58)$$

With these inputs, solving (1) gives

$$\begin{aligned} T_{i,t+1} &= \beta_2 \sum_{j=0}^{t} \beta_1^j R_{i,t} + \sum_{j=0}^{t} \beta_1^j \lambda_{t-j} + \sum_{j=0}^{t} \beta_1^j u_{it+1} \\ &= \beta_2 \sum_{j=0}^{t} \beta_1^j R_{i,t-j}^0 + \beta_2 \delta_{ri}' \sum_{j=0}^{t} \beta_1^j U_{gt-j} + \beta_2 \sum_{j=0}^{t} \beta_1^j U_{it-j}^P + \sum_{j=0}^{t} \beta_1^j \left(\bar{\gamma}_0 + u_{\lambda t-j}\right) + \sum_{j=0}^{t} \beta_1^j u_{it+1-j} \end{aligned}$$

$$\begin{aligned} &= \frac{\bar{\gamma}_0}{1-\beta_1} + \beta_2 \delta_{ri}' \sum_{j=0}^{t} \beta_1^j U_{gt-j} + \beta_2 \sum_{j=0}^{t} \beta_1^j U_{it-j}^P + \sum_{j=0}^{\infty} \beta_1^j \left(u_{\lambda t-j} + u_{it+1-j} + \beta_2 R_{i,t-j}^0\right) + O_p\left(|\beta_1|^t\right) \\ &= \frac{\bar{\gamma}_0}{1-\beta_1} + \frac{\beta_2}{1-\beta_1} \delta_{ri}' U_{gt} + \frac{\beta_2}{1-\beta_1} U_{it}^P + v_{Tit+1}, \end{aligned} \quad (59)$$

where

$$v_{Tit+1} = u_{it+1} + \sum_{j=0}^{\infty} \beta_1^j \left(u_{\lambda t-j} + \beta_1 u_{it-j} + \beta_2 R_{i,t-j}^0 - \frac{\beta_1}{1-\beta_1} u_{gt-j} - \frac{\beta_1}{1-\beta_1} u_{it-k}^P \right) + O_p\left(|\beta_1|^t\right) \quad (60)$$

is asymptotically stationary. Line (59) follows because, just as in the derivation of (53),

$$\begin{aligned} \sum_{j=0}^{t} \beta_1^j U_{gt-j} &= \sum_{j=0}^{t} \beta_1^j \sum_{s=1}^{t-j} u_{gs} = \sum_{s=1}^{t} u_{gs} \sum_{j=0}^{t-s} \beta_1^j = \sum_{s=1}^{t} u_{gs} \frac{1-\beta_1^{t-s+1}}{1-\beta_1} \\ &= \frac{1}{1-\beta_1} \sum_{s=1}^{t} u_{gs} - \frac{1}{1-\beta_1} \sum_{s=1}^{t} u_{gs} \beta_1^{t-s+1} = \frac{1}{1-\beta_1} U_{gt} - \frac{\beta_1}{1-\beta_1} \sum_{k=0}^{t-1} u_{gt-k} \beta_1^k \\ &= \frac{1}{1-\beta_1} U_{gt} - \frac{\beta_1}{1-\beta_1} \sum_{k=0}^{\infty} u_{gt-k} \beta_1^k + O_p\left(|\beta_1|^t\right), \end{aligned}$$

and in a similar way

$$\begin{aligned} \sum_{j=0}^{t} \beta_1^j U_{it-j}^P &= \sum_{j=0}^{t} \beta_1^j \sum_{s=1}^{t-j} u_{is}^P = \sum_{s=1}^{t} u_{is}^P \sum_{j=0}^{t-s} \beta_1^j = \sum_{s=1}^{t} u_{is}^P \frac{1-\beta_1^{t-s+1}}{1-\beta_1} \\ &= \frac{1}{1-\beta_1} U_{it}^P - \frac{\beta_1}{1-\beta_1} \sum_{k=0}^{\infty} u_{it-k}^P \beta_1^k + O_p\left(|\beta_1|^t\right). \end{aligned}$$

Again with notation $\tilde{T}_{it} = T_{it} - \bar{T}_t$, $\tilde{R}_{it} = R_{it} - \bar{R}_t$, and $\tilde{\delta}_{ir} = \delta_{ir} - \bar{\delta}_r$, we have

$$\begin{aligned} T_{it} &= \frac{\bar{\gamma}_0}{1-\beta_1} + \frac{\beta_2}{1-\beta_1}\delta'_{ri}U_{gt-1} + v_{Tit} = \frac{\bar{\gamma}_0}{1-\beta_1} + \frac{\beta_2}{1-\beta_1}\delta'_{ri}U_{gt} + v_{Tit} - \frac{\beta_2}{1-\beta_1}\delta'_{ri}u_{gt} \\ &= \frac{\bar{\gamma}_0}{1-\beta_1} + \frac{\beta_2}{1-\beta_1}\delta'_{ri}U_{gt} + u_{Tit} \end{aligned} \tag{61}$$

where $u_{Tit} := v_{Tit} - \frac{\beta_2}{1-\beta_1}\delta'_{ri}u_{gt}$. From (58) and (61), we have the following explicit expressions for $\left(\tilde{T}_{it}, \tilde{R}_{it}\right)$

$$\tilde{T}_{it} = \frac{\beta_2}{1-\beta_1}\tilde{\delta}'_{ri}U_{gt} + \tilde{u}_{Tit}, \tag{62}$$

$$\tilde{R}_{it} = \tilde{\delta}^0_{ri} + \tilde{\delta}'_{ri}U_{gt} + \tilde{U}^P_{it}. \tag{63}$$

Define $\tilde{X}_{i,t} = \left(\tilde{T}_{it}, \tilde{R}_{it}\right)'$ and write (62)-(63) in vector form as

$$\tilde{X}_{it} = \begin{bmatrix} \tilde{T}_{it} \\ \tilde{R}_{it} \end{bmatrix} = \begin{bmatrix} \frac{\beta_2}{1-\beta_1}\tilde{\delta}'_{ri}U_{gt} + \tilde{u}_{Tit} \\ \tilde{\delta}^0_{ri} + \tilde{\delta}'_{ri}U_{gt} + \tilde{U}^P_{it} \end{bmatrix} = \tilde{\delta}^0_{xi} + \delta_x\tilde{\delta}'_{ri}U_{gt} + \tilde{u}_{xit},$$

where $\tilde{\delta}^0_{xi} = \left(0, \tilde{\delta}^0_{ri}\right)'$, $\delta_x = \left(\frac{\beta_2}{1-\beta_1}, 1\right)'$, and $\tilde{u}_{xit} = \left(\tilde{u}_{Tit}, \tilde{U}^P_{it}\right)$. Importantly, we note that

$$\left(1-\beta_1\right)\tilde{T}_{it} - \beta_2\tilde{R}_{it} = \left(1-\beta_1\right)\tilde{\delta}^0_{ri} + \left(1-\beta_1\right)\tilde{u}_{Tit} - \beta_2\tilde{U}^P_{it} \tag{64}$$

which corresponds to the cointegrating relation for the aggregate time series shown in (21)

$$\left(\beta_1 - 1\right)\delta'_T + \beta_2\delta'_r = 0.$$

But in the panel data case the linear combination $\left(1-\beta_1\right)\tilde{T}_{it} - \beta_2\tilde{R}_{it}$ also involves the nonstationary component $\tilde{U}^P_{it} = \sum_{k=1}^t \tilde{u}^P_{ik}$ when station-level trends are present in R_{it}.

9.3.3 Proof of Part (a)

To establish Part (a) of the Theorem 2, we begin by developing limit theory for the sample covariance element

$$\begin{aligned} \frac{1}{n\sqrt{N}}\sum_{t=1}^n\sum_{i=1}^N \tilde{T}_{it}u_{it+1} &= \frac{1}{n\sqrt{N}}\sum_{t=1}^n\sum_{i=1}^N\left[\frac{\beta_2}{1-\beta_1}\tilde{\delta}'_{ri}U_{gt} + \tilde{u}_{Tit}\right]u_{it+1} \\ &= \frac{\beta_2}{1-\beta_1}\frac{1}{n}\sum_{t=1}^n\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N\tilde{\delta}'_{ri}u_{it+1}\right)U_{gt} + \frac{1}{\sqrt{N}}\sum_{i=1}^N\left(\frac{1}{n}\sum_{t=1}^n\tilde{u}_{Tit}u_{it+1}\right) \end{aligned}$$

$$= \frac{\beta_2}{1-\beta_1}\frac{1}{n}\sum_{t=1}^{n}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\tilde{\delta}_{ri}'u_{it+1}\right)U_{gt}+o_p(1) \tag{65}$$

$$= \frac{\beta_2}{1-\beta_1}\frac{1}{n}\sum_{t=1}^{n}\xi_{rt+1}'U_{gt}+o_p(1) \text{ as } N\to\infty \tag{66}$$

$$\Rightarrow \frac{\beta_2}{1-\beta_1}\int_0^1 dB_{\xi_r}'(a)\,B_g(a) \equiv \frac{\beta_2}{1-\beta_1}\mathcal{MN}\left(0,\sigma_u^2\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(a)\,B_g(a)'\,da\right\}\right) \tag{67}$$

The $o_p(1)$ error in (65) holds because $\mathbb{E}(\tilde{u}_{Tit}u_{it+1}) = 0$ and $\frac{1}{n}\sum_{t=1}^{n}\tilde{u}_{Tit}u_{it+1} \to_p 0$. The stochastic integral limit in (67) holds because: (i) by the martingale CLT $\xi_{rN} := \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\tilde{\delta}_{ri}u_{it+1} \Rightarrow \xi_{rt+1} \sim_{iid} \mathcal{N}(0,\sigma_u^2\Sigma_r)$ over t, where we can replace the weak convergence by convergence in probability in a suitably defined probability space at this point in the argument without affecting the final weak convergence result (67); (ii) by virtue of A(i), $\left\{\tilde{\delta}_{ri},u_{it+1}\right\}$ is independent of U_{gs} for all (i,t,s); and (iii) $\sum_{t=1}^{n}\frac{\xi_{r,t+1}'}{\sqrt{n}}\frac{U_{gt}}{n} \Rightarrow \int_0^1 dB_{\xi_r}'(s)\,B_g(s)$ by martingale convergence to a stochastic integral (Ibragimov and Phillips, 2008). Note that the square bracket (conditional variance) matrix of $\int_0^1 dB_{\xi_r}'B_g(r)$ is

$$\left[\int_0^1 dB_{\xi_r}'B_g(r)\right] = \sigma_u^2\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)\,B_g(s)'\,ds\right\},$$

corresponding to the variance mixture process in (67).

The second sample covariance element is

$$\frac{1}{n\sqrt{N}}\sum_{t=1}^{n}\sum_{i=1}^{N}\tilde{R}_{i,t}u_{it+1} = \frac{1}{n\sqrt{N}}\sum_{t=1}^{n}\sum_{i=1}^{N}\left[\tilde{\delta}_{ri}^0+\tilde{\delta}_{ri}'U_{gt}+\tilde{U}_{it}^P\right]u_{it+1}$$

$$= \frac{1}{n}\sum_{t=1}^{n}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\tilde{\delta}_{ri}'u_{it+1}\right)U_{gt}+\frac{1}{n\sqrt{N}}\sum_{t=1}^{n}\sum_{i=1}^{N}\tilde{U}_{it}^P u_{it+1}+\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\tilde{\delta}_{ri}^0\left(\frac{1}{n}\sum_{t=1}^{n}u_{it+1}\right)$$

$$= \frac{1}{n}\sum_{t=1}^{n}\xi_{rt+1}'U_{gt}+\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left\{\frac{1}{n}\sum_{t=1}^{n}\left(\sum_{k=1}^{t}\tilde{u}_{ik}^P\right)u_{it+1}\right\}+o_p(1) \tag{68}$$

$$= \frac{1}{n}\sum_{t=1}^{n}\xi_{rt+1}'U_{gt}+\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\int_0^1 B_{u_i}^P(s)\,dB_{u_i}(s)\right)+o_p(1) \tag{69}$$

$$= \frac{1}{n}\sum_{t=1}^{n}\xi_{rt+1}'U_{gt}+\eta_P+o_p(1) \tag{70}$$

$$\Rightarrow \int_0^1 dB_{\xi_r}'(s)\,B_g(s)+\eta_P \equiv \mathcal{MN}\left(0,\sigma_u^2\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)\,B_g(s)'\,ds\right\}+\frac{1}{2}\sigma_u^2\omega_{uP}^2\right) \tag{71}$$

The $o_p(1)$ errors in (68) - (70) hold for the same reasons as in (65) and (66) above; and by the Lindeberg Lévy CLT $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\int_0^1 B_{u_i}^P(s)\,dB_{u_i}(s)\right) \Rightarrow \eta_P \equiv \mathcal{N}\left(0,\frac{1}{2}\sigma_u^2\omega_{uP}^2\right)$ because

$\mathbb{E}\left(\int_0^1 B_{u_i}^P(s)\,dB_{u_i}(s)\right) = 0$ and

$$\sigma_u^2 \mathbb{E}\left(\int_0^1 B_{u_i}^P(s)\,dB_{u_i}(s)\,ds\right)^2 = \sigma_u^2\left(\int_0^1 \mathbb{E}\left(B_{u_i}^P(s)\right)^2 ds\right) = \frac{1}{2}\sigma_u^2\omega_{uP}^2.$$

Finally, we note that $\int_0^1 dB_{\xi_r}'(s)\,B_g(s)$ and η_P are independent, leading to the mixed normal limit (71) because B_{ξ_r}' is independent of η_P and both are independent of B_g. The latter holds by assumption A(i) and the former because

$$\begin{aligned}
& \mathbb{E}\left\{\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\tilde{\delta}_{ri}' u_{it+1}\right)\times\frac{1}{\sqrt{N}}\sum_{j=1}^{N}\left\{\frac{1}{n}\sum_{s=1}^{n}\left(\sum_{k=1}^{s}\tilde{u}_{jk}^P\right)u_{js+1}\right\}\right\} \\
= \;& \mathbb{E}\left\{\frac{1}{n^{3/2}}\sum_{t=1}^{n}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}u_{it+1}^2\tilde{\delta}_{ri}'\right)\times\frac{1}{\sqrt{N}}\left(\sum_{k=1}^{t}\tilde{u}_{ik}^P\right)\right\} \\
= \;& \mathbb{E}\left\{\frac{1}{n^{3/2}}\sum_{t=1}^{n}\left(\frac{1}{N}\sum_{i=1}^{N}u_{it+1}^2\left(\sum_{k=1}^{t}\tilde{u}_{ik}^P\right)\tilde{\delta}_{ri}'\right)\right\} \\
= \;& \frac{1}{n^{3/2}}\sum_{t=1}^{n}\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left\{u_{it+1}^2\left(\sum_{k=1}^{t}\tilde{u}_{ik}^P\right)\tilde{\delta}_{ri}'\right\} \\
= \;& \frac{1}{n^{3/2}}\sum_{t=1}^{n}\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left(u_{it+1}^2\right)\mathbb{E}\left(\sum_{k=1}^{t}\tilde{u}_{ik}^P\right)\mathbb{E}\left(\tilde{\delta}_{ri}'\right) = 0.
\end{aligned}$$

We are now ready to consider the sample covariance. In the argument that follows we proceed as earlier, using $\frac{1}{N}\sum_{i=1}^{N}\tilde{\delta}_{ri}\tilde{\delta}_{ri}' \to_{a.s.} \Sigma_r$ and where convenient replacing weak convergence by convergence in probability in a suitably defined probability space without affecting the final weak convergence result in the original space. Thus

$$\begin{aligned}
& \frac{1}{n\sqrt{N}}\sum_{t=1}^{n}\sum_{i=1}^{N}\tilde{T}_{i,t,R}u_{it+1} = \frac{1}{n\sqrt{N}}\sum_{t=1}^{n}\sum_{i=1}^{N}\left\{\tilde{T}_{i,t} - \frac{\frac{\beta_2}{1-\beta_1}\mathrm{tr}\left\{\Sigma_r\left(\frac{1}{n^2}\sum_{t=1}^{n}U_{gt}U_{gt}'\right)\right\} + o_p(1)}{\mathrm{tr}\left\{\Sigma_r\left(\frac{1}{n^2}\sum_{t=1}^{n}U_{gt}U_{gt}'\right)\right\} + \frac{1}{2}\omega_{uP}^2 + o_p(1)}\tilde{R}_{i,t}\right\}u_{it+1} \\
= \;& \frac{1}{n\sqrt{N}}\sum_{t=1}^{n}\sum_{i=1}^{N}\left\{\tilde{T}_{i,t} - \frac{\frac{\beta_2}{1-\beta_1}\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)\,B_g(s)'\,ds\right\}}{\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)\,B_g(s)'\,ds\right\} + \frac{1}{2}\omega_{uP}^2}\tilde{R}_{i,t}\right\}u_{it+1} + o_p(1) \\
= \;& \frac{1}{n\sqrt{N}}\sum_{t=1}^{n}\sum_{i=1}^{N}\left\{\left[\frac{\beta_2}{1-\beta_1}\tilde{\delta}_{ri}'U_{gt} + \tilde{u}_{Tit}\right]\right. \\
& \left. -\frac{\frac{\beta_2}{1-\beta_1}\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)\,B_g(s)'\,ds\right\}}{\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)\,B_g(s)'\,ds\right\} + \frac{1}{2}\omega_{uP}^2}\left[\left[\tilde{\delta}_{ri}^0 + \tilde{\delta}_{ri}'U_{gt} + \tilde{U}_{it}^P\right]\right]\right\}u_{it+1} + o_p(1)
\end{aligned}$$

$$= \frac{1}{n\sqrt{N}}\sum_{t=1}^{n}\sum_{i=1}^{N}\left\{\frac{\beta_2}{1-\beta_1}\left[1-\frac{\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)B_g(s)'ds\right\}}{\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)B_g(s)'ds\right\}+\frac{1}{2}\omega_{uP}^2}\right]\tilde{\delta}_{ri}'u_{it+1}U_{gt}\right.$$
$$\left.-\frac{\frac{\beta_2}{1-\beta_1}\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)B_g(s)'ds\right\}}{\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)B_g(s)'ds\right\}+\frac{1}{2}\omega_{uP}^2}\tilde{U}_{it}^P u_{it+1}\right\}+o_p(1)$$
$$= \frac{\frac{\beta_2}{1-\beta_1}}{\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)B_g(s)'ds\right\}+\frac{1}{2}\omega_{uP}^2}\frac{1}{n\sqrt{N}}\sum_{t=1}^{n}\sum_{i=1}^{N}\left\{\left(\frac{1}{2}\omega_{uP}^2\right)\tilde{\delta}_{ri}'u_{it+1}U_{gt}\right.$$
$$\left.-\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)B_g(s)'ds\right\}\tilde{U}_{it}^P u_{it+1}\right\}+o_p(1)$$

$$\sim_a \frac{\frac{\beta_2}{1-\beta_1}}{\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)B_g(s)'ds\right\}+\frac{1}{2}\omega_{uP}^2}\left\{\left[\left(\frac{1}{2}\omega_{uP}^2\right)\int_0^1 dB_{\xi_r}'(s)B_g(s)\right]\right.$$
$$\left.-\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)B_g(s)'ds\right\}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\int_0^1 B_{u_i}^P(s)dB_{u_i}(s)\right)\right\} \tag{72}$$
$$\sim_a \frac{\frac{\beta_2}{1-\beta_1}}{\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)B_g(s)'ds\right\}+\frac{1}{2}\omega_{uP}^2}$$
$$\times\mathcal{MN}\left(0,\sigma_u^2\left[\left(\frac{1}{2}\omega_{uP}^2\right)^2\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)B_g(s)'ds\right\}+\frac{1}{2}\omega_{uP}^2\left(\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)B_g(s)'ds\right\}\right)^2\right]\right)$$

$$= \frac{\frac{\beta_2}{1-\beta_1}\sigma_u\left\{\left(\frac{1}{2}\omega_{uP}^2\right)\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)B_g(s)'ds\right\}\right\}^{1/2}}{\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)B_g(s)'ds\right\}+\frac{1}{2}\omega_{uP}^2}\times\mathcal{MN}\left(0,\frac{1}{2}\omega_{uP}^2+\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)B_g(s)'ds\right\}\right)$$
$$= \mathcal{MN}\left(0,\frac{\left(\frac{\beta_2}{1-\beta_1}\right)^2\sigma_u^2\left(\frac{1}{2}\omega_{uP}^2\right)\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)B_g(s)'ds\right\}}{\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)B_g(s)'ds\right\}+\frac{1}{2}\omega_{uP}^2}\right). \tag{73}$$

The limit theory for $\hat{\beta}_1$ now follows. The estimation error is

$$\hat{\beta}_1-\beta_1=\frac{\sum_{t=1}^{n}\sum_{i=1}^{N}\tilde{T}_{i,t,R}u_{it+1}}{\sum_{t=1}^{T}\sum_{i=1}^{N}\tilde{T}_{i,t,R}^2}, \tag{74}$$

and it has already been shown in part (iv) of Lemma A1 that

$$\frac{1}{n^2N}\sum_{t=1}^{n}\sum_{i=1}^{N}\tilde{T}_{i,t,R}^2\Rightarrow\frac{\left(\frac{1}{2}\omega_{uP}^2\right)\left(\frac{\beta_2}{1-\beta_1}\right)^2\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)B_g(s)'ds\right\}}{\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)B_g(s)'ds\right\}+\frac{1}{2}\omega_{uP}^2}. \tag{75}$$

Thus, using (73) and (75) we find that

$$\frac{\frac{1}{n\sqrt{N}}\sum_{t=1}^{n}\sum_{i=1}^{N}\tilde{T}_{i,t,R}u_{it+1}}{\frac{1}{n^2N}\sum_{t=1}^{n}\sum_{i=1}^{N}\tilde{T}_{i,t,R}^2} \Rightarrow \frac{\mathcal{MN}\left(0, \frac{\left(\frac{\beta_2}{1-\beta_1}\right)^2\sigma_u^2\left(\frac{1}{2}\omega_{uP}^2\right)\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)B_g(s)'ds\right\}}{\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)B_g(s)'ds\right\}+\frac{1}{2}\omega_{uP}^2}\right)}{\frac{\left(\frac{1}{2}\omega_{uP}^2\right)\left(\frac{\beta_2}{1-\beta_1}\right)^2\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)B_g(s)'ds\right\}}{\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)B_g(s)'ds\right\}+\frac{1}{2}\omega_{uP}^2}}$$

$$= \mathcal{MN}\left(0, \sigma_u^2\frac{\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)\,B_g(s)'\,ds\right\}+\frac{1}{2}\omega_{uP}^2}{\left(\frac{1}{2}\omega_{uP}^2\right)\left(\frac{\beta_2}{1-\beta_1}\right)^2\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)\,B_g(s)'\,ds\right\}}\right). \tag{76}$$

The required result now follows from (74) and (76), viz.,

$$\sqrt{n^2N}\left(\hat{\beta}_1-\beta_1\right) = \frac{\frac{1}{n\sqrt{N}}\sum_{t=1}^{n}\sum_{i=1}^{N}\tilde{T}_{i,t,R}u_{it+1}}{\frac{1}{n^2N}\sum_{t=1}^{T}\sum_{i=1}^{N}\tilde{T}_{i,t,R}^2}$$

$$\sim_a \left(\frac{\left(\frac{1}{2}\omega_{uP}^2\right)\left(\frac{\beta_2}{1-\beta_1}\right)^2\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)\,B_g(s)'\,ds\right\}}{\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)\,B_g(s)'\,ds\right\}+\frac{1}{2}\omega_{uP}^2}\right)^{-1}\left\{\frac{\frac{\beta_2}{1-\beta_1}}{\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)\,B_g(s)'\,ds\right\}+\frac{1}{2}\omega_{uP}^2}\right\}$$

$$\times\left\{\left(\frac{1}{2}\omega_{uP}^2\right)\int_0^1 dB_{\xi_r}'(s)\,B_g(s)-\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)\,B_g(s)'\,ds\right\}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\int_0^1 B_{u_i}^P(s)\,dB_{u_i}(s)\right)\right\}$$

$$\Rightarrow \frac{\left(\frac{1}{2}\omega_{uP}^2\right)\int_0^1 dB_{\xi_r}'(s)\,B_g(s)-\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)\,B_g(s)'\,ds\right\}\eta_P}{\left(\frac{1}{2}\omega_{uP}^2\right)\left(\frac{\beta_2}{1-\beta_1}\right)\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)\,B_g(s)'\,ds\right\}} \tag{77}$$

$$= \mathcal{MN}\left(0, \frac{\frac{\sigma_u^2\left(\frac{1}{2}\omega_{uP}^2\right)^2\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)B_g(s)'ds\right\}+\frac{1}{2}\sigma_u^2\omega_{uP}^2\left[\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)B_g(s)'ds\right\}\right]^2}{\left(\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)B_g(s)'ds\right\}+\frac{1}{2}\omega_{uP}^2\right)^2}}{\left(\frac{1}{2}\omega_{uP}^2\right)^2\left(\frac{\beta_2}{1-\beta_1}\right)^2\left(\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)\,B_g(s)'\,ds\right\}\right)^2}\right)$$

$$\equiv \mathcal{MN}\left(0, \sigma_u^2\frac{\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)\,B_g(s)'\,ds\right\}+\frac{1}{2}\omega_{uP}^2}{\left(\frac{1}{2}\omega_{uP}^2\right)\left(\frac{\beta_2}{1-\beta_1}\right)^2\mathrm{tr}\left\{\Sigma_r\int_0^1 B_g(s)\,B_g(s)'\,ds\right\}}\right), \tag{78}$$

which uses the independence of $\int_0^1 dB_{\xi_r}'(s)\,B_g(s)$ and η_P established earlier. The representation (77) is useful later and the final line (78) delivers the required result. ■

9.3.4 Proof of Part (b)

To establish Part (b), we note first that

$$\sqrt{n^2 N}\left(\hat{\beta}_2-\beta_2\right)=\frac{\frac{1}{n\sqrt{N}}\sum_{t=1}^{n}\sum_{i=1}^{N}\tilde{R}_{i,t,T}u_{it+1}}{\frac{1}{n^2 N}\sum_{t=1}^{T}\sum_{i=1}^{N}\tilde{R}_{i,t,T}^2},$$

where $\tilde{R}_{i,t,T}=\tilde{R}_{i,t}-\frac{\sum_{s=1}^{T}\sum_{j=1}^{N}\tilde{T}_{j,s}\tilde{R}_{j,s}}{\sum_{s=1}^{n}\sum_{j=1}^{N}\tilde{T}_{j,s}^2}\tilde{T}_{j,t}$. Using similar arguments to those in the proof of Part (a), we find the following limit behavior of the sample covariance

$$\begin{aligned}
&\frac{1}{n\sqrt{N}}\sum_{t=1}^{n}\sum_{i=1}^{N}\tilde{R}_{i,t,T}u_{it+1}=\frac{1}{n\sqrt{N}}\sum_{t=1}^{n}\sum_{i=1}^{N}\left\{\tilde{R}_{i,t}-\frac{\frac{\beta_2}{1-\beta_1}\mathrm{tr}\left\{\Sigma_r\left(\frac{1}{n^2}\sum_{t=1}^{n}U_{gt}U_{gt}'\right)\right\}+o_p(1)}{\left(\frac{\beta_2}{1-\beta_1}\right)^2\mathrm{tr}\left\{\Sigma_r\frac{1}{n^2}\sum_{t=1}^{n}\left(U_{gt}U_{gt}'\right)\right\}+o_p(1)}\tilde{T}_{i,t}\right\}u_{it+1}\\
=&\ \frac{1}{n\sqrt{N}}\sum_{t=1}^{n}\sum_{i=1}^{N}\left\{\left[\left[\tilde{\delta}_{ri}^0+\tilde{\delta}_{ri}'U_{gt}+\tilde{U}_{it}^P\right]\right]-\left(\frac{\beta_2}{1-\beta_1}\right)^{-1}\left[\frac{\beta_2}{1-\beta_1}\tilde{\delta}_{ri}'U_{gt}+\tilde{u}_{Tit}\right]\right\}u_{it+1}+o_p(1)\\
=&\ \frac{1}{n\sqrt{N}}\sum_{t=1}^{n}\sum_{i=1}^{N}\tilde{U}_{it}^P u_{it+1}+o_p(1)=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\sum_{t=1}^{n}\frac{\tilde{U}_{it}^P}{\sqrt{n}}\frac{u_{it+1}}{\sqrt{n}}\right)+o_p(1)\\
=&\ \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\int_0^1 B_{ui}^P(s)\,dB_{ui}(s)+o_p(1)\Rightarrow\mathcal{N}\left(0,\frac{1}{2}\sigma_u^2\omega_{uP}^2\right),
\end{aligned}\tag{79}$$

since $\int_0^1 B_{ui}^P(s)\,dB_{ui}(s)\sim_{iid}\left(0,\mathbb{E}\left(\int_0^1 B_{ui}^P(s)\,dB_{ui}(s)\right)^2\right)=\left(0,\frac{1}{2}\sigma_u^2\omega_{uP}^2\right)$ and (79) follows by the Lindeberg Lévy CLT. The stated result follows from (79) and Lemma A1(v) viz.,

$$\sqrt{n^2 N}\left(\hat{\beta}_2-\beta_2\right)=\frac{\frac{1}{n\sqrt{N}}\sum_{t=1}^{n}\sum_{i=1}^{N}\tilde{R}_{i,t,T}u_{it+1}}{\frac{1}{n^2 N}\sum_{t=1}^{T}\sum_{i=1}^{N}\tilde{R}_{i,t,T}^2}\Rightarrow\frac{1}{\frac{1}{2}\omega_{uP}^2}\mathcal{N}\left(0,\frac{1}{2}\sigma_u^2\omega_{uP}^2\right)=\mathcal{N}\left(0,2\frac{\sigma_u^2}{\omega_{uP}^2}\right).$$

■

9.4 Proof of Theorem 3

Part (a): We start by tackling the degeneracy in the limit of the signal matrix $\sum_{t=1}^{n}\widetilde{\widetilde{W}}_t\widetilde{\widetilde{W}}_t'$ in the cointegrating regression (38). Recall that the regressor vector $\bar{W}_t=\delta_w+\beta_w U_{wt}+u_{wt}$ and so $\widetilde{\widetilde{W}}_t=\beta_w\tilde{U}_{wt}+\tilde{u}_{wt}$, where we use the notation $\widetilde{\bar{A}}_t=\bar{A}_t-n^{-1}\sum_{t=1}^{n}\bar{A}_t$. By assumption B(ii), a standardized form of the partial sum process $U_{wt}=a'U_{gt}$ satisfies the functional law

$$n^{-1/2}U_{w\lfloor nr\rfloor}=n^{-1/2}a'U_{g\lfloor nr\rfloor}\Rightarrow a'B_g(r)=:B_w(r),$$

where a is defined in (23). Then

$$\frac{1}{n^2}\sum_{t=1}^{n}\left(\bar{W}_t-\delta_w\right)\left(\bar{W}_t-\delta_w\right)'=\frac{1}{n}\sum_{t=1}^{n}\frac{\bar{W}_t-\delta_w}{\sqrt{n}}\frac{\left(\bar{W}_t-\delta_w\right)'}{\sqrt{n}}\Rightarrow\beta_w\beta_w'\int_0^1 B_w\left(r\right)^2dr, \quad (80)$$

and, using the notation $\tilde{A}\left(s\right)=A\left(s\right)-\int_0^1 A\left(p\right)dp$, we have

$$\begin{aligned} n^{-2}\sum_{t=1}^{n}\widetilde{\bar{W}}_t\widetilde{\bar{W}}_t' &= \beta_w\beta_w'\left\{\frac{1}{n}\sum_{t=1}^{n}\frac{a'\bar{U}_{gt}}{\sqrt{n}}\frac{\bar{U}_{gt}'a}{\sqrt{n}}+o_p\left(1\right)\right\} \\ \Rightarrow \quad \beta_w\beta_w'\int_0^1\tilde{B}_w\left(s\right)^2ds &= \beta_w\beta_w'\left(\int_0^1 a'\tilde{B}_g\left(s\right)\tilde{B}_g\left(s\right)'ads\right), \end{aligned}$$

which is asymptotically singular, due to cointegration in the component processes of $\bar{W}_t$. Note that $\int_0^1\tilde{B}_w\left(s\right)^2ds=\int_0^1 a'\tilde{B}_g\left(s\right)\tilde{B}_g\left(s\right)'ads>0$ $a.s.$ since $a\neq 0$ (see Phillips and Hansen, 1990). As usual with asymptotically singular moment matrices of cointegrated processes (Phillips, 1995) we proceed by rotation methods as follows.

In particular, to address the asymptotic degeneracy in (80) we rotate coordinates to isolate the direction of cointegration and the orthogonal direction. Define the orthogonal matrix $H=\left[\beta_w,\beta_\perp\right]$, where $\beta_\perp$ is an orthogonal complement matrix of β_w. let $D_n=\text{diag}\left(n,\sqrt{n}I_2\right)$. Then

$$\begin{aligned} D_nH'\left(\sum_{t=1}^{n}\widetilde{\bar{W}}_t\widetilde{\bar{W}}_t'\right)^{-1}HD_n &= \left(D_n^{-1}\left(\sum_{t=1}^{n}H'\widetilde{\bar{W}}_t\widetilde{\bar{W}}_t'H\right)D_n^{-1}\right)^{-1} \\ &= \begin{bmatrix} n^{-2}\sum_{t=1}^{n}\beta_w'\widetilde{\bar{W}}_t\widetilde{\bar{W}}_t'\beta_w & n^{-3/2}\sum_{t=1}^{n}\beta_w'\widetilde{\bar{W}}_t\widetilde{\bar{W}}_t'\beta_\perp \\ n^{-3/2}\sum_{t=1}^{n}\beta_\perp'\widetilde{\bar{W}}_t\widetilde{\bar{W}}_t'\beta_w & n^{-1}\sum_{t=1}^{n}\beta_\perp'\widetilde{\bar{W}}_t\widetilde{\bar{W}}_t'\beta_\perp \end{bmatrix}^{-1} \\ &= \begin{bmatrix} n^{-2}\sum_{t=1}^{n}\beta_w'\widetilde{\bar{W}}_t\widetilde{\bar{W}}_t'\beta_w & O_p\left(n^{-1/2}\right) \\ O_p\left(n^{-1/2}\right) & n^{-1}\sum_{t=1}^{n}\beta_\perp'\bar{W}_t\bar{W}_t'\beta_\perp \end{bmatrix}^{-1} \\ &= \begin{bmatrix} \left(n^{-2}\sum_{t=1}^{n}\beta_w'\widetilde{\bar{W}}_t\widetilde{\bar{W}}_t'\beta_w\right)^{-1} & o_p\left(1\right) \\ o_p\left(1\right) & \left(n^{-1}\sum_{t=1}^{n}\beta_\perp'\widetilde{\bar{W}}_t\widetilde{\bar{W}}_t'\beta_\perp\right)^{-1} \end{bmatrix} \\ &\Rightarrow \begin{bmatrix} \left(\int_0^1\tilde{B}_w\left(r\right)^2dr\right)^{-1} & o_p\left(1\right) \\ o_p\left(1\right) & \left(\beta_\perp'\left\{\mathbb{E}\left(u_{wt}u_{wt}'\right)\right\}\beta_\perp\right)^{-1} \end{bmatrix}, \end{aligned}$$

since $\beta_{\perp}'\widetilde{\overline{W}}_t = \beta_{\perp}'\tilde{u}_{wt}$ and $n^{-1}\sum_{t=1}^{n}\beta_{\perp}'\tilde{u}_{wt}\tilde{u}_{wt}'\beta_{\perp} \to_p \beta_{\perp}'\left\{\mathbb{E}\left(u_{wt}u_{wt}'\right)\right\}\beta_{\perp}$. Then, we have

$$
\begin{aligned}
n\left(\sum_{t=1}^{n}\widetilde{\overline{W}}_t\widetilde{\overline{W}}_t'\right)^{-1} &= \sqrt{n}HD_n^{-1}\left[D_nH'\left(\sum_{t=1}^{n}\widetilde{\overline{W}}_t\widetilde{\overline{W}}_t'\right)^{-1}HD_n\right]D_n^{-1}H'\sqrt{n} \\
&= \sqrt{n}HD_n^{-1}\left[\left(D_n^{-1}H'\sum_{t=1}^{n}\widetilde{\overline{W}}_t\widetilde{\overline{W}}_t'D_n^{-1}H'\right)^{-1}\right]D_n^{-1}H'\sqrt{n} \\
&= \sqrt{n}HD_n^{-1}\begin{bmatrix}\left(n^{-2}\sum_{t=1}^{n}\beta_w'\widetilde{\overline{W}}_t\widetilde{\overline{W}}_t'\beta_w\right)^{-1} & o_p(1) \\ o_p(1) & \left(n^{-1}\sum_{t=1}^{n}\beta_{\perp}'\widetilde{\overline{W}}_t\widetilde{\overline{W}}_t'\beta_{\perp}\right)^{-1}\end{bmatrix}D_n^{-1}H'\sqrt{n} \\
&= H\begin{bmatrix}\frac{1}{n}\left(n^{-2}\sum_{t=1}^{n}\beta_w'\widetilde{\overline{W}}_t\widetilde{\overline{W}}_t'\beta_w\right)^{-1} & o_p(1) \\ o_p(1) & \left(n^{-1}\sum_{t=1}^{n}\beta_{\perp}'\widetilde{\overline{W}}_t\widetilde{\overline{W}}_t'\beta_{\perp}\right)^{-1}\end{bmatrix}H' \\
&\to_p \beta_{\perp}\left(\beta_{\perp}'\left\{\mathbb{E}\left(u_{wt}u_{wt}'\right)\right\}\beta_{\perp}\right)^{-1}\beta_{\perp}'.
\end{aligned}
$$

Next consider the sample covariance $\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\beta_{\perp}'\widetilde{\overline{W}}_t\xi_{t+1} = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\beta_{\perp}'\tilde{u}_{wt}\xi_{t+1}$ where ξ_{t+1} is defined in (35) by $N^{-1/2}\sum_{i=1}^{N}u_{i,t+1} \to_{a.s.} \xi_{t+1}$ in a suitably expanded probability space. Observe that

$$
\begin{aligned}
\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\beta_{\perp}'\tilde{u}_{wt}\xi_{t+1} &= \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\beta_{\perp}'u_{wt}\xi_{t+1} - \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\beta_{\perp}'\left(n^{-1}\sum_{s=1}^{n}u_{ws}\right)\xi_{t+1} \\
&= \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\beta_{\perp}'u_{wt}\xi_{t+1} + o_p(1),
\end{aligned}
$$

and, as we now show, $\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\beta_{\perp}'u_{wt}\xi_{t+1}$ satisfies the martingale CLT.

First, ξ_t is *iid* over t since it inherits the time series properties of the panel $\{u_{it}\}_{i=1}^{\infty}$ and $u_{it} \sim_{iid} \left(0, \sigma_u^2\right)$ over all (i,t). It follows that the sequence $\{\beta_{\perp}'u_{wt}\xi_{t+1}\}$ is a martingale difference sequence (mds) with respect to the filtration $\mathcal{F}_t = \sigma\left\{u^t, u^{t-1}, ...; \delta^{\infty}\right\}$ with $u^t = \left(\{u_{it}\}_{i=1}^{\infty}, u_{gt}, u_{ct}, \left\{u_{it}^P\right\}_{i=1}^{\infty}\right)$ and $\delta^{\infty} = \{\delta_i\}_{i=1}^{\infty}$, because u_{wt} is $\mathcal{F}_t$ measurable, $\mathbb{E}\left(\beta_{\perp}'u_{wt}\xi_{t+1}|\mathcal{F}_t\right) = 0$, and $\mathbb{E}\left(\beta_{\perp}'u_{wt}u_{wt}'\beta_{\perp}\xi_{t+1}^2|\mathcal{F}_t\right) = \sigma_u^2\beta_{\perp}'u_{wt}u_{wt}'\beta_{\perp}$. Second, the martingale conditional variance is

$$
\left\langle\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\beta_{\perp}'u_{wt}\xi_{t+1}\right\rangle = \frac{1}{n}\sum_{t=1}^{n}\beta_{\perp}'u_{wt}u_{wt}'\beta_{\perp}\sigma_u^2 \to_p \beta_{\perp}'\left\{\mathbb{E}\left(u_{wt}u_{wt}'\right)\right\}\beta_{\perp} > 0,
$$

showing that the martingale stability condition is satisfied. The Lindeberg condition holds because

$$\begin{aligned}
& \frac{1}{n}\sum_{t=1}^{n}\mathbb{E}\left[\left\|\beta_{\perp}' u_{wt}\xi_{t+1}\right\|^2 \mathbf{1}\left\{\left\|\beta_{\perp}' u_{wt}\xi_{t+1}\right\| > \epsilon\sqrt{n}\right\}|\mathcal{F}_t\right] \\
\leq\ & \frac{1}{n}\sum_{t=1}^{n}\left\|\beta_{\perp}' u_{wt}\right\|^2 \mathbb{E}\left[\left\|\xi_{t+1}\right\|^2 \mathbf{1}\left\{\left\|\beta_{\perp}' u_{wt}\xi_{t+1}\right\| > \epsilon\sqrt{n}\right\}|\mathcal{F}_t\right] \\
\leq\ & \frac{1}{n}\sum_{t=1}^{n}\left\|\beta_{\perp}' u_{wt}\right\|^2 \mathbf{1}\left\{\left\|\beta_{\perp}' u_{wt}\right\| > \epsilon^{1/2}n^{1/4}\right\}\left\{\mathbb{E}\left[\left\|\xi_{t+1}\right\|^2 |\mathcal{F}_t\right]\right\} \\
& +\frac{1}{n}\sum_{t=1}^{n}\left\|\beta_{\perp}' u_{wt}\right\|^2 \mathbb{E}\left[\left\|\xi_{t+1}\right\|^2 \mathbf{1}\left\{\left\|\xi_{t+1}\right\| > \epsilon^{1/2}n^{1/4}\right\}|\mathcal{F}_t\right] \\
=\ & \left[\frac{1}{n}\sum_{t=1}^{n}\left\|\beta_{\perp}' u_{wt}\right\|^2 \mathbf{1}\left\{\left\|\beta_{\perp}' u_{wt}\right\| > \epsilon^{1/2}n^{1/4}\right\}\right]\left\{\mathbb{E}\left[\left\|\xi_{1}\right\|^2 |\mathcal{F}_0\right]\right\} \\
& +\left[\frac{1}{n}\sum_{t=1}^{n}\left\|\beta_{\perp}' u_{wt}\right\|^2\right]\mathbb{E}\left[\left\|\xi_{1}\right\|^2 \mathbf{1}\left\{\left\|\xi_{1}\right\| > \epsilon^{1/2}n^{1/4}\right\}|\mathcal{F}_0\right] \\
\to_{L_1}\ & 0.
\end{aligned}$$

We therefore have

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\beta_{\perp}' u_{wt}\xi_{t+1} \Rightarrow \mathcal{N}\left(0, \sigma_u^2\beta_{\perp}'\left\{\mathbb{E}\left(u_{wt}u_{wt}'\right)\right\}\beta_{\perp}\right),$$

and so, since $\beta_{\perp}'\widetilde{\bar{W}}_t = \beta_{\perp}'\tilde{u}_{wt} = \beta_{\perp}' u_{wt} - \beta_{\perp}'\left(n^{-1}\sum_{s=1}^{n} u_{ws}\right) = \beta_{\perp}' u_{wt} + o_p(1)$,

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\beta_{\perp}'\widetilde{\bar{W}}_t\xi_{t+1} = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left[\beta_{\perp}' u_{wt}\right]\xi_{t+1} + o_p(1) \Rightarrow \mathcal{N}\left(0, \sigma_u^2\beta_{\perp}'\left\{\mathbb{E}\left(u_{wt}u_{wt}'\right)\right\}\beta_{\perp}\right).$$

In a similar way, since $\beta_{\perp}'\widetilde{\bar{W}}_t = \beta_{\perp}'\bar{W}_t - n^{-1}\sum_{t=1}^{n}\beta_{\perp}'\bar{W}_t = \beta_{\perp}' u_{wt} - n^{-1}\sum_{t=1}^{n}\beta_{\perp}' u_{wt} = \beta_{\perp}' u_{wt} + o_{a.s.}(1)$, we have

$$\frac{1}{n}\sum_{t=1}^{n}\beta_{\perp}'\widetilde{\bar{W}}_t\widetilde{\bar{W}}_t'\beta_{\perp} = \frac{1}{n}\sum_{t=1}^{n}\beta_{\perp}' u_{wt}u_{wt}'\beta_{\perp}' + o_p(1) \to_p \beta_{\perp}'\left\{\mathbb{E}\left(u_{wt}u_{wt}'\right)\right\}\beta_{\perp}.$$

It follows that

$$\sqrt{nN}\left(\hat{\gamma}-\gamma\right) = H\left(\frac{1}{n}\sum_{t=1}^{n}H'\widetilde{\bar{W}}_t\widetilde{\bar{W}}_t'H\right)^{-1}\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}H'\widetilde{\bar{W}}_t\widetilde{\xi}_{N,t+1}+o_p\left(1\right)\right)$$

$$= H\left(\frac{1}{n}\sum_{t=1}^{n}H'\widetilde{\bar{W}}_t\widetilde{\bar{W}}_t'H\right)^{-1}\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}H'\widetilde{\bar{W}}_t\xi_{t+1}+o_p\left(1\right)\right)$$

$$= H\left\{\begin{bmatrix}\left(\beta_\perp'\left\{\mathbb{E}\left(u_{wt}u_{wt}'\right)\right\}\beta_\perp\right)^{-1}+o_p\left(1\right) & o_p\left(1\right)\\ o_p\left(1\right) & o_p\left(1\right)\end{bmatrix}\right\}\left\{\begin{bmatrix}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left[\beta_\perp'u_{wt}\right]\xi_{t+1}+o_p\left(1\right)\\ o_p\left(1\right)\end{bmatrix}\right\}$$

$$\Rightarrow \mathcal{N}\left(0,\sigma_u^2\beta_\perp\left(\beta_\perp'\left\{\mathbb{E}\left(u_{wt}u_{wt}'\right)\right\}\beta_\perp\right)^{-1}\beta_\perp'\right). \tag{81}$$

which proves part (a).

Part (b): In the alternate direction of β_w we have

$$\sqrt{n^2N}\beta_w'\left(\hat{\gamma}-\gamma\right) = \beta_w'H\left(\frac{1}{n^2}\sum_{t=1}^{n}H'\widetilde{\bar{W}}_t\widetilde{\bar{W}}_t'H\right)^{-1}\left(\frac{1}{n}\sum_{t=1}^{n}H'\widetilde{\bar{W}}_t\widetilde{\xi}_{N,t+1}+o_p\left(1\right)\right)$$

$$= \beta_w'H\left(\frac{1}{n^2}\sum_{t=1}^{n}H'\widetilde{\bar{W}}_t\widetilde{\bar{W}}_t'H\right)^{-1}\left(\frac{1}{n}\sum_{t=1}^{n}H'\widetilde{\bar{W}}_t\xi_{t+1}+o_p\left(1\right)\right).$$

As in part (a), we partition the scaled inverse of the signal matrix as follows

$$n^2\beta_w'\left(\sum_{t=1}^{n}\widetilde{\bar{W}}_t\widetilde{\bar{W}}_t'\right)^{-1}\beta_w = n^2\beta_w'HD_n^{-1}\left(D_n^{-1}H'\sum_{t=1}^{n}\widetilde{\bar{W}}_t\widetilde{\bar{W}}_t'HD_n^{-1}\right)^{-1}D_n^{-1}H'\beta_w$$

$$= n^2\left[1,O_{1\times2}\right]D_n^{-1}\left[D_nH'\left(\sum_{t=1}^{n}\widetilde{\bar{W}}_t\widetilde{\bar{W}}_t'\right)^{-1}HD_n\right]D_n^{-1}\begin{bmatrix}1\\ O_{2\times1}\end{bmatrix}$$

$$= n^2\left[1,O_{1\times2}\right]D_n^{-1}\left(D_n^{-1}H'\sum_{t=1}^{n}\widetilde{\bar{W}}_t\widetilde{\bar{W}}_t'HD_n^{-1}\right)^{-1}D_n^{-1}\begin{bmatrix}1\\ O_{2\times1}\end{bmatrix}$$

$$= n^2\left[1,O_{1\times2}\right]D_n^{-1}\begin{bmatrix}\left(n^{-2}\sum_{t=1}^{n}\beta_w'\widetilde{\bar{W}}_t\widetilde{\bar{W}}_t'\beta_w\right)^{-1} & o_p\left(1\right)\\ o_p\left(1\right) & \left(n^{-1}\sum_{t=1}^{n}\beta_\perp'\widetilde{\bar{W}}_t\widetilde{\bar{W}}_t'\beta_\perp\right)^{-1}\end{bmatrix}D_n^{-1}\begin{bmatrix}1\\ O_{2\times1}\end{bmatrix}$$

$$= \left(n^{-2}\sum_{t=1}^{n}\beta_w'\widetilde{\bar{W}}_t\widetilde{\bar{W}}_t'\beta_w\right)^{-1}$$

$$\Rightarrow \left(\int_0^1 a'B_g\left(s\right)B_g\left(s\right)'a\beta_w'\bar{\beta}_w\right)^{-1} = \left(\int_0^1 a'B_g\left(s\right)B_g\left(s\right)'a\right)^{-1} \tag{82}$$

The sample covariance term is

$$\frac{1}{n}\sum_{t=1}^{n} H'\widetilde{\overline{W}}_t\xi_{t+1} = \left[\begin{array}{c} \frac{1}{n}\sum_{t=1}^{n}\left[a'\tilde{U}_{gt}+\beta_w'\tilde{u}_{wt}\right]\xi_{t+1}+o_p(1) \\ \frac{1}{n}\sum_{t=1}^{n}\left[\beta_\perp'\tilde{u}_{wt}\right]\xi_{t+1}+o_p(1)\end{array}\right] = \left[\begin{array}{c} \frac{1}{n}\sum_{t=1}^{n}\left[a'\tilde{U}_{gt}\right]\xi_{t+1}+o_p(1) \\ o_p(1)\end{array}\right]$$

$$\Rightarrow \left[\begin{array}{c} \int_0^1 a'B_g(s)\,dB_\xi(s) \\ O_{2\times 1}\end{array}\right] \equiv \left[\begin{array}{c} \mathcal{MN}\left(0,\sigma_u^2\mathrm{tr}\left\{\Sigma_r\int_0^1 a'\tilde{B}_g(s)\tilde{B}_g(s)'\,ads\right\}\right) \\ O_{2\times 1}\end{array}\right] \quad (83)$$

The weak convergence $\frac{1}{n}\sum_{t=1}^{n}\tilde{U}_{gt}\xi_{t+1}\Rightarrow\int_0^1\tilde{B}_g(s)\,dB_\xi(s)$ holds just as in the argument leading to (67) above because ξ_{t+1} is independent of U_{gs} for all (t,s); and $\frac{1}{n}\sum_{t=1}^{n}\tilde{u}_{wt}\xi_{t+1} = \frac{1}{n}\sum_{t=1}^{n}u_{wt}\xi_{t+1}-\left(\frac{1}{n}\sum_{t=1}^{n}u_{wt}\right)\left(\frac{1}{n}\sum_{t=1}^{n}\xi_{t+1}\right)=o_p(1)$ because ξ_{t+1} is independent of the non-negligible stationary components $\{u_{cs},u_{gs}\}_{-\infty}^{t}$ of u_{wt}, so that $\frac{1}{n}\sum_{t=1}^{n}u_{wt}\xi_{t+1}\to_p 0$. Recall that $\xi_{t+1}\sim_{iid}\left(0,\sigma_u^2\right)$

Let

$$\begin{aligned} & A_{ww.\perp} \\ = {} & n^{-2}\sum_{t=1}^{n}\beta_w'\widetilde{\overline{W}}_t\widetilde{\overline{W}}_t'\beta_w-\frac{1}{n}\left(n^{-1}\sum_{t=1}^{n}\beta_w'\widetilde{\overline{W}}_t\widetilde{\overline{W}}_t'\beta_\perp\right)\left(n^{-1}\sum_{t=1}^{n}\beta_\perp'\widetilde{\overline{W}}_t\widetilde{\overline{W}}_t'\beta_\perp\right)^{-1}\left(n^{-1}\sum_{t=1}^{n}\beta_\perp'\widetilde{\overline{W}}_t\widetilde{\overline{W}}_t'\beta_w\right) \\ = {} & n^{-2}\sum_{t=1}^{n}\beta_w'\widetilde{\overline{W}}_t\widetilde{\overline{W}}_t'\beta_w+O_p\left(\frac{1}{n}\right), \end{aligned}$$

we have

$$\begin{aligned} & \sqrt{n^2N}\beta_w'(\hat{\gamma}-\gamma) = \beta_w'H\left(\frac{1}{n^2}\sum_{t=1}^{n}H'\widetilde{\overline{W}}_t\widetilde{\overline{W}}_t'H\right)^{-1}\left(\frac{1}{n}\sum_{t=1}^{n}H'\widetilde{\overline{W}}_t\widetilde{\xi}_{N,t+1}+o_p(1)\right) \\ = {} & [1,O_{1\times 2}]\left(\frac{1}{n^2}\sum_{t=1}^{n}H'\widetilde{\overline{W}}_t\widetilde{\overline{W}}_t'H\right)^{-1}\left(\frac{1}{n}\sum_{t=1}^{n}H'\widetilde{\overline{W}}_t\xi_{t+1}+o_p(1)\right) \\ = {} & [1,O_{1\times 2}]\left[\begin{array}{cc} n^{-2}\sum_{t=1}^{n}\beta_w'\widetilde{\overline{W}}_t\widetilde{\overline{W}}_t'\beta_w & n^{-2}\sum_{t=1}^{n}\beta_w'\widetilde{\overline{W}}_t\widetilde{\overline{W}}_t'\beta_\perp \\ n^{-2}\sum_{t=1}^{n}\beta_\perp'\widetilde{\overline{W}}_t\widetilde{\overline{W}}_t'\beta_w & n^{-2}\sum_{t=1}^{n}\beta_\perp'\widetilde{\overline{W}}_t\widetilde{\overline{W}}_t'\beta_\perp \end{array}\right]^{-1} \\ & \times\left[\begin{array}{c} \frac{1}{n}\sum_{t=1}^{n}\left[a'\tilde{U}_{gt}+\beta_w'\tilde{u}_{wt}\right]\xi_{t+1}+o_p(1) \\ \frac{1}{n}\sum_{t=1}^{n}\left[\beta_\perp'\tilde{u}_{wt}\right]\xi_{t+1}+o_p(1)\end{array}\right] \\ = {} & [1,O_{1\times 2}]\left[\begin{array}{cc} A_{ww.\perp}^{-1} & o_p(1) \\ o_p(1) & o_p(1)\end{array}\right]\times\left[\begin{array}{c} \frac{1}{n}\sum_{t=1}^{n}\left[a'\tilde{U}_{gt}\right]\xi_{t+1}+o_p(1) \\ o_p(1)\end{array}\right] \end{aligned}$$

$$
\begin{aligned}
&= \left(\frac{1}{n^2}\sum_{t=1}^{n}\beta_w'\widetilde{\overline{W}}_t\widetilde{\overline{W}}_t'\beta_w\right)^{-1}\left(\frac{1}{n}\sum_{t=1}^{n}\left[a'\tilde{U}_{gt}\right]\xi_{t+1}+o_p(1)\right)\\
&\Rightarrow \left(\int_0^1 a'\tilde{B}_g(s)\,\tilde{B}_g(s)'\,ads\right)^{-1}\int_0^1 a'\tilde{B}_g(s)\,dB_\xi(s)\\
&\equiv \mathcal{MN}\left(0,\frac{\sigma_u^2}{a'\int_0^1\tilde{B}_g(s)\,\tilde{B}_g(s)'\,dsa}\right),
\end{aligned}
$$

giving the stated result.

In addition, suppose we take the linear combination $b_w'\gamma$ where $b_w = \mu\beta_w$ for some scalar $\mu \neq 0$. Then, using the above result, we have

$$
\sqrt{n^2N}b_w'(\hat{\gamma}-\gamma)\Rightarrow\mathcal{MN}\left(0,\frac{\mu^2\sigma_u^2}{\int_0^1 a'\tilde{B}_g(s)\,\tilde{B}_g(s)'\,ads}\right)
$$

Just as in the derivation of (82), we find that

$$
\begin{aligned}
& n^2b_w'\left(\sum_{t=1}^{n}\widetilde{\overline{W}}_t\widetilde{\overline{W}}_t'\right)^{-1}b_w' = \mu^2\beta_w'HD_n^{-1}\left[D_nH'\left(\sum_{t=1}^{n}\widetilde{\overline{W}}_t\widetilde{\overline{W}}_t'\right)^{-1}HD_n\right]D_n^{-1}H'\beta_w\\
&= \mu^2n^2\left[1,O_{1\times2}\right]D_n^{-1}\left(D_n^{-1}H'\sum_{t=1}^{n}\widetilde{\overline{W}}_t\widetilde{\overline{W}}_t'HD_n^{-1}\right)^{-1}D_n^{-1}\begin{bmatrix}1\\O_{2\times1}\end{bmatrix}\\
&= \mu^2n^2\left[1,O_{1\times2}\right]D_n^{-1}\begin{bmatrix}\left(n^{-2}\sum_{t=1}^{n}\beta_w'\widetilde{\overline{W}}_t\widetilde{\overline{W}}_t'\beta_w\right)^{-1} & o_p(1)\\ o_p(1) & \left(n^{-1}\sum_{t=1}^{n}\beta_\perp'\widetilde{\overline{W}}_t\widetilde{\overline{W}}_t'\beta_\perp\right)^{-1}\end{bmatrix}D_n^{-1}\begin{bmatrix}1\\O_{2\times1}\end{bmatrix}\\
&= \mu^2\left(n^{-2}\sum_{t=1}^{n}\beta_w'\widetilde{\overline{W}}_t\widetilde{\overline{W}}_t'\beta_w\right)^{-1}\\
&\Rightarrow \mu^2\left(\int_0^1 a'B_g(s)\,B_g(s)'\,a\right)^{-1}, \qquad (84)
\end{aligned}
$$

which enables inference and the construction of confidence intervals in the usual manner.
■

9.5 Proof of Theorem 4

We derive the limit distribution of $TCS = f(\theta) = \frac{\gamma_3}{1-\beta_1-\gamma_1}\ln(2)$ using the delta method. Complications arise because of differing rates of convergence in the elements of $\theta =$

$(\beta_1, \gamma_1, \gamma_3)'$ and degeneracies in the limit theory. Since $\hat{\theta}$ is consistent for θ and f is continuously differentiable when $|\beta_1 + \gamma_1| < 1$, we use the following Taylor expansion, denoting first derivatives evaluated at the true value of θ by subscripts,

$$\begin{aligned}
&\widehat{TCS} - TCS = f\left(\hat{\beta}_1, \hat{\gamma}_1, \hat{\gamma}_3\right) - f\left(\beta_1, \gamma_1, \gamma_3\right) \\
&= f_{\beta_1}\left(\hat{\beta}_1 - \beta_1\right) + f_{\gamma_1}\left(\hat{\gamma}_1 - \gamma_1\right) + f_{\gamma_3}\left(\hat{\gamma}_3 - \gamma_3\right) + o_p\left(|\hat{\gamma}_1 - \gamma_1| + |\hat{\gamma}_3 - \gamma_3| + \left|\hat{\beta}_1 - \beta_1\right|\right) \\
&= f_{\beta_1}\left(\hat{\beta}_1 - \beta_1\right) + f_{\gamma_1}\left(\hat{\gamma}_1 - \gamma_1\right) + f_{\gamma_3}\left(\hat{\gamma}_3 - \gamma_3\right) + o_p\left(\frac{1}{\sqrt{nN}}\right) \qquad (85) \\
&= f_{\gamma_1}\left(\hat{\gamma}_1 - \gamma_1\right) + f_{\gamma_3}\left(\hat{\gamma}_3 - \gamma_3\right) + o_p\left(\frac{1}{\sqrt{nN}}\right), \qquad (86)
\end{aligned}$$

which accounts for the differential convergence rates of $\left|\hat{\beta}_1 - \beta_1\right| = O_p\left(1/\sqrt{n^2N}\right) = o_p\left(1/\sqrt{nN}\right)$ and $|\hat{\gamma} - \gamma| = O_p\left(1/\sqrt{nN}\right)$. Importantly, the rate $|\hat{\gamma} - \gamma| = O_p\left(1/\sqrt{nN}\right)$ applies for all linear combinations $b'\gamma$for which $b'\beta_\perp \neq 0$. In such cases, using the limit normal distribution of $b'\left(\hat{\gamma} - \gamma\right)$ given in Theorem 3(a), we have

$$\sqrt{nN}\left(\widehat{TCS} - TCS\right) \Rightarrow \mathcal{N}\left(0, \sigma_u^2 b'\beta_\perp\left[\beta_\perp'\mathbb{E}\left(u_{wt}u_{wt}'\right)\beta_\perp\right]^{-1}\beta_\perp' b\right) \qquad (87)$$

with derivative vector

$$\begin{aligned}
b' &= \left(f_{\gamma_1}, 0, f_{\gamma_3}\right) = \ln(2) \times \left(\frac{\gamma_3}{\left(1 - \beta_1 - \gamma_1\right)^2}, 0, \frac{1}{1 - \beta_1 - \gamma_1}\right) \\
&= \frac{\ln(2)}{1 - \beta_1 - \gamma_1} \times \left(\frac{\gamma_3}{1 - \beta_1 - \gamma_1}, 0, 1\right). \qquad (88)
\end{aligned}$$

The limit theory (87) then holds whenever $b'\beta_\perp \neq 0$.

By direct calculation using (88) and the matrix β_γ in (22) for the cointegration space spanned by $\beta_\perp$, we obtain

$$\begin{aligned}
b'\beta_\gamma &= \frac{\ln(2)}{1 - \beta_1 - \gamma_1}\left(\frac{\gamma_3}{1 - \beta_1 - \gamma_1}, 0, 1\right)\begin{bmatrix} \gamma_1 & \beta_1 - 1 \\ \gamma_2 & \beta_2 \\ \gamma_3 & 0 \end{bmatrix} = \frac{\ln(2)}{1 - \beta_1 - \gamma_1}\left[\gamma_3 + \frac{\gamma_3\gamma_1}{1 - \beta_1 - \gamma_1}, \frac{\gamma_3\left(\beta_1 - 1\right)}{1 - \beta_1 - \gamma_1}\right] \\
&= \frac{\ln(2)}{1 - \beta_1 - \gamma_1}\left[\frac{\gamma_3\left(1 - \beta_1\right)}{1 - \beta_1 - \gamma_1}, \frac{\gamma_3\left(\beta_1 - 1\right)}{1 - \beta_1 - \gamma_1}\right] = \frac{\ln(2)\gamma_3\left(1 - \beta_1\right)}{\left(1 - \beta_1 - \gamma_1\right)^2}\left[1, -1\right]. \qquad (89)
\end{aligned}$$

Hence $b'\beta_\gamma = b'\beta_\perp = 0$ requires $\gamma_3 = 0$ when $|\beta_1| < 1$. Thus, the limit theory (87) applies except when $\gamma_3 = 0$, in which case there is no global GHG impact in the energy balance equation (2). In that event, $TCS = 0$ and there is no climate sensitivity to CO_2.

However, in this special case where $\gamma_3 = 0$ and $b' = \frac{\ln(2)}{1-\beta_1-\gamma_1} \times (0,0,1) = \frac{\ln(2)}{1-\beta_1-\gamma_1}\beta_w$ we have from Theorem 3(b)

$$\begin{aligned}\sqrt{n^2N}b'(\hat{\gamma}-\gamma) &\Rightarrow \mu\left(\int_0^1 a'\tilde{B}_g(s)\tilde{B}_g(s)'a\right)^{-1}\int_0^1 a'\tilde{B}_g(s)\,dB_\xi(s)\\ &\equiv \mathcal{MN}\left(0,\sigma_u^2\mu^2\left(\int_0^1 a'\tilde{B}_g(s)\tilde{B}_g(s)'ads\right)^{-1}\right). \qquad (90)\end{aligned}$$

Thus, in this case, $\hat{\gamma}$ and $\hat{\beta}_1$ have the same convergence rate and in place of (86) we now have the Taylor expansion

$$\begin{aligned}\widehat{TCS}-TCS &= f\left(\hat{\beta}_1,\hat{\gamma}_1,\hat{\gamma}_3\right)-f(\beta_1,\gamma_1,\gamma_3)\\ &= f_{\beta_1}\left(\hat{\beta}_1-\beta_1\right)+f_{\gamma_1}(\hat{\gamma}_1-\gamma_1)+f_{\gamma_3}(\hat{\gamma}_3-\gamma_3)+o_p\left(\frac{1}{\sqrt{n^2N}}\right)\\ &= f_{\gamma_3}(\hat{\gamma}_3-\gamma_3)+o_p\left(\frac{1}{\sqrt{n^2N}}\right), \qquad (91)\end{aligned}$$

since $f_{\gamma_1} = \frac{\gamma_3}{(1-\beta_1-\gamma_1)^2} = 0$, $f_{\beta_1} = \frac{\gamma_3}{(1-\beta_1-\gamma_1)^2} = 0$, and $f_{\gamma_3} = \frac{\ln(2)}{1-\beta_1-\gamma_1} \neq 0$ when $\gamma_3 = 0$. Then, with $b' = \frac{\ln(2)}{1-\beta_1-\gamma_1} \times (0,0,1)$, we use the limit theory representation of (90), viz.,

$$\begin{aligned}\sqrt{n^2N}b'(\hat{\gamma}-\gamma) &= \frac{\ln(2)}{1-\beta_1-\gamma_1}\sqrt{n^2N}(\hat{\gamma}_3-\gamma_3)\\ &\Rightarrow \frac{\ln(2)}{1-\beta_1-\gamma_1}\left(\int_0^1 a'\tilde{B}_g(s)\tilde{B}_g(s)'ads\right)^{-1}\int_0^1 a'\tilde{B}_g(s)\,dB_\xi(s). \qquad (92)\end{aligned}$$

It follows from (91) and (92) that

$$\begin{aligned}\sqrt{n^2N}\left(\widehat{TCS}-TCS\right) &\Rightarrow \frac{\ln(2)}{1-\beta_1-\gamma_1}\left(\int_0^1 a'\tilde{B}_g(s)\tilde{B}_g(s)'ads\right)^{-1}\int_0^1 a'\tilde{B}_g(s)\,dB_\xi(s)\\ &\equiv \mathcal{MN}\left(0,\sigma_u^2\left(\frac{\ln(2)}{1-\beta_1-\gamma_1}\right)^2\left(\int_0^1 a'\tilde{B}_g(s)\tilde{B}_g(s)'ads\right)^{-1}\right),\end{aligned}$$

giving the stated result in part (b). ■

10 References

Andreae M. O., C. D. Jones & P. M. Cox (2005). "Strong present-day aerosol cooling implies a hot future," *Nature,* 435, 1187–1190.

Flato, G., J. Marotzke, B. Abiodun, P. Braconnot, S. C. Chou, W. Collins, P. Cox, F. Driouech, S. Emori, S., V. Eyring, C. Forest, P. Gleckler, E. Guilyardi, C. Jakob, V. Kattsov, C. Reason & M. Rummukainen (2013). "Evaluation of Climate Models, In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change," *Cambridge University Press* 9, 741–866.

Gilgen, H. & A. Ohmura (1999). "The Global Energy Balance Archive," *B Am Meteorol Soc,* 80, 831–850.

Hansen, J., M. Sato, P. Kharecha, Pushker % K. von Schuckmann (2011). "Earth's energy imbalance and implications," *Atmospheric Chemistry and Physics,* 11, 13421–13449.

Harris, I. P. D. J., P. D. Jones, T. J. Osborn & D. H. Lister (2014) "Updated high-resolution grids of monthly climatic observations–the CRU TS3. 10 Dataset, " *International Journal of Climatology*, Vol. 34, No. 3, 2016, pp. 632-642.

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.

Ibragimov, R. and P. C. B. Phillips (2008). "Regression Asymptotics Using Martingale Convergence Methods," *Econometric Theory*, 24, pp. 888-947.

Lamarque, J-F, T. C. Bond, V. Eyring, C. Granier, A. Heil, Z. Klimont, D. Lee, C. Liousse, A. Mieville, B. Owen, M. G. Schultz, D. Shindell, S. J. Smith, E. Stehfest, J. Van Ardenne, O. R. Cooper, M. Kainuma, N. Mahowald, J. R. McConnell, V. Naik, K. Riahi, & D. P. Van Vuuren (2010) "Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application, " *Atmospheric Chemistry and Physics,* 10, 7017–7039.

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. (1988). "Multiple regression with integrated processes." In N. U. Prabhu, (ed.), *Statistical Inference from Stochastic Processes, Contemporary Mathematics* 80, 79–106.

Phillips, P. C. B. (1995). “Fully modified least squares and vector autoregression,” *Econometrica,* 63, 1023-1078.

Phillips, P. C. B. and B. E. Hansen (1990). “Statistical inference in instrumental variables regression with I(1) processes,” *Review of Economic Studies* 57, 99–125.

Phillips, P. C. B. and M. Loretan (1991). “Estimating Long-Run Economic Equilibria,” *Review of Economic Studies* 59, 407–436.

Phillips, P. C. B. and H. R. Moon (1999) : Linear Regression Limit Theory for Nonstationary Panel Data,” *Econometrica,* 67, 1057-1111.

Saikkonen, P. (1991). “Asymptotically Efficient Estimation of Cointegration Regressions,” *Econometric Theory* 7, 1–21.

Stock, J. H. and M. W. Watson (1993). “A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems,” *Econometrica* 61, 783–821.

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.

Taylor, K. E., R. J. Stouffer & G. A. Meehl (2012). “An overview of CMIP5 and the experiment design, ” *Bulletin of the American Meteorological Society,*, 93, 485–498.