Econometric Estimates of Earth's Transient Climate Sensitivity*

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

March 5, 2019

Abstract

How sensitive is Earth's climate to a given increase in atmospheric greenhouse gas (GHG) concentrations? This long-standing 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). Those 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 regularity conditions and asymptotic theory justifying the use of time series cointegration-based methods of estimation when there are both stochastic process and deterministic trends in the global forcing variables, such as GHGs, and station-level trend effects from such sources as local aerosol pollutants. The asymptotics validate estimation and confidence interval construction for econometric measures of Earth's transient climate sensitivity (TCS). The methods are applied to observational data and to data generated from several groups of global climate models (GCMs) that are sampled spatio-temporally and aggregated in the same way as the empirical observations for the time period 1964 - 2005. The findings indicate that 7 out of 9 of the GCM reported TCS values lie within the 95% empirical confidence interval computed econometrically from the GCM output. The analysis shows the potential of econometric methods to provide empirical estimates and confidence limits for TCS, to calibrate GCM simulation output against observational data in terms of the implied TCS estimates obtained via the econometric model, 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

*We thank three referees and the guest editors for extremely helpful comments on earlier versions of the paper. The work was presented at ANZESG in February 2018, NESG in May 2018 and ESAM in July 2018. Phillips acknowledges support from the Kelly Fund at the University of Auckland.

†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 challenging. 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. If aerosol cooling is underestimated, climate might appear less sensitive to CO_2 than it really is (Kiehl, 2007). 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 promise for climate studies and is worth pursuing, we acknowledge that in order to arrive at inferences concerning climate sensitivity using these 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 together with spatial and temporal dependence, econometric analysis requires a full development of asymptotic theory of estimation and inference in the presence of such trends and dependence whilst allowing for variable co-movement that may be governed by energy balance considerations.

The present paper contributes by addressing some of 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 and

deterministic nonstationarity in the data; (ii) asymptotic theory is developed for estimation and inference in the context of this expanded model using time series cointegration-based methods with an explicit focus on 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 groups of leading GCMs for the time period 1964 - 2005. The developments in (ii) are novel in econometrics because they allow treatment of nonstationarity with cointegrated regressors at both the individual station level data and the global aggregate level. The application in (iii) innovates by analyzing GCM simulated data by econometric methods and by carefully matching GCM simulated data at times and spatial locations for which observational data are available. This matching helps to assess 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 against observed data enables us to identify potential GCM model shortcomings which may not be so apparent in standard GCM validation exercises.

Section 2 provides an introduction to the spatial econometric framework for modeling key climate variables observable over time at specific station locations. Extensions to the model are made to accommodate stochastic driver variables that include both global forcing variables and station-specific aerosol pollution trends. Regularity conditions are provided and cross section aggregation effects are discussed. Some econometric implications of the expanded model are explored in Section 3, including the cointegrating structures that arise from energy balance/imbalance considerations at the individual station and global levels. Asymptotic theory for the time series estimation of the global relationship is developed which leads to global climate sensitivity estimation and an asymptotically valid confidence interval for this composite parameter. Section 4 reports an empirical application of this methodology to observational data related to that employed previously in Storelvmo et al. (2016) and to matching simulation output data from three groups of GCM models. Section 5 provides an overview of the model set up, the aggregation process leading to the fitted time series model, and computations involved in estimating the transient climate sensitivity parameter and a confidence interval for this parameter. This section is intended to provide a roadmap of the implementation for readers with less interest in the econometric methodology and asymptotic theory. These readers may find it easier to read the paper by omitting Section 3. Summary conclusions are given in Section 6. Relevant technical material, proofs of results, additional tables, and some further discussion are provided in Appendices A and B.

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

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

where the α_i are station-level effects, β_1 and β_2 are parameters, and the time specific effect

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

depends on the spatial aggregate variables[1] $\left(\bar{T}_t, \bar{R}_t\right) = \left(N^{-1}\sum_{i=1}^{N} T_{it}, N^{-1}\sum_{i=1}^{N} R_{it}\right)$ and the logarithm[2] of the CO_2 equivalent series, $\ln(CO_{2,t})$.

Equation (2) may be interpreted as an energy balancing relationship that captures the global linkage between temperature, radiation and greenhouse gas atmospheric influences. The balance or imbalance in these global elements is assumed to be one of the drivers impacting local temperature (T_i) in the subsequent time period via the time specific effect λ_t in equation (1). The dynamic panel regression equation (1) therefore characterizes the dynamic adjustment mechanism of station level temperature T_{it+1} as an autoregression on past temperature T_{it}, radiation R_{it}, and the time specific global energy effects embodied in λ_t. When λ_t is nonstationary, these effects are imported as time specific trends to the local temperature series $T_{i,t+1}$, a mechanism that carries the effects of untempered rising aggregate CO_2 levels through to local temperature. This formulation accords with empirical climate science research by Hansen et al. (2005; 2011) which discusses how the climate system is increasingly out of equilibrium. In effect, energy entering and leaving the climate system is unequal at present and the differential has been growing over time. In equation (1) the impact of this energy differential or imbalance is transported to station level temperature via the time specific effect λ_t.

2.1 Model Extensions and Assumptions

In what follows we extend this model to accommodate stochastic forcing variables at both the station-specific and global levels. Neither Magnus et al (2011) nor Storelvmo et al.

[1] The aggregate variables $\left(\bar{T}_t, \bar{R}_t\right)$ are strictly speaking arrays that also depend on N. However, under ergodicity (or weaker conditions involving uniform integrability) over the spatial index i these spatial aggregates will converge to population aggregate variables that are measurable with respect to the invariant event algebra at time t, so that $\left(\bar{T}_t, \bar{R}_t\right) \to_{a.s.} (T_t, R_t) = \lim_{N\to\infty}\left(N^{-1}\sum_{i=1}^{N} E_t T_{it}, E_t R_{it}\right)$ where E_t denotes expectation conditional on the invariant sigma algebra at time t.

[2] Use of a logarithmic transformation of CO_2 in the energy balance relation (2) is based on technical considerations associated with the way in which atmospheric concentrations of CO_2 impart climate forcing – see Huang and Bani Shahabadi (2014) for a recent analysis and justification.

(2016) provided a complete model capturing the linkages of the station level data to the energy balance relationship in a way that accommodates potential stochastic nonstationarity in the variables and additional forcing variables at both station and global levels. In what follows, 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 to broadly match the characteristics of the observed data. The model may be developed further, for instance to allow for time varying individual effects or idiosyncratic loading effects of the global factors embodied in λ_t. For instance, in place of (1) we may have

$$T_{i,t+1} = \alpha_i + \beta_1 T_{i,t} + \beta_2 R_{i,t} + \phi_i \lambda_t + u_{it+1}, \quad i = 1, ..., N \text{ and } t = 1, ..., n, \tag{3}$$

where the ϕ_i are factor loading coefficients. This extension allows for the impact of the global drivers in λ_t to vary spatially over i with weights ϕ_i. Such spatial dependence may be important empirically. For instance, the global forcing variables may influence local temperature dynamics in regions of different latitudes in different ways. The time series methods that we use in the present paper are unaffected by this generalization of the model because the climate sensitivity parameter of interest is a global, not a local, parameter and the methods are applied to spatially aggregated data, as will become apparent later. Global averages of the ϕ_i then have only scale effects on the parameters $\{\gamma_i : i = 0, 1, 2, 3\}$ in (2) and these parameters are then interpreted as embodying these scale effects.

The model (1) and (2) can be used to measure transient climate sensitivity for land (TCS_L) to CO_2 emissions, which is the main focus of the present work. TCS_L is defined as the change in global mean land surface air temperature after a doubling of CO_2 (at the time of doubling) and in the context of the above model has the following analytic form (Magnus et al., 2011; and Storelvmo et al., 2016) which is derived in Appendix A

$$TCS_L = \frac{\gamma_3}{1 - \beta_1 - \gamma_1} \times \ln(2). \tag{4}$$

One way to calculate TCS_L is for the full model (1) and (2) to be estimated with observed data using dynamic panel regression methods. Under certain conditions a valid confidence interval for the resulting estimate of TCS_L can then be obtained using the asymptotic theory for the parameter estimate of TCS_L that is implied by the joint limit theory of the estimates of the components $(\beta_1, \gamma_1, \gamma_3)$ of the model as $(N, n) \to \infty$. That approach was used in Storelvmo et al (2016) without any formal development of the limit theory and has the (apparent) advantage that it reveals the explicit dependence on the parameters of the full panel regression model (1) and (2). However, standard conditions for the validity of dynamic panel regression methods are stringent when applied in this context as they require temporal and spatial independence of the innovations $u_{i,t+1}$ to avoid bias and inconsistency and to ensure correct standard error and confidence interval calculations. In addition to these limitations, standard methods of panel inference do not presently allow for the complex mix of stochastically and deterministically trending variables and co-moving time series that are manifest in the data.

The present paper does not address these complications that arise in the application of dynamic panel methods. Instead, in what follows we introduce a direct time series approach to the estimation of TCS_L using data on the global variables $\left(\bar{T}_t, \bar{R}_t, \ln(CO_{2,t})\right)$. This approach has the advantage of accommodating the temporal and spatial dependence issues that affect dynamic panel estimation by utilizing cointegration-based methods that can more easily than disaggregated analysis deal with these complications as well as more directly address potential co-moving trending behavior in the data. The approach also allows for factor model extensions such as (3) of (1).

First, we complete the specification of the system and impose conditions on the parameters and stochastic elements that match the characteristics of the available data on the variables $(T_{it}, R_{it}, CO_{2t})$. The framework is expanded by adding the following mechanisms for the generation of local radiation effects and global CO_2

$$R_{it} = R_{it}^0 + \delta_{ri}' U_{gt} + P_{it}, \tag{5}$$

$$\ln(CO_{2,t}) = \delta_{c0} + \delta_{c1} t + \delta_c' U_{gt} + u_{ct}, \tag{6}$$

$$U_{gt} = \sum_{s=1}^{t} u_{gs}. \tag{7}$$

In (5) local radiation R_{it} is assumed to have a component R_{it}^0 that embodies stationary fluctuations about some mean level $\mathbb{E}\left(R_{it}^0\right)$, U_{gt} is an m_g- vector of latent global forcing variables that are stochastically nonstationary, driven by global shocks u_{gt}, and δ_{ri} is an idioscyncratic vector factor loading parameter vector that captures any station level effects arising from the common global shocks embodied in U_{gt}. Not all the components of U_{gt} can be expected to influence R_{it}. But just as CO_2 levels in the atmosphere have been rising, so have other gases that are precursors to atmospheric aerosol particles (e.g. sulphur dioxide from anthropogenic sources) that in turn negatively influence downwelling solar radiation reaching the Earth's surface. The term P_{it} complements these components and represents any 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 components of the global common shock U_{gt}. Both U_{gt} and P_{it} can be considered latent variables within R_{it} and, consequently, T_{it+1} in (1).

The generating process (6) for the logarithm of the CO_2 equivalent series $\ln(CO_{2,t})$ is in trend components form comprising a stochastic (unit root) trend with a deterministic drift and a stationary shock u_{ct}. The linear drift term $\delta_{c1} t$ captures the steady annual accumulation of atmospheric greenhouse gases (currently measured around 36 gigatonnes[3]). The stochastic trend component $\delta_c' U_{gt}$ involves a factor loading vector δ_c that captures the global impact on $\ln(CO_2)$ of the accumulated shocks from various greenhouse gas and other sources that are embodied in the latent variable U_{gt}. Importantly, some elements of δ_c may be zero and the corresponding components of U_{gt} do not then figure in determining

[3]National Oceanic & Atmospheric Laboratory (NOAA): (www.esrl.noaa.gov).

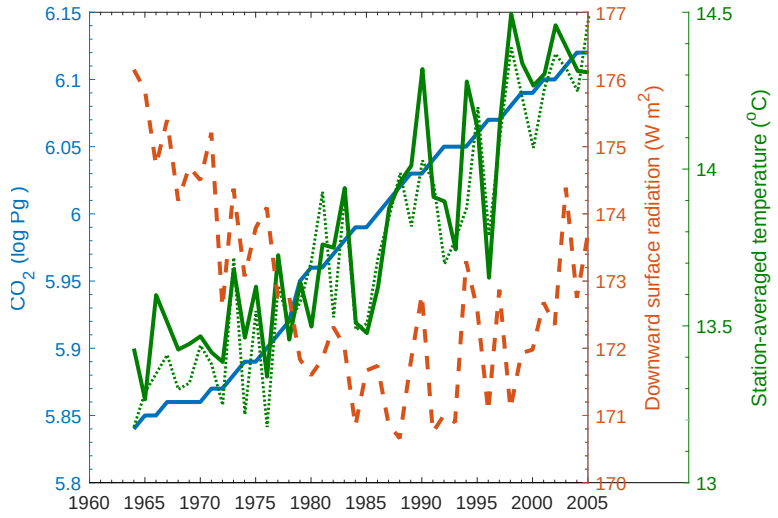

Figure 1: Station-averaged temperature (°Celsius; green solid), Global land temperature (°Celsius; green dotted), Downward surface radiation (Watts per m^2; red dashed), and logarithms of CO_2 (Pg: metric gigatons; blue solid) 1964-2005

$\ln(CO_2)$. This specification allows flexibility so that the determinants in $\delta_c' U_{gt}$ may differ entirely or possibly partly overlap with those that appear in $\delta_{ri}' U_{gt}$ and determine R_{it}.

The formulation in (6) is compatible with the observed data on $\ln(CO_{2t})$ over the historical period 1964-2005, as is apparent from the plot of the time series shown in Figure 1 and the tests reported in Storelvmo et al (2016). Similarly, the random wandering character of the observed aggregate time series $\bar{R}_t$ shown in Figure 1 and the tests in Storelvmo et al (2016) are compatible with the formulation (5). Figure 1 also graphs the surface-averaged (over stations) land temperature series $\bar{T}_t$, which closely follows the stochastic drift in $\ln(CO_{2t})$ with subperiod volatility that mirrors the wandering character of $\bar{R}_t$, particularly over the latter half of the sample period. The time paths of the global variables $(\bar{T}_t, \bar{R}_t, CO_{2t})$ shown in Figure 1 provide visual evidence of linkages in these nonstationary variables, corroborating the evidence for cointegration in these aggregate time series that was found in Storelvmo et al (2016).

Figure 1 includes two curves for global land temperature, both measured on the same (extreme) right hand scale. The solid green curve shows station-averaged surface temperature observations over the land stations used in the present study. The dotted green curve shows the global land temperature time series calculated from data obtained on the NOAA website[4]. The proximity of the two curves indicates that the number of stations and the spatial coverage of these stations are sufficient for the station-averaged quantities to provide a good representation of the general land temperature time series reported in

[4]NOAA: https://www.ncdc.noaa.gov/cag/global/time-series.

NOAA[5].

Table B1 in Appendix B provides residual based tests for cointegration among the aggregate variables $(\bar{T}_t, \bar{R}_t, CO_{2t})$. These results are strongly confirmatory of a long run linkage among these three variables taken together but show no direct linkage between the two component variables $(\bar{R}_t, CO_{2t})$ or between $(\bar{T}_t, CO_{2t})$. This confirms the role that $\bar{R}_t$ and CO_{2t} play jointly in the long run determination of $\bar{T}_t$.

To develop an asymptotic theory for estimation of the key parameter TCS we make the following high level assumptions about the aggregration processes leading to the global variables $(\bar{T}_t, \bar{R}_t, CO_{2t})$ and the component U_{gt} of latent global forcing variables. These conditions will be valid under a wide range of primitive conditions that allow for weak temporal dependence, some cross section dependence, and cross section heterogeneity in the various innovation sequences driving both station-level and global shocks.

Assumption A

(i) $\bar{R}_t^0 := N^{-1}\sum_{i=1}^N R_{it}^0 \to_{a.s.} R^0, \quad N^{-1/2}\sum_{i=1}^N \left(R_{it}^0 - R^0\right) = O_p(1),$
$\bar{\delta}_r := N^{-1}\sum_{i=1}^N \delta_{ri} \to_{a.s.} \delta_r, \quad N^{-1/2}\sum_{i=1}^N (\delta_{ri} - \delta_r) = O_p(1),$
and the u_{it} *satisfy* $(nN)^{-1/2}\sum_{t=1}^n \sum_{i=1}^N u_{it} = O_p(1)$.

(ii) $P_{it} = P_{it}^0 + \sum_{k=1}^t u_{ik}^P =: P_{it}^0 + U_{it}^P, \quad \bar{P}_t^0 = N^{-1}\sum_{i=1}^N P_{it}^0 \to_{a.s.} P^0,$
$N^{-1/2}\sum_{i=1}^N \left(P_{it}^0 - P^0\right) = O_p(1)$ *for all* t, *and the partial sums* U_{it}^P *satisfy the invariance principle* $n^{-1/2}U_{it}^P \Rightarrow U_i^P(r) \equiv BM\left(\sigma_{iP}^2\right)$, *a sequence of independent scalar Brownian motions with variance* $\sigma_{iP}^2 > 0$.

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

(iv) $|\beta_1| < 1$, $|\beta_1 + \gamma_1| < 1$, and $T_0 = O_p(1)$.

(v) $\frac{n}{N} \to 0$ *as* $n, N \to \infty$.

The first part of Assumption A(i) is a strong law for global averages of the stationary component R_{it}^0 of local downwelling radiation and a strong law for global averages of the idiosyncratic loadings δ_{ri}. These hold under standard moment and cross-section dependence conditions. The second part of A(i) is satisfied under related conditions that ensure the operation of a central limit theorem (CLT) for these cross section averages of R_{it}^0 and

[5]The time series graphed as the dotted curve in Figure 1 is obtained from the global land temperature time series in NOAA which is recorded in annual increments. To normalize this incremental series with the station-averaged temperature series graphed as the solid green curve in Figure 1, the station averaged time series observation for 1964 is added to the incremental time series for global land temperature from NOAA.

δ_{ri}. The final part of A(i) holds if the equation errors u_{it} in (1) satisfy a joint central limit theorem of the form $(nN)^{-1/2}\sum_{t=1}^{n}\sum_{i=1}^{N}u_{it} \Rightarrow \mathcal{N}(0,\sigma_u^2)$, jointly as $(n,N)\to\infty$, for some $\sigma_u^2 > 0$, which will be so under general primitive conditions on moments and cross section and time series dependence (e.g., Phillips and Solo, 1992; and Phillips and Moon, 1999).

Assumption A(ii) decomposes P_{it} into two components. The stationary part P_{it}^0 satisfies a cross-section strong law and CLT; the nonstationary part $\sum_{k=1}^{t}u_{ik}^P$ is a partial sum process that satisfies an invariance principle upon suitable normalization. This assumption also holds under standard primitive conditions on the component shocks of the system. A(iii) places a similar invariance principle requirement on the partial sum processes U_{gt} and U_{ct}. This specification accords with a body of empirical evidence supporting unit autoregressive roots in observed climate data (Storelvmo et al, 2016; Kaufmann et al, 2006a, 2006b, 2011, 2013) and the limit theory in the following section is developed under these conditions.

Some relaxation of the unit root condition is possible. But, as in standard cointegrating regression, pivotal inference in the case of near integrated or mildly integrated regressors (Phillips and Magdalinos, 2007) is challenging. One successful approach that leads to standard inferental methods relies on the use of endogenous (so-called IVX) instrumental variables (Phillips and Magdalinos, 2009; Kostakis et al, 2015). This method may be used in the present context to achieve greater generality but the limit theory for such regressions needs to be developed to accommodate co-movement in the regressors and a mixture of trending mechanisms, a task that is not undertaken here and is left for future research.

A(iv) imposes a system-wide transient adjustment condition on the parameters β_1 and γ_1 and a stable initial condition $T_0 = O_p(1)$ for global temperature. A(v) is a simple rate condition requiring the number of stations N to dominate the time series sample size n asymptotically. This condition simplifies the limit theory and appears acceptable in the present context where the actual sample sizes are $N = 1484, n = 42$.

A final matter of importance is that the time series for $\ln(CO_{2,t})$ shows clear evidence of a deterministic drift upwards over the entire sample period, as is evident in Figure 1. That drift is manifestly associated with the long run behavior of global temperature $\bar{T}_t$. It also affects the asymptotic theory and convergence rates of the various parameter estimates, particularly in view of the presence of both stochastic and deterministic trends in the component processes.

2.2 Implications of Aggregation

The panel regression equation (1) implies that station level temperature adjusts to past temperature, local radiation $R_{i,t}$, station level idiosyncratic effects α_i, global influences imported via λ_t, and the panel system errors u_{it+1}. Upon station averaging of (5), we have

$$\bar{R}_t = \bar{R}_t^0 + \bar{\delta}_r' U_{gt} + \bar{P}_t,$$

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 $\bar{\delta}_r = N^{-1}\sum_{i=1}^N \delta_{ri}$. Under A(i) and A(ii) and using the fact that $U_{gt} = O_p\left(\sqrt{t}\right)$ in view of A(iii), the global radiation effect is measured asymptotically as $N \to \infty$ by

$$\bar{R}_t = R^0 + \delta_r' U_{gt} + P_t + O_p\left(\sqrt{\frac{n}{N}}\right) =: R_t + O_p\left(\sqrt{\frac{n}{N}}\right), \tag{8}$$

which evidently imports the nonstationarity of the partial sum process U_{gt} but with a (possibly small) average coefficient effect, measured by the parameter δ_r. To justify (8), note that the local idiosyncratic trend component P_{it} has the stochastic trend representation $P_{it} = P_{it}^0 + U_{it}^P$. The partial sum component $U_{it}^P = \sum_{k=1}^t u_{ik}^P$ is assumed in A(ii) to satisfy a functional law. But at the global level these station specific trends are subject to cross section averaging, so that

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

provided $n/N \to 0$ as $N \to \infty$, which is assumed in A(v).

Thus, when $n/N \to 0$ station-specific stochastic trends such as aerosol pollution average out through global averaging to some mean global level P_t^0. In effect, global averaging of the local pollution trends P_{it} to some mean level implies that some areas may be cleaning up while others are deteriorating over time, leading to a net average effect that is negligible or constant.[6] If there is any general global trend in pollution (say) then it can be considered part of one of the components of the global effect U_{gt}. Thus, any common aerosol pollution trends that may be present in local radiation are assumed to be absorbed in the latent common global shock U_{gt} and manifest via the individual factor loading δ_{ri}. Unlike the local trend effects in $\bar{P}_t$ that average out asymptotically, common trends that are embodied in U_{gt} do have persistent effects in the model.

It follows that the extended model (5) for local radiation impacts global radiation effects in a form that can be represented under the above assumptions as

$$R_t = \delta_{r0} + \delta_r' U_{gt} + O_p\left(\sqrt{\frac{n}{N}}\right), \quad \text{where } \delta_{r0} = R^0 + P_0. \tag{9}$$

Under A(vi) (9) implies that $R_t = \delta_{r0} + \delta_r' U_{gt} + o_p(1)$. These conditions mean that global downwelling radiation is modeled as a unit root stochastic trend driven by the common

[6] Assumption A(ii) may be weakened to allow for stationarity in P_{it}^0 across section only, so that $\bar{P}_t^0 \to_p P_t^0$ with $P_t^0 = O_p(1)$. In this case, the results obtained below continue to apply with some modification in the derivations because P_t^0 is absorbed by a weakly dependent equation error in the aggregate relation and any trend in pollution is carried via the global trend effect $\delta_r' U_{gt}$ in (8).

global stochastic trend U_{gt} with average loading factor δ_r, and initial conditions determined by a linear combination of mean local radiation (R^0) and aerosol pollution (P^0). The stochastic trend limit representation (9) of global downwelling radiation accords well with the observed time series aggregate shown in Figure 1.

3 Econometric Implications

3.1 Global Cointegrating Linkage

Aggregating (1) over stations gives

$$\bar{T}_{t+1} = \bar{\alpha} + \beta_1 \bar{T}_t + \beta_2 \bar{R}_t + \lambda_t + \frac{1}{N}\sum_{i=1}^{N} u_{it+1} = \bar{\alpha} + \beta_1 + \beta_2 \bar{R}_t + \lambda_t + \bar{u}_{\cdot t+1}, \qquad (10)$$

where $\bar{\alpha} = N^{-1}\sum_{i=1}^{N}\alpha_i$ and $\bar{u}_{\cdot t+1} = N^{-1}\sum_{i=1}^{N} u_{it+1}$. Following standard practice for identification purposes, it is convenient to set $\bar{\alpha} = 0$. Substituting (2) then gives the global relationship

$$\bar{T}_{t+1} = \gamma_0 + (\beta_1 + \gamma_1)\,\bar{T}_t + (\beta_2 + \gamma_2)\,\bar{R}_t + \gamma_3 \ln(CO_{2,t}) + \bar{u}_{\cdot t+1}. \qquad (11)$$

Setting $\theta_1 = \beta_1 + \gamma_1$ and $\theta_2 = \beta_2 + \gamma_2$, write (11) as

$$\bar{T}_{t+1} = \gamma_0 + \theta_1 \bar{T}_t + \theta_2 \bar{R}_t + \gamma_3 \ln(CO_{2,t}) + \bar{u}_{\cdot t+1},$$

Using the fact that $|\theta_1| < 1$ under A(v), (11) may be solved to give a stochastic trend representation of $\bar{T}_t$ and to deliver a long run cointegrating relationship among the global variables $\left(\bar{T}_t, \bar{R}_t, \ln(CO_{2,t})\right)$. With some manipulations the results for the common trend expressions can be obtained and are detailed in the following result.

Theorem 1 (Common trend drivers) *Under Assumption A, $W_t = \left(\bar{T}_t, \bar{R}_t, \ln(CO_{2,t})\right)'$ is a vector of stochastically and deterministically trending time series driven by U_{gt} of the form $W_t = \delta_{w0} + \delta_{w1} t + \Delta_w U_{gt} + u_{wt}^{+}$ where $u_{wt}^{+} = (u_{Tt}, 0, u_{ct})' + o_p(1)$ is asymptotically stationary. Specifically*

$$W_t = \begin{bmatrix} \bar{T}_t \\ \bar{R}_t \\ \ln(CO_{2,t}) \end{bmatrix} = \begin{bmatrix} \delta_{T0} + \delta_{T1} t + \delta_T' U_{gt} + u_{Tt}^{+} \\ \delta_{r0} + \delta_r' U_{gt} + O_p\left(\sqrt{\frac{n}{N}}\right) \\ \delta_{c0} + \delta_{c1} t + \delta_c' U_{gt} + u_{ct} \end{bmatrix} =: \delta_{w0} + \delta_{w1} t + D_w U_{gt} + u_{wt}^{+}, \qquad (12)$$

where

$$\delta_{T1} = \frac{\gamma_3\delta_{c1}}{1-\theta_1},\ \delta_T = \frac{\theta_2\delta_r+\gamma_3\delta_c}{1-\theta_1}, \tag{13}$$

$$u_{Tt} = \gamma_3\sum_{j=0}^{\infty}\theta_1^j u_{ct-1-j} - \frac{\theta_1}{1-\theta_1}\sum_{k=0}^{\infty}\theta_1^k\left[\theta_2\delta_r+\gamma_3\delta_c\right]' u_{gt-1-k} - \delta_T' u_{gt}, \tag{14}$$

$$u_{Tt}^{+} = u_{Tt} + O_p\left(\frac{1}{\sqrt{N}}+\sqrt{\frac{n}{N}}+t\left|\theta_1\right|^t\right), \tag{15}$$

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

Remarks

1. Since $\delta_T = \frac{\theta_2\delta_r+\gamma_3\delta_c}{1-\theta_1}$, it is apparent from (12) that $\bar{T}_t$ is cointegrated with $\left(\bar{R}_t, \ln(CO_{2,t})\right)$. In particular, we have

$$\begin{aligned}
&(1-\theta_1)\bar{T}_t = \gamma_3\delta_{c1}t + \left[\theta_2\delta_r+\gamma_3\delta_c\right]' U_{gt} + \delta_{T0}(1-\theta_1) + u_{Tt}^{+}(1-\theta_1)\\
=\ &\theta_2\left(\delta_{r0}+\delta_r' U_{gt}\right) + \gamma_3\left(\delta_{c0}+\delta_{c1}t+\delta_c' U_{gt}\right) + \left\{\delta_{T0}(1-\theta_1) - \theta_2\delta_{r0} - \gamma_3\delta_{c0}\right\} + u_{Tt}^{+}(1-\theta_1)\\
=\ &\theta_2\bar{R}_t + \gamma_3\ln(CO_{2,t}) + \mu + \zeta_t,
\end{aligned} \tag{16}$$

   with $\mu = \delta_{T0}(1-\theta_1) - \theta_2\delta_{r0} - \gamma_3\delta_{c0}$ and $\zeta_t = u_{Tt}(1-\theta_1) - \gamma_3 u_{ct} + o_p(1)$. We may write this global cointegrating relation as

$$\bar{T}_t = \frac{\theta_2}{1-\theta_1}\bar{R}_t + \frac{\gamma_3}{1-\theta_1}\ln(CO_{2,t}) + \frac{\mu}{1-\theta_1} + \frac{1}{1-\theta_1}\zeta_t. \tag{17}$$

   Importantly for applied work on climate sensitivity, the coefficient of $\ln(CO_{2,t})$ in this relationship gives the transient climate sensitivity parameter

$$TCS = \frac{\gamma_3}{1-\theta_1}\times\ln(2) = \frac{\gamma_3}{1-\beta_1-\gamma_1}\times\ln(2) \tag{18}$$

   upon scaling by $\ln(2)$.

2. The cointegrating relation (17) involves deterministic and stochastic cointegration. The linear drift in global temperature $\bar{T}_t$ moves with the linear drift in $\ln(CO_{2,t})$, giving deterministic drift co-movement between these two variables. The stochastic trends in $\ln(CO_{2,t})$ and $\bar{R}_t$ co-move with the stochastic trend in $\bar{T}_t$, producing stochastic cointegration among these three variabes. Equation (17) therefore describes how the trends that are inherent in greenhouse gases and downwelling radiation impact global temperature in the long run. It is this relationship rather than equation (2)

that captures the cointegrating linkage among the global variables. Assuming radiation is constant, the formula (18) for climate sensitivity to CO_2 is an immediate consequence of this long run relationship among the global variables. Estimation of the TCS parameter is then most conveniently performed by direct estimation of (17).

3. This framework and Theorem 1 therefore support a single long run cointegrating relationship among the global variables $\left(\bar{T}_t, \bar{R}_t, \ln(CO_{2,t})\right)$. This relationship is given by (17) so that

$$(1-\beta_1)\,\bar{T}_t - \beta_2 \bar{R}_t - \lambda_t \simeq_a I(0),$$

where $\simeq_a I(0)$ signifies asymptotically integrated of order zero. As noted, the time specific effect λ_t carries both a deterministic trend and stochastic trend. The dynamic panel regression equation (1) then describes transient adjustments of station-specific temperature T_{it+1} about a long run trend driven by global components in the time specific effect λ_t and idioscyncratic trends that may be present in local radiation R_{it}.

4. Define the cointegrating vector

$$b_\gamma' = \left[\begin{array}{ccc} \theta_1 - 1 & \theta_2 & \gamma_3 \end{array}\right] \tag{19}$$

for which $b_\gamma' \delta_{w1} = 0$ and $b_\gamma' D_w = 0$, so that the $3 \times m_g$ matrix D_w has rank 2. Let b_w be an orthonormal matrix complement of b_γ, and write the $3 \times m_g$ matrix D_w in (12) in the outer product form $D_w = b_w a'$ for some 3×2 matrix b_w and $2 \times m_g$ matrix a'. Then,

$$W_t = \delta_{w0} + \delta_{w1} t + b_w a' U_{gt} + u_{wt}^+, \tag{20}$$

and

$$b_\gamma' W_t = b_\gamma' \delta_{w0} + b_\gamma' u_{wt}^+ \simeq_a I(0)$$

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

$$W_t = \delta_w + \delta_{w1} t + b_w U_{wt} + u_{wt}^+. \tag{21}$$

3.2 Estimation and Asymptotic Theory

The goal of the present paper is to estimate the transient climate sensitivity parameter TCS. This may be achieved most simply by estimating the cointegrating relation (17) and scaling the coefficient of $\ln(CO_{2,t})$ by $\ln(2)$. In addition to simplicity, this approach has the advantage that it allows for general assumptions concerning the cross section and time series dependence properties of the station level data and innovations. The latter is helpful

since dynamic panel GMM regression techniques rely for consistent estimation on much stronger assumptions such as serial independence of the equation errors in (1).[7]

Asymptotically efficient estimates of the coefficients in (17) can be obtained under quite general assumptions on the errors ζ_t by a variety of methods such as fully modified least squares (FM-OLS: Phillips and Hansen, 1990), dynamic least squares (DOLS: Saikonnen, 1991; Phillips and Loretan, 1991; Stock and Watson, 1993), or trend instrumental variable regression (TIV: Phillips, 2014).

It is convenient to write the cointegrating regression equation (17) in the form

$$\begin{aligned} \bar{T}_t = \;& d_0 + d_1 \bar{R}_t + d_2 \ln(CO_{2,t}) + \zeta_{d,t} =: d_0 + d' x_t + \zeta_{d,t}, \\ d_0 \quad = \quad & \frac{\mu}{1-\theta_1},\; d_1 = \frac{\theta_2}{1-\theta_1},\; d_2 = \frac{\gamma_3}{1-\theta_1},\; \zeta_{dt} = \frac{1}{1-\theta_1}\zeta_t. \end{aligned} \tag{22}$$

The regressor $x_t = \left(\bar{R}_t, \ln(CO_{2,t})\right)'$ in (22) has common trend representation given in (12), which it is convenient to write in the subset form

$$x_t = \left[\begin{array}{c} \delta_{r0} + \delta_r' U_{gt} + O_p\left(\sqrt{\frac{n}{N}}\right) \\ \delta_{c0} + \delta_{c1} t + \delta_c' U_{gt} + u_{ct} \end{array} \right] =: \delta_{x0} + \delta_{x1} t + D_x U_{gt} + u_{xt},$$

with $\delta_{x0}' = (\delta_{r0}, \delta_{c0})$, $\delta_{x1}' = (0, \delta_{c1})$, $D_x = [\delta_r, \delta_c]'$, and $u_{xt} = \left(O_p\left(\sqrt{\frac{n}{N}}\right), u_{ct}\right)'$.

The asymptotic theory for the regression fitting of (22) needs to take into account the form of the common trend drivers of the two regressors in x_t, the presence of a linear trend in $\ln(CO_{2,t})$, the absence of cointegration between the two regressors, and the presence of an intercept in the regression. The linear trend in the regressor x_t is specific to the component regressor $\ln(CO_{2,t})$ and therefore dominates the asymptotic theory of estimates of the regression coefficient d_2, leading to a convergence rate of $n^{3/2}$ for both least squares and efficient estimates of this coefficient. However, as shown in Park and Phillips (1988) the presence of a deterministic drift in the regressors complicates multivariate regression asymptotics with unit root regressors (as distinct from regressions with a single regressor) because of the presence of full rank (non-cointegrated) stochastic trends in x_t. This complication means that the limit theory is mixed normal ($\mathcal{MN}$) and involves projection residuals of the dominating (continuous time) linear trend on the (limiting) stochastic trend components of x_t as well as the constant intercept in (22). The result in Theorem 2 below gives the required limit theory for the FM-OLS estimates of the slope coefficients in (22). Other asymptotically efficient estimators such as DOLS and TIV have the same limit distribution.

The FM-OLS estimator of the slope coefficient vector d in (22) takes the usual form

$$\hat{d}^+ = \left(\tilde{X}'\tilde{X}\right)^{-1}\left(\tilde{X}'\widehat{\tilde{y}_t^+} - n\hat{\Delta}_{x\zeta}^+\right). \tag{23}$$

[7] On the other hand, when those additional assumptions are valid, panel cross section averaging has the advantage that it can raise the rate of convergence of the coefficient estimates by a factor of $\sqrt{N}$.

In (23) we use the tilde affix to denote demeaned variables. Thus, $\tilde{X}$ is the $n \times 2$ moment matrix of demeaned observations $\tilde{x}_t = x_t - \bar{x}$, and $\widehat{\tilde{y}_t^+}$ is the vector of observations of the endogeneity corrected demeaned observations of the endogenous variable. More specifically, $\widehat{\tilde{y}_t^+} = \tilde{y}_t - \hat{\Omega}_{\zeta x}\hat{\Omega}_{xx}^{-1}\widetilde{\Delta x}_t$, where $\left(\hat{\Omega}_{\zeta x}, \hat{\Omega}_{xx}\right)$ are consistent estimators of the long run covariance matrices $(\Omega_{\zeta x}, \Omega_{xx}) = \left(\sum_{h=-\infty}^{\infty} \mathbb{E}\left(\zeta_{d0} u'_{gh} D_x\right), \sum_{h=-\infty}^{\infty} D_x \mathbb{E}\left(u_{g0} u'_{gh}\right) D'_x\right)$. The quantity $n\hat{\Delta}_{x\zeta}^+$ in (23) is a second-order bias correction term, where $\hat{\Delta}_{x\zeta}^+ := \hat{\Delta}_{x0} - \hat{\Delta}_{xx}\hat{\Omega}_{xx}^{-1}\hat{\Omega}_{x\zeta}$, and $\left(\hat{\Delta}_{x0}, \hat{\Delta}_{xx}\right)$ are consistent estimates of the one-sided long run covariance matrices

$$(\Delta_{x0}, \Delta_{xx}) = \left(\sum_{h=0}^{\infty} D_x \mathbb{E}\left(u_{g,0}\zeta_{d,h}\right), \sum_{h=0}^{\infty} D_x \mathbb{E}\left(u_{g,0} u'_{g,h}\right) D'_x\right).$$

These consistent estimates are obtained in the usual manner by lag kernel methods applied to sample covariances of the residuals, including those obtained from a first stage consistent cointegrating regression on (22) by least squares. Readers are referred to Phillips and Hansen (1990) and Phillips (1995) for details.

Theorem 2 (FM-OLS Limit theory) *Under Assumption A, we have:*

(i) $n\left(\hat{d}_1^+ - d_1\right) \rightsquigarrow \left(\int_0^1 \underline{\tilde{U}}_R(p)^2 dp\right)^{-1} \left\{\int_0^1 \underline{\tilde{U}}_R(p) dB_{\zeta_d \cdot x}\right\} =_d \mathcal{MN}\left(0, \omega_{\zeta_d \cdot x}^2 \left(\int_0^1 \underline{\tilde{U}}_R(p)^2 dp\right)^{-1}\right)$,

(ii) $n^{3/2}\left(\hat{d}_2^+ - d_2\right) \rightsquigarrow \left(\delta_{c1}^2 \int_0^1 \tilde{r}_x(p)^2 dp\right)^{-1} \left\{\delta_{c1} \int_0^1 \tilde{r}_x(p) dB_{\zeta_d \cdot x}\right\} =_d \mathcal{MN}\left(0, \omega_{\zeta_d \cdot x}^2 \left(\delta_{c1}^2 \int_0^1 \tilde{r}_x(p)^2 dp\right)^{-1}\right)$,

where $\rightsquigarrow$ signifies weak convergence, $=_d$ denotes equivalence in distribution,

$$\begin{aligned}
\underline{\tilde{U}}_R(p) &= \tilde{U}_R(p) - \left(\int_0^1 \tilde{U}_R(s)\tilde{r}(s) ds\right)\left(\int_0^1 \tilde{r}^2(s) ds\right)^{-1} \tilde{r}(p), \\
\tilde{r}_x(p) &= \tilde{r}(p) - \int_0^1 s\tilde{U}_R(s) ds \left(\int_0^1 \tilde{U}_R(s)^2 ds\right)^{-1} \tilde{U}_R(p), \\
\tilde{r}(p) &= p - \int_0^1 s ds,\ \tilde{U}_R(p) = U_R(p) - \int_0^1 U_R(s) ds,\ U_R(p) = \delta_r' U_g(p),
\end{aligned}$$

and $B_{\zeta_d \cdot x}(r)$ is Brownian motion with variance $\omega_{\zeta_d \cdot x}^2$ where

$$\omega_{\zeta \cdot x}^2 = \omega_\zeta^2 - \Omega_{\zeta x}\Omega_{xx}^{-1}\Omega_{x\zeta},\ \omega_{\zeta_d \cdot x}^2 = \frac{\omega_{\zeta \cdot x}^2}{(1-\theta_1)^2},\ \omega_\zeta^2 = \sum_{h=-\infty}^{\infty} \mathbb{E}\left(\zeta_0 \zeta_h\right). \tag{24}$$

Remarks

5. Long run covariance matrices (24) rather than variances appear in the limit formulae (i) and (ii) and are associated with the Brownian motion $B_{\zeta_d \cdot x}$. These covariance matrices are needed, even under *iid* conditions on the shocks (u_{ct}, u_{gt}) in Assumption A, because the constituent error processes $\zeta_t = u_{Tt}(1-\theta_1) - \gamma_3 u_{ct} + o_p(1)$ and $\zeta_{dt} = \frac{1}{1-\theta_1}\zeta_t$ that enter the cointegrating regression (22) depend on the full history of the shocks (u_{ct}, u_{gt}) rather than contemporaneous shocks. In particular, ζ_t depends on u_{Tt}, as well as u_{ct}, and by definition (14) we have $u_{Tt} = \gamma_3 \sum_{j=0}^{\infty} \theta_1^j u_{ct-1-j} - \frac{\theta_1}{1-\theta_1} \sum_{k=0}^{\infty} \theta_1^k \left[\theta_2 \delta_r + \gamma_3 \delta_c\right]' u_{gt-1-k} - \delta_T' u_{gt}$. In view of the simple autoregressive linear process form of u_{Tt}, it is clear that Theorem 2 continues to hold after suitable adjustment in the long run variance formulae when the component shocks (u_{ct}, u_{gt}) are themselves temporally dependent linear processes under weak summability and moment conditions on the components. The conditional long run variance $\omega^2_{\zeta_d \cdot x}$ in (24) takes account of the regressor endogeneity and corresponds to standard limit theory for efficient cointegrating regressions (Phillips, 1991).

6. The time series $n^{3/2}$ rate of convergence in (ii) is the consequence of two special circumstances in the model: (a) the presence of the linear trend in the generating mechanism (6) of $\ln(CO_{2t})$; and (b) the fact that the pertinent coefficient of interest is d_2, which is the coefficient of $\ln(CO_{2t})$ in the cointegrating regression (22), whereas the other regressor $\bar{R}_t$ in the equation does not have a linear trend by virtue of the generating mechanism (5). Thus, the direction of the linear trend in the model is specific to that of the regressor $\ln(CO_{2t})$.

7. Asymptotic standard errors for the components of the FM-OLS estimate $\hat{d}^+ = \left(\hat{d}_1^+, \hat{d}_2^+\right)'$ are obtained in the standard way using the square roots of the diagonal elements of the matrix $\hat{\omega}^2_{\zeta_d \cdot x} \left(\tilde{X}'\tilde{X}\right)^{-1}$ in spite of asymptotic singularity in the moment matrix. The long run conditional variance $\omega^2_{\zeta_d \cdot x}$ is estimated in the usual manner by $\hat{\omega}^2_{\zeta_d \cdot x}$ by using residuals from the fitted FM-OLS regression. As shown in the proof of Theorem 2, these elements provide consistent estimates of the limit variances in (i) and (ii). In particular, from (56) in the proof we have

$$\left(\tilde{X}'\tilde{X}\right)^{-1} = \begin{pmatrix} \left(\tilde{X}_r' Q_c \tilde{X}_r\right)^{-1} & -\left(\tilde{X}_r'\tilde{X}_r\right)^{-1} \tilde{X}_r'\tilde{X}_c \left(\tilde{X}_c' Q_r \tilde{X}_c\right)^{-1} \\ -\left(\tilde{X}_c' Q_r \tilde{X}_c\right)^{-1} \tilde{X}_c'\tilde{X}_r \left(\tilde{X}_r'\tilde{X}_r\right)^{-1} & \left(\tilde{X}_c' Q_r \tilde{X}_c\right)^{-1} \end{pmatrix},$$

which employs the projection matrix notation $Q_a = I - \tilde{X}_a \left(\tilde{X}_a'\tilde{X}_a\right)^{-1} \tilde{X}_a'$ for $a \in \{r, c\}$. Then, from (57) and (58) in the proof we deduce that

$$n^2 \left(\tilde{X}_r' Q_c \tilde{X}_r\right)^{-1} \rightsquigarrow \left(\int_0^1 \underline{\tilde{U}}_R(r)^2 \, dr\right)^{-1}, \text{ and } n^3 \left(\tilde{X}_c' Q_r \tilde{X}_c\right)^{-1} \rightsquigarrow \left(\delta_{c1}^2 \int_0^1 \tilde{r}_x(p)^2 \, dp\right)^{-1},$$

as required for consistent estimation of the variances by standard regression formulae. Accordingly, asymptotic $100\,(1-\alpha)\,\%$ confidence regions for d_1 and d_2 are given by

$$\hat{d}_1^+ \pm z_a \left\{ \hat{\omega}^2_{\zeta_d \cdot x} \left(\tilde{X}_r' Q_c \tilde{X}_r \right)^{-1} \right\}^{1/2}, \quad \hat{d}_2^+ \pm z_a \left\{ \hat{\omega}^2_{\zeta_d \cdot x} \left(\tilde{X}_c' Q_r \tilde{X}_c \right)^{-1} \right\}^{1/2} \tag{25}$$

where z_a is the $100\,(1-\alpha/2)\,\%$ percentile of the standard normal distribution.

8. The corresponding estimate of the TCS parameter is $\widehat{TCS} = \hat{d}_2^+ \ln(2)$ and an asymptotic $100\,(1-\alpha)\,\%$ confidence region for TCS is then

$$\widehat{TCS} \pm z_a \ln(2) \left\{ \hat{\omega}^2_{\zeta_d \cdot x} \left(\tilde{X}_c' Q_r \tilde{X}_c \right)^{-1} \right\}^{1/2}. \tag{26}$$

9. The limit theory (i) and (ii) holds as both $n \to \infty$ and $N \to \infty$ under the condition A(v). With this condition it follows from Theorem 1 that the cross section averages $\left(\bar{T}_t, \bar{R}_t\right)$ take the following form

$$\begin{aligned} \bar{T}_t = &\quad \delta_{T0} + \delta_{T1} t + \delta_T' U_{gt} + u_{Tt} + O_p\left(t\,|\theta_1|^t\right) =: T_t + O_p\left(t\,|\theta_1|^t\right), \\ \bar{R}_t = &\quad \delta_{r0} + \delta_r' U_{gt} + O_p\left(\sqrt{\frac{n}{N}}\right) =: R_t + O_p\left(\sqrt{\frac{n}{N}}\right), \end{aligned}$$

and the asymptotic properties of the cointegrating regression are therefore determined by the relationship

$$T_t^A = d_0 + d_1 R_t^A + d_2 \ln(CO_{2,t}) + \zeta_{d,t} + O_p\left(\sqrt{\frac{n}{N}} + t\,|\theta_1|^t\right) \tag{27}$$

among the (full) global time series aggregates $(T_t, R_t, \ln(CO_{2,t}))$. The error in (27) can be neglected asymptotically as $N \to \infty$ because $\frac{n}{N} \to 0$ by A(v) and the cross moments of the time trend t and stochastic trend U_{gt} with the $O\left(t\,|\theta_1|^t\right)$ error are all negligible asymptotically since the sum $\sum_{t=1}^n t^k\,|\theta_1|^t = O(1)$ for all $k \geq 1$ when $|\theta_1| < 1$ and, consequently, $\sum_{t=1}^n tU_{gt}\,|\theta_1|^t = O_p(1)$ as $n \to \infty$, which makes the cross moment components negligible in asymptotic moment calculations. The cointegrating equation $T_t = d_0 + d_1 R_t + d_2 \ln(CO_{2,t}) + \zeta_{d,t}$ may therefore be interpreted as a full cross section limiting average relationship linking these global aggregate time series[8].

Table B2 reports results from the estimation of (22) by the three time series methods FM-OLS, DOLS and OLS as well as the panel within group (WG) regression method.

[8]Under spatial ergodicity as $N \to \infty$, the quantities (T_t, R_t) may be interpreted as limiting conditional expectations given the invariant sigma algebra at time t and the relationship $T_t = d_0 + d_1 R_t + d_2 \ln(CO_{2,t}) + \zeta_{d,t}$ describes how these limiting global aggregates cointegrate with $\ln(CO_{2,t})$.

FM-OLS is computed using an automated lag length selector (with AR prefiltering and recoloring) to determine the number of lags used in the calculation of the long run variance $\hat{\omega}^2_{\zeta_d \cdot x}$. DOLS(m,m) is run with $m \in \{1, 2, 3\}$ lags and leads in the regression. As is apparent from Table B2, the TCS parameter estimates by the different methods are similar, ranging from $\widehat{TCS} = 3.1637$ and 2.9175 for DOLS(3) and DOLS(1) to $\widehat{TCS} = 2.8020$ for FM-OLS, and $\widehat{TCS} = 2.7012$ for WG estimation. From standard asymptotic theory (Phillips and Durlauf, 1986), OLS is known to suffer from second order bias because of endogeneity and serial correlation and the present results indicate that the OLS bias is in the upwards direction, since the OLS estimate exceeds that of the FM-OLS estimate, which is known to be asymptotically unbiased and efficient (Phillips and Hansen, 1990). As might be expected DOLS(1,1) produces an estimate that is closest to OLS. The FM-OLS estimated standard errors are uniformly smaller than those of DOLS and the latter increase with the number of leads and lags included in the DOLS regression, which is explained by increasing multicollinearity in the DOLS regression as the number of lead and lag differences included in the DOLS regression rises. The 95% confidence interval $(2.360, 3.243)$ for TCS from FM-OLS regression is the shortest interval, consonant with the fact that FM-OLS is asymptotically efficient. The corresponding intervals for the DOLS regressions are much wider in view of the higher standard errors.

An alternative estimation approach is to use dynamic panel regression methods to estimate both (1) and (2). This was the approach used in Storelvmo et al. (2016). Various methods are available, including within-group (WG) least squares estimation, difference GMM (diff-GMM; Arellano and Bond, 1991) and system GMM (sys-GMM; Blundell and Bond, 1998) with various choices of instruments in the GMM regressions. Sys-GMM methods were used by Magnus et al. (2011) and Storelvmo et al. (2016) on the grounds that this method is considered preferable in cases where the panel autoregressive coefficient β_1 is close to unity[9] because such cases typically lead to the use of weak instruments in diff-GMM, thereby reducing efficiency. However, as might be expected from the global coverage of the station locations in the present application, there is considerable heterogeneity in the fixed effects α_i of the dynamic panel regression (1). This feature is known to produce sys-GMM estimates of the coefficients in dynamic panel regression that are substantially biased and can even be inconsistent (Hayakawa, 2007, 2015).

The properties of these panel regression methods, including asymptotic theory, are analyzed in other ongoing work and are not needed for the present study. We mention here that simulations based on data generated from (1) and (2) and using the actual data for R_{it} and $\ln(CO_{2t})$ as inputs showed that sys-GMM produced estimates of β_1 that were heavily biased upwards (by almost 700%), whereas WG and diff-GMM delivered far better performance with little bias. Recent work (Phillips, 2018) has shown that these dynamic panel regression methods do provide robust estimates of the TCS parameter, even though

[9] Magnus et al. (2011) and Storelvmo et al. (2016) obtained estimates (0.9063 and 0.9212, respectively) of β_1 that were both in the vicinity of unity.

the individual parameter estimates (such as β_1) of the dynamic panel regression and energy balance relationship (such as γ_1) differ considerably with each other. For comparison purposes we therefore report here in Table 2B the WG estimate, 2.7012, of the transient climate sensitivity parameter which is comparable to the FM-OLS estimate but with a wider confidence interval.

4 Inference on Earth's Climate Sensitivity

This section of the paper reports applications of the above methods to the study of Earth's transient climate sensitivity. These applications use two sources of information: (i) spatio-temporal empirical observations; and (ii) climate model simulation data computed using the same spatio-temporal coordinates with outputs from several leading climate models. The data sources are described below. Some summary statistics of the main features of the data are provided before reporting the results of the econometric analysis.

4.1 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 to ensure a balanced panel 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.

4.1.1 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 1484 land-based 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 stations in which data for 24 consecutive months were available at least once during the time span 1964-2005. Any data gaps in the remaining months were filled using a machine learning approach ('random forests, see e.g. Breiman, 2001), which is described in more detail in Storelvmo et al. (2018). The missing values are spread unevenly over time, and often only one month is missing for a given year. The infilling of missing values allowed us to increase the number of GEBA stations included in

the study by approximately 200 compared to Storelvmo et al. (2016).We thus obtained a matrix of 1484x42 annual mean surface radiation observations.

4.1.2 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 4.00), available at a spatial resolution of 0.5°. Each of the 1484 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 1484×42 data matrix.

4.1.3 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 climate forcing 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 tens to hundreds 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 global CO_2 time series.

4.2 Climate Model Data

We use data from nine of the GCMs that participated in the Coupled Model Intercomparison Project - Phase 5 (CMIP5, Taylor et al., 2012), see Table 1 for the models and their salient features. We use data from their historical simulations, run from 1850 to 2005, forced with changing GHG and aerosol concentrations (Lamarque et al., 2010). While some of the models produced only one historical simulation, others produced ensembles of simulations. The ensemble members 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.

Table 1: Overview of salient GCM features: horizontal atmospheric model resolution (Resolution), global transient climate sensitivity (TCS_G) and ratio of warming over ocean to that over land, derived from the CMIP5 historical simulations ($\frac{\Delta T_{O,hist}}{\Delta T_{L,hist}}$).

Short Name	Long name (Country)	Resolution	TCS_G	$\frac{\Delta T_{O,hist}}{\Delta T_{L,hist}}$
BCC	BCC-CSM1.1 (China)	$\sim 2.8 \times 2.8^\circ$	1.7K	0.56
BNU	BNU-ESM (China)	$\sim 2.8 \times 2.8^\circ$	2.6K	0.74
CNRM	CNRM-CM5 (France)	$\sim 1 \times 1^\circ$	2.1K	0.88
CSIRO	CSIRO-Mk3.6.0 (Australia)	$\sim 1.9 \times 1.9^\circ$	1.8K	0.86
GFDL-2M	GFDL-ESM2M (USA)	$\sim 2.5 \times 2^\circ$	1.3K	0.56
GFDL-2G	GFDL-ESM2G (USA)	$\sim 2.5 \times 2^\circ$	1.1K	0.71
HadGEM2	HadGEM2-ES (UK)	$\sim 1.9 \times 1.3^\circ$	2.5K	1.03
INM	INM-CM4 (Russia)	$\sim 2 \times 1.5^\circ$	1.3K	0.61
MPI	MPI-ESM-MR (Germany)	$\sim 1.9 \times 1.9^\circ$	2.0K	0.64

The CMIP5 data archive contains output from a total of more than 30 different GCMs. For the present analysis we have selected nine of these and grouped them into three categories (High-TCS, Medium-TCS and Low-TCS) according to their model-reported TCS_G values (see Flato et al., 2013).[10]

4.3 Econometric Analysis of Observational and Climate Model Data

A primary empirical motivation for the present study was to determine whether econometric analysis applied to panel observations as in Magnus et al. (2011) and Storelvmo et al. (2016) could successfully determine TCS_G estimates corresponding to those reported for the GCMs. The empirical exercise applies the same analysis to the GCM data with GCM simulation output being included only where observational data (as used in the observational study) is available. One measure of success in this exercise is the extent to which the TCS_G emerging from the econometric analysis agrees with the reported value for the GCM in question – more specifically the extent to which the GCM reported TCS_G values lie within the 95% confidence interval calculated from observational data as implied by Theorem 2 and indicated in (26) of Remark 8. Furthermore, differences between observed and modeled sensitivities to radiation and equivalent CO_2, as measured by differences in the estimates of TCS_L and the parameters of equations (1) and (2), can reveal possible

[10]The model-reported TCS_G values were calculated numerically by running a GCM simulation in which atmospheric CO_2 was increased by 1% per year until doubling was reached (after 70 years). The TCS_G was then calculated as the global mean temperature difference between the last and the first decade of simulation.

Table 2: Mean, standard deviation, minimum and maximum for the annual change in station-averaged temperature (upper table) and downwelling solar radiation at the surface (lower table) from observations and for GCMs with high reported TCS_G (Group A).

	Temperature (K)			
	Mean	St.dev.	Min.	Max.
Observations	0.021	0.239	-0.471	0.535
HadGEM2 r_1	0.019	0.151	-0.224	0.308
HadGEM2 r_2	0.010	0.190	-0.366	0.381
HadGEM2 r_3	0.017	0.133	-0.309	0.356
HadGEM2 r_4	0.015	0.162	-0.286	0.364
BNU	0.023	0.306	-0.780	0.636
CNRM	0.018	0.159	-0.372	0.495
	Radiation ($\mathbf{Wm^{-2}}$)			
	Mean	St.dev.	Min.	Max.
Observations	-0.059	1.147	-2.595	2.394
HadGEM2 r_1	-0.028	1.071	-1.983	2.459
HadGEM2 r_2	0.001	1.273	-2.919	2.784
HadGEM2 r_3	0.018	1.394	-3.609	3.553
HadGEM2 r_4	-0.017	1.086	-3.567	2.281
BNU	-0.010	1.528	-3.034	2.950
CNRM	-0.023	1.027	-2.067	2.392

GCM model shortcomings with respect to policy-important parameters that may not be so readily evident with standard procedures or simple graphical comparisons.

Tables 2, 3 and 4 report summary statistics for mean annual changes in temperature and radiation for the three TCS categories (high, medium and low groupings). The mean annual change in observed temperature is 0.021°C, with an estimated standard deviation of 0.239°C. The observed mean change in temperature (0.021) is similar to the mean change in the GCM simulations in most cases, which ranges from 0.009 (INM) to 0.023 (GFDL 2G). The standard deviations for the GCM simulations are broadly consistent with the observed standard deviation, ranging from 0.126 (CSIRO r_1) to 0.306 (BNU). These descriptive figures indicate that the GCMs fit observed global average temperature reasonably well although some GCMs tend above and others below observed temperature. This finding is corroborated by the curves shown in Figure 2, which trace the simulated and observed evolution in temperatures for the time period 1964 - 2005.

The mean annual change in the observed downward solar radiation is -0.059, with a standard deviation of 1.147. As the radiation time series in Figure 3 shows, there is a negative drift in radiation until the early nineties, when the drift shifts to become positive with mild fluctuations. This observed pattern impacts the sample mean, as the annual

Table 3: Mean, standard deviation, minimum and maximum for the annual change in station-averaged temperature (upper table) and downwelling solar radiation at the surface (lower table) from observations and for GCMs with medium reported TCS_G (Group B).

	Temperature (K)			
	Mean	St.dev.	Min.	Max.
Observations	0.021	0.239	-0.471	0.535
CSIRO r_1	0.011	0.137	-0.265	0.231
CSIRO r_2	0.016	0.126	-0.217	0.349
CSIRO r_3	0.014	0.127	-0.237	0.292
CSIRO r_4	0.015	0.135	-0.357	0.302
MPI	0.025	0.151	-0.334	0.282
BCC r_1	0.020	0.173	-0.546	0.323
BCC r_2	0.025	0.191	-0.630	0.321
BCC r_3	0.019	0.203	-0.513	0.453
	Radiation ($\mathbf{Wm^{-2}}$)			
	Mean	St.dev.	Min.	Max.
Observations	-0.059	1.147	-2.595	2.394
CSIRO r_1	0.012	0.855	-2.447	1.879
CSIRO r_2	0.011	0.959	-2.267	2.112
CSIRO r_3	-0.015	1.534	-2.715	3.759
CSIRO r_4	0.012	1.233	-2.841	3.940
MPI	-0.039	1.265	-2.479	2.038
BCC r_1	0.042	1.108	-2.575	2.250
BCC r_2	0.047	1.398	-2.121	4.704
BCC r_3	0.027	1.405	-2.830	4.986

Table 4: Mean, standard deviation, minimum and maximum for the annual change in station-averaged temperature (upper table) and downwelling solar radiation at the surface (lower table) from observations and for GCMs with low reported TCS_G (Group C).

	Temperature (K)			
	Mean	St.dev.	Min.	Max.
Observations	0.021	0.239	-0.471	0.535
GFDL 2G	0.038	0.234	-0.737	0.405
GFDL 2M	0.018	0.164	-0.326	0.406
INM	0.009	0.119	-0.261	0.217
	Radiation ($\mathbf{Wm^{-2}}$)			
	Mean	St.dev.	Min.	Max.
Observations	-0.059	1.147	-2.595	2.394
GFDL 2G	-0.035	1.271	-3.522	2.509
GFDL 2M	0.014	1.491	-4.079	2.987
INM	-0.019	1.035	-2.345	2.573

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, about half of the models actually produce a positive mean annual change, and none of the models produce a negative trend of the magnitude seen in the observations. MPI comes closest, with a mean annual change of -0.039Wm^{-2} per year. This apparent shortcoming in the GCM simulations in reproducing the observed pattern of radiation trend 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 subset of GCMs considered here, and more generally for the entire CMIP5 archive in Storelvmo et al. (2018), therefore suggests that these models may be underestimating the aerosol 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 finding appears important for the global climate modeling community and will be examined further in a more extensive investigation of all simulations in the CMIP5 archive and eventually also in CMIP6 (Eyring et al., 2016) in our future research.

We next turn to the estimated TCS_L parameter. The TCS_L parameter is a nonlinear function of the panel system and energy balance parameters that is obtained as a cointegrating autoregressive solution of the system, as shown in Section 3. The TCS_L parameter

Table 5: Parameter estimates, standard errors, and 95% confidence intervals of Transient Climate Sensitivity for land (TCS_L) using observational and GCM simulation data (Group A) by FM-OLS regression, and estimated and reported TCS_G.

	TCS_L Parameter: $\frac{\gamma_3}{1-\beta_1-\gamma_1} \times ln(2)$			
	Estimate	St. Error	95% Confidence Interval	Estimated (reported) TCS_G
Observ.	2.8021	0.2254	(2.360,3.244)	2.05
BNU r_1	2.7618	0.3011	(2.171,3.352)	2.25 (2.6)
CNRM r_1	1.6361	0.3242	(1.001,2.272)	1.51 (2.1)
HadGEM2 r_1	2.5308	0.4398	(1.668,3.392)	2.58 (2.5)
HadGEM2 r_2	1.6232	0.3739	(0.890,2.356)	1.65 (2.5)
HadGEM2 r_3	1.5239	0.1952	(1.141,1.906)	1.55 (2.5)
HadGEM2 r_4	1.0874	0.4837	(0.139,2.035)	1.11 (2.5)

Table 6: Parameter estimates, standard errors, and 95% confidence intervals of Transient Climate Sensitivity (TCS) using observational and GCM simulation data (Group B) by FM-OLS regression, and estimated and reported TCS_G.

	TCS_L Parameter: $\frac{\gamma_3}{1-\beta_1-\gamma_1} \times ln(2)$			
	Estimate	St. Error	95% Confidence Interval	Estimated (reported) TCS_G
Observ.	2.8021	0.2254	(2.360,3.244)	2.05
BCC r_1	1.9200	0.2771	(1.376,2.463)	1.32 (1.7)
BCC r_2	1.9823	0.1900	(1.609,2.354)	1.37 (1.7)
BCC r_3	1.8932	0.4065	(1.096,2.690)	1.30 (1.7)
CSIRO r_1	1.4328	0.3066	(0.831,2.033)	1.30 (1.8)
CSIRO r_2	1.3619	0.3551	(0.665,2.057)	1.23 (1.8)
CSIRO r_3	2.2217	0.2559	(1.720,2.723)	2.01 (1.8)
CSIRO r_4	1.5137	0.2256	(1.072,1.956)	1.37 (1.8)
MPI r_1	2.1539	0.4777	(1.216,3.089)	1.61 (2.0)

Table 7: Parameter estimates, standard errors, and 95% confidence intervals of Transient Climate Sensitivity for land (TCS_L) using observational and GCM simulation data (Group C) by FM-OLS regression, and estimated and reported TCS_G.

	TCS_L Parameter: $\frac{\gamma_3}{1-\beta_1-\gamma_1} \times ln(2)$			
	Estimate	St. Error	95% Confidence Interval	Estimated (reported) TCS_G
Observ.	2.8021	0.2254	(2.360,3.244)	2.05
GFDL2G	2.7110	0.6325	(1.471,3.950)	2.16 (1.1)
GFDL2M	1.8461	0.3465	(1.166,2.525)	1.28 (1.3)
INM r_1	0.5792	0.3377	(-0.082,1.241)	0.58 (1.3)

is $ln(2)$ times the linear coefficient of the $ln(CO_{2t})$ regressor in this cointegrating equation. The linear and stochastic trends in $ln(CO_{2t})$ ensure a strong signal from this regressor that reduces the standard error of the TCS_L estimate and narrows the confidence intervals for both the observational data and the GCM output data, as implied by Theorem 2(ii) and the form of (26).

Tables 5-7 report the empirical findings for FM-OLS estimates of the TCS_L parameter obtained from the observational data, the observational TCS_G values and the observational and reported TCS_G values from GCM simulations that fall in the High-TCS, Medium-TCS and Low-TCS categories, respectively. Note that because we only have data over land, we convert to a global estimate (TCS_G) by following the procedure of Storelvmo et al. (2016), but here use an observed ratio of ocean to land warming of 0.62 (see Appendix C for details). This ratio is an average based on four observational data sets of land and ocean warming over the past century, as reported in Hartmann et al. (2013). Using this ratio yields an observationally estimated TCS_G value of 2.05K with a corresponding 95% confidence interval $(1.73, 2.37)$. This estimated TCS_G value is remarkably similar to the 2.0K estimate of Storelvmo et al. (2016), despite the many revisions to the estimation framework described above, in addition to changes in the length and station-width of the data record and the use of machine learning to fill gaps in the observations. The observational estimate is also consistent with the most sensitive models among the latest generation of GCMs (Flato et al., 2013) and falls well within, but in the upper end of, the range of TCS_G estimates compiled in a recent comprehensive review of all published TCS_G estimates available (Knutti et al., 2017). Similar conversions are necessary to estimate TCS_G values for the GCMs. The ratio of land to ocean warming for each of the models is based on their respective historical simulations for the time period 1900 to 2000. The resulting values are listed in Table 1. Both for the observations and the GCMs, there

is uncertainty associated with the ratio of ocean to land warming, which we take into account by adding a generous error of +/- 0.05 to the ratio (see Appendix C for the exact procedure). In Figure 4 we show TCS_G estimates for the observations and all nine GCMs (and their respective ensemble members when relevant), along with the reported TCS_G estimates from the GCMs. Also shown for the observational estimates are error bars that here represent the 95% confidence interval and the added uncertainty due to the ratio of ocean to land warming.

For 7 out of 9 GCMs the reported TCS values lie within the estimated error bars, which is encouraging. The two models for which there is disagreement between the reported and estimated TCS values are HadGEM2 and INM. The estimated TCS value is in excellent agreement with the reported value (2.6 vs. 2.5K) for one of the HadGEM2 ensemble members ($r1$). Accounting for the uncertainty in the land/ocean warming ratio, which widens the confidence interval slightly, the estimated TCS for HadGEM2 r_2 is also just inside the uncertainty bounds. But the reported TCS values are outside of the confidence intervals for the other two ensemble members ($r3$ and $r4$). This is a puzzling result. Different ensemble members should differ only in their internal variability as a consequence of their slightly different initial conditions, but otherwise should be expected to have similar long-term trends. As evident from Table 2 and Figures 2 and 3, the station-averaged annual mean changes and trends appear different for the four HadGEM2 ensemble members. Perhaps not surprisingly, $r1$ with a strong positive temperature trend and a relatively strong negative radiation trend, yields a high estimated TCS value in the econometric analysis. In contrast, the other three HadGEM2 ensemble members yield much too weak station-averaged temperature trends, and little or no radiation trend. It is not surprising that simulation data with this combination of temperature and radiation trends are interpreted econometrically as being generated by a model with low TCS. It is possible that the limited number of stations and the relatively short observational time period available for this study causes ensemble members from the same model to appear more different than when seen through a global-mean and longer-term lens. Testing how sensitive the estimated TCS values are to additional data in both time and space will be an interesting extension to this study. Finally, although the estimated TCS_G value for INM is the lowest of all, consistent with its reported TCS value being the lowest one, the estimated confidence interval does not include the reported value. It is generally noteworthy that the estimated values are lower than the reported ones in 15 of the 17 analysed simulations. This outcome could be the result of the method by which the reported estimate is computed. Specifically, there is now considerable evidence that climate feedbacks are state dependent (e.g., Armour, 2017), and that the TCS inferred from historical simulations may consistently be lower than that evaluated from a simulation that is run to a state with doubled atmospheric CO_2. It is therefore possible that the observational TCS estimate reported here is a conservative one, but this warrants more detailed study.

5 Overview of Methods and Computation

In contrast to Storelvmo et al (2016), which employed dynamic panel methods to estimate the parameters of the spatial model (1), this paper uses a time series approach to estimation and confidence interval construction for the transient climate sensitivity parameter TCS. By its nature TCS is a global climate parameter. The analysis in Section 2 shows how TCS arises as one of the coefficients in an aggregate balancing relationship by solving the dynamic equation (1) for local temperature and then spatially aggregating. The resulting global equation relates the three key variables of the study $\left(\bar{T}_t, \bar{R}_t, \ln(CO_{2,t})\right)$ in the linear relation

$$\bar{T}_t = d_0 + d_1 \bar{R}_t + d_2 \ln(CO_{2,t}) + \zeta_{d,t} =: d_0 + d' x_t + \zeta_{d,t}, \tag{28}$$

from which the TCS parameter is obtained by a simple scaling of the coefficient d_2 as $TCS = d_2 \times \ln(2)$.

Equation (28) defines a global long run relationship among the observable aggregate variables, linking temperature $\bar{T}_t$ to downwelling radiation $\bar{R}_t$ and $CO_{2,t}$, subject to an unobserved disturbance $\zeta_{d,t}$. The relationship characterizes the aggregate impact of $CO_{2,t}$ with its deterministic and stochastic trend properties on global temperature $\bar{T}_t$, while controlling for the effects of downwelling radiation $\bar{R}_t$. With this formulation based on an underlying model that respects spatial behavior and trend characteristics in the data, the time series approach attempts to accommodate the main factors involved in determining the transient effects (at the time of doubling) of $CO_{2,t}$ on temperature.

Practical implementation involves the use of data – either empirical observation or GCM simulation data – to estimate equation (28) by regression. Simple least squares methods may be used. But due to the trending nature of the data, the joint dependence of temperature and the explanatory variables, and serial dependence in both the data and the equation errors, some modifications to least squares are needed to achieve good asymptotic properties (as $n \to \infty$) that validate simple confidence interval construction for estimates of TCS. The method used in the asymptotic development here to validate inference is fully modified least squares (FM-OLS) but another method that is commonly used in applied econometric work is dynamic OLS (DOLS). These two methods make simple adjustments to least squares by means of corrections for potential bias and serial correlation. Both methods are available in standard software packages and matrix programming language routines[11]; and both methods are used in the reported empirical application.

As shown in Section 3.2, when equation (28) is estimated by FM-OLS or DOLS, the resulting estimate of TCS is $\widehat{TCS} = \hat{d}_2^+ \times \ln(2)$, where $\hat{d}_2^+$ is the estimate of the regression coefficient $\hat{d}_2$. An asymptotically valid 95% confidence interval for TCS may then be

[11] The R package 'cointReg' has routines for both FM-OLS and DOLS, as well as other methods. See https://cran.r-project.org/web/packages/cointReg/.

constructed as follows

$$\widehat{TCS} \pm 1.96 \times \ln(2) \left\{ \hat{\omega}^2_{\zeta_d \cdot x} \left(\tilde{X}_c' Q_r \tilde{X}_c \right)^{-1} \right\}^{1/2}. \tag{29}$$

In this expression, the estimated variance (within the braces in (29)) relies on two quantities. The first, $\hat{\omega}_{\zeta_d \cdot x}$, is an estimate of the conditional variance of the equation error $\zeta_{d,t}$ in (28) which takes into account long run effects over time (notably the serial dependence in $\zeta_{d,t}$) and the joint dependence (endogeneity) of the regressors $\left(\bar{R}_t, CO_{2,t}\right)$ with $\zeta_{d,t}$. The second quantity, $\tilde{X}_c' Q_r \tilde{X}_c$, is the sample second moment of the regressor $\ln(CO_{2,t})$ adjusted for the presence of the intercept and the regressor $\bar{R}_t$ in the fitted equation. These quantities are now standard expressions and extend the usual formula for the estimated variance of a least squares regression coefficient by taking account of the special features present in this application: the nonstationarity in the data, the endogeneity of the regressors, and the serial dependence in the errors.

Figure 5 provides a visual roadmap of the sequence of operations involved in the transition from the dynamic panel model through the aggregation process to the fitted regression equation and the computation of the TCS estimate and the 95% confidence interval. The implementation is the same for both the observed data and the simulated GCM data.

6 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 nine leading global climate models (some with multiple ensemble members) generally do reproduce actual temperature trajectories over nearly half a century to 2005, when averaged across the approximately 1400 surface stations considered in this study. Our analysis further revealed that even though simulated station-averaged temperature trends are broadly consistent with observations, the negative trend in incoming solar radiation over the same stations and time period was manifestly underestimated in the GCM simulation data. The analysis supports the notion that this could be explained by an underestimation of aerosol cooling in the GCMs, which would allow the GCMs to reproduce the observed temperature record with an incorrect TCS.

The present application provides an observational-data-based confidence interval for Earth's transient climate sensitivity to greenhouse gas emissions, and a best observational TCS estimate of 2.05K using aggregate time series methodology. The same method was

applied to GCM output from the CMIP5 archive, and for some models produced estimates that lie at the margin or outside of the 95% confidence interval for TCS, pointing to some potential misspecifications in these models. The fact that the TCS values reported for the GCMs generally lie within the 95% TCS confidence interval estimated based on the model output instills confidence in the ability of the econometric framework to correctly diagnose TCS, though a few notable exceptions will require follow-up research. A more extensive investigation of this matter will be undertaken in a subsequent study that extends this analysis to the full archive of climate models and provides estimates of the full system parameters as well as TCS.

In modeling climate data, theory restrictions that incorporate global energy forces play a key role in empirical modeling. As shown in the present study, the 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 regularity conditions that allow for weak temporal dependence and cross section dependence in the model innovations. 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 and station-level variables. The asymptotic results indicate that, with appropriate methods and regularity conditions, inference from spatially aggregated climate data is possible that allows for signal degeneracies, trend effects, and co-movement in the data at both the transient and aggregate levels.

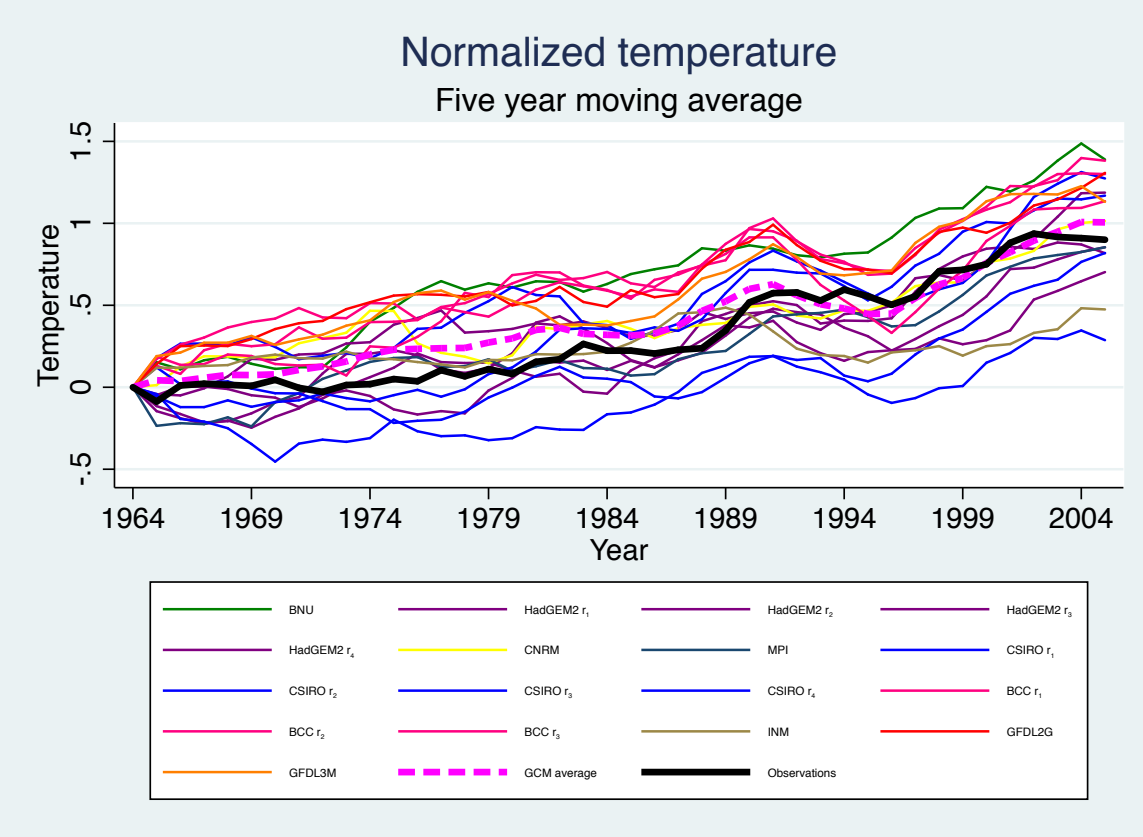

Figure 2: Station-averaged temperature change (in Kelvin) for observations and GCMs.

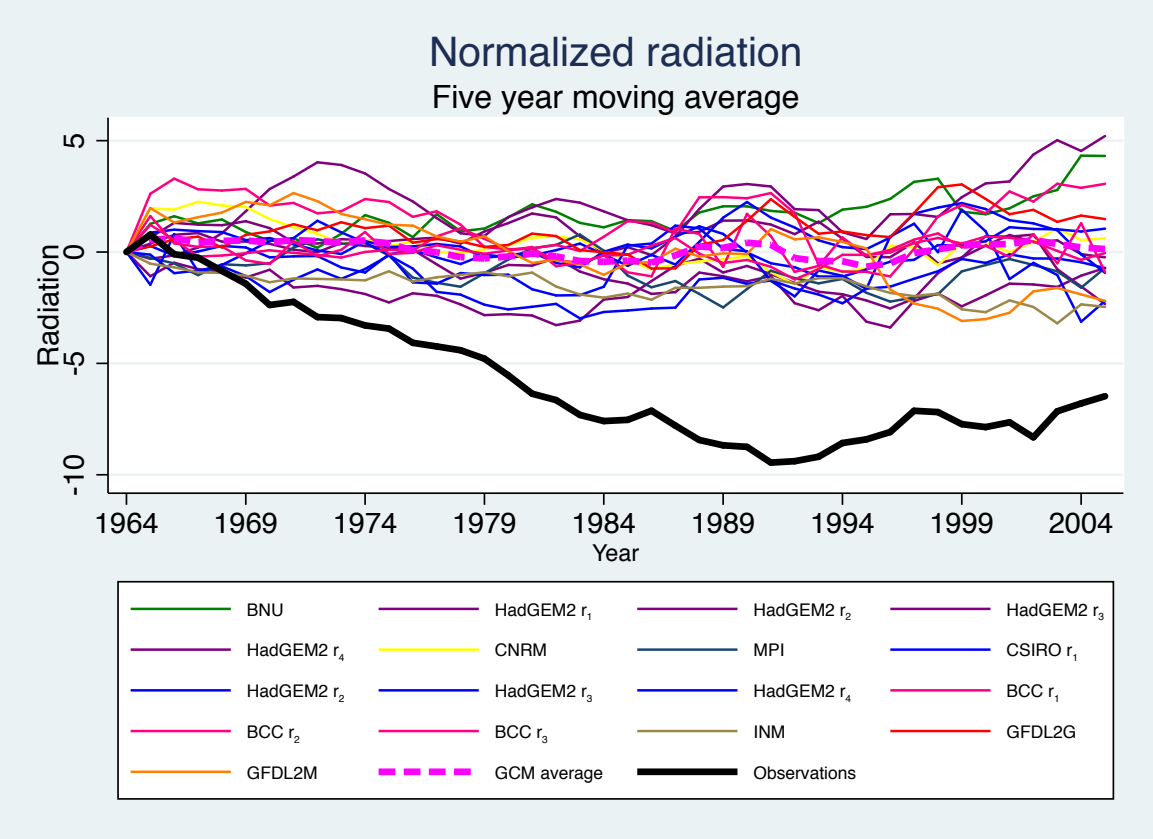

Figure 3: Station-averaged radiation change (in Wm^{-2}) for observations and GCMs.

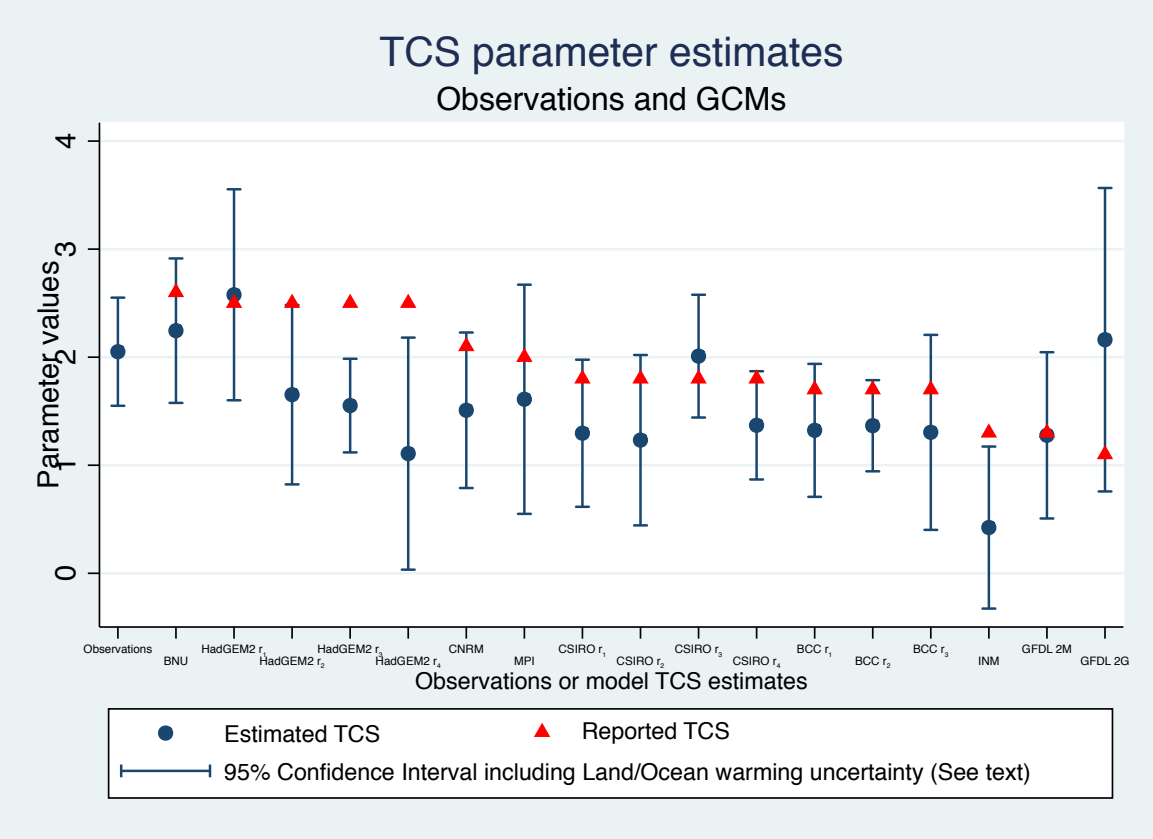

Figure 4: Estimated TCSs (blue symbols) and associated confidence intervals resulting from the econometric analysis, and reported TCS values for the GCMs (red symbols).

7 Appendix A

7.1 The TCS Formula

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

$$\begin{aligned} \bar{T}_{t+1} &= \bar{\alpha} + \beta_1 + \beta_2 \bar{R}_t + \lambda_t + o_p(1) \\ &= \bar{\alpha} + \beta_1 + \beta_2 \bar{R}_t + \gamma_0 + \gamma_1 + \gamma_2 R_t + \gamma_3 \ln(CO_{2,t}) + o_p(1), \end{aligned} \tag{30}$$

since $N^{-1}\sum_{i=1}^{N} u_{it+1} \to_p 0$ under Assumption A(i). Solving (30) leads to the equilibrium solution

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

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

$$d\bar{T} = \frac{\beta_2 + \gamma_2}{1 - \beta_1 - \gamma_1} d\bar{R} + \frac{\gamma_3 d\ln(CO_2)}{1 - \beta_1 - \gamma_1}, \tag{31}$$

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 (31) by

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

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

$$TCS = \frac{\gamma_3}{1 - \beta_1 - \gamma_1} \times \ln(2), \tag{32}$$

giving (4).

In a similar way spatial aggregation of model (3) leads to

$$\bar{T}_{t+1} = \bar{\alpha} + \beta_1 + \beta_2 \bar{R}_t + \bar{\phi}\{\gamma_0 + \gamma_1 + \gamma_2 R_t + \gamma_3 \ln(CO_{2,t})\} + o_p(1),$$

where $\bar{\phi} = N^{-1}\sum_{i=1}^{N} \phi_i$. Then

$$TCS = \frac{\bar{\gamma}_3}{1 - \beta_1 - \bar{\gamma}_1} \times \ln(2), \tag{33}$$

where $\bar{\gamma}_i = \bar{\phi}\gamma_i$ for $\{i = 0, 1, 2, 3\}$.

7.2 Proof of Theorem 1

Setting $\theta_1 = \beta_1 + \gamma_1$ and $\theta_2 = \beta_2 + \gamma_2$, write the aggregate relation (11) as

$$\bar{T}_{t+1} = \gamma_0 + \theta_1 + \theta_2 \bar{R}_t + \gamma_3 \ln(CO_{2,t}) + \bar{u}_{\cdot t+1}. \tag{34}$$

The stochastic trend representations of $\bar{R}_t$ and $\ln(CO_{2,t})$ follow directly from (9) and (6), giving

$$\begin{bmatrix} \bar{R}_t \\ \ln(CO_{2,t}) \end{bmatrix} = \begin{bmatrix} \delta_{r0} + \delta_r' U_{gt} + O_p\left(\sqrt{\frac{n}{N}}\right) \\ \delta_{c0} + \delta_{c1} t + \delta_c' U_{gt} + u_{ct} \end{bmatrix}, \quad U_{gt} = \sum_{s=1}^{t} u_{gs}. \tag{35}$$

To establish the stated representation, we proceed as follows. Solving (34) for $\bar{T}_{t+1}$ in terms of the past history of the inputs $\left(\bar{R}_t, \ln(CO_{2,t})\right)$ leads to the representation

$$\begin{aligned} \bar{T}_{t+1} &= \gamma_0 \sum_{j=0}^{t} \theta_1^j + \theta_2 \sum_{j=0}^{t} \theta_1^j \bar{R}_{t-j} + \gamma_3 \sum_{j=0}^{t} \theta_1^j \ln(CO_{2,t-j}) + \theta_1^{t+1} T_0 + \sum_{j=0}^{t} \theta_1^j \bar{u}_{\cdot t+1} \\ &= \frac{\gamma_0}{1-\theta_1} + \theta_2 \sum_{j=0}^{t} \theta_1^j \bar{R}_{t-j} + \gamma_3 \sum_{j=0}^{t} \theta_1^j \ln(CO_{2,t-j}) + O_p\left(\frac{1}{\sqrt{N}} + |\theta_1|^t\right), \end{aligned} \tag{36}$$

since $\gamma_0 \sum_{j=0}^{t} \theta_1^j = \frac{\gamma_0}{1-\theta_1} - \gamma_0 \frac{\theta_1^{t+1}}{1-\theta_1}$, $\overline{T}_0 = O_p(1)$, and

$$\mathbb{E}\left(\sum_{j=0}^{t} \theta_1^j \bar{u}_{\cdot t+1}\right)^2 = \frac{\sigma_u^2}{N^2} \sum_{i=1}^{N} \sum_{j=0}^{t} \theta_1^2 = \frac{\sigma_u^2}{N} \frac{1-\theta_1^{2(t+1)}}{1-\theta_1^2} = O\left(\frac{1}{N}\right).$$

Next substitute the stochastic trend representations of R_t and $\ln(CO_{2,t})$ from (35) into (36), giving

$$\begin{aligned} \bar{T}_{t+1} &= \frac{\gamma_0}{1-\theta_1} + \theta_2 \sum_{j=0}^{t} \theta_1^j \left\{\delta_{r0} + \delta_r' U_{gt-j} + O_p\left(\sqrt{\frac{n}{N}}\right)\right\} \\ &\quad + \gamma_3 \sum_{j=0}^{t} \theta_1^j \left\{\delta_{c0} + \delta_{c1} t + \delta_c' U_{gt-j} + u_{ct-j}\right\} + O_p\left(\frac{1}{\sqrt{N}} + |\theta_1|^t\right) \\ &= \frac{\gamma_0 + \theta_2 \delta_{r0} + \gamma_3 \delta_{c0}}{1-\theta_1} + \sum_{j=0}^{t} \theta_1^j \left[\theta_2 \delta_r + \gamma_3 \delta_c\right]' U_{gt-j} + \gamma_3 \delta_{c1} \sum_{j=0}^{t} \theta_1^j (t-j) \\ &\quad + \gamma_3 \sum_{j=0}^{t} \theta_1^j u_{ct-j} + O_p\left(\frac{1}{\sqrt{N}} + \sqrt{\frac{n}{N}} + |\theta_1|^t\right) \end{aligned}$$

$$
\begin{aligned}
&= \frac{\gamma_0 + \theta_2\delta_{r0} + \gamma_3\delta_{c0} - \frac{\theta_1}{1-\theta_1}}{1-\theta_1} + \frac{\gamma_3\delta_{c1}}{1-\theta_1}t + \frac{[\theta_2\delta_r + \gamma_3\delta_c]' U_{gt}}{1-\theta_1} + u_{Tt} \\
&\quad + O_p\left(\frac{1}{\sqrt{N}} + \sqrt{\frac{n}{N}} + t\left|\theta_1\right|^t\right), \qquad (37)
\end{aligned}
$$

where u_{Tt} is defined below in (48). The final line (37) is now demonstrated in several steps. Observe first that

$$
\begin{aligned}
\sum_{j=0}^{t}\theta_1^j(t-j) &= t\frac{1-\theta_1^{t+1}}{1-\theta_1} - \sum_{j=0}^{t}\theta_1^j j = \frac{t}{1-\theta_1} + \frac{\theta_1^{t+1}(1-\theta_1 t)}{(1-\theta_1)^2} - \frac{\theta_1}{(1-\theta_1)^2} \\
&= \frac{t}{1-\theta_1} - \frac{\theta_1}{(1-\theta_1)^2} + O\left(t\left|\theta_1\right|^t\right),
\end{aligned}
$$

and

$$
\begin{aligned}
\sum_{j=0}^{t}\theta_1^j U_{gt-j} &= \sum_{j=0}^{t}\theta_1^j\sum_{s=1}^{t-j}u_{gs} = \sum_{s=1}^{t}u_{gs}\sum_{j=0}^{t-s}\theta_1^j = \sum_{s=1}^{t}u_{gs}\frac{1-\theta_1^{t-s+1}}{1-\theta_1} \\
&= \frac{1}{1-\theta_1}\sum_{s=1}^{t}u_{gs} - \frac{\theta_1}{1-\theta_1}\sum_{s=1}^{t}u_{gs}\theta_1^{t-s} = \frac{1}{1-\theta_1}\sum_{s=1}^{t}u_{gs} - \frac{\theta_1}{1-\theta_1}\sum_{k=0}^{t-1}u_{gt-k}\theta_1^k \\
&= \frac{1}{1-\theta_1}\sum_{s=1}^{t}u_{gs} - \frac{\theta_1}{1-\theta_1}\sum_{k=0}^{\infty}u_{gt-k}\theta_1^k + O_p\left(\left|\theta_1\right|^t\right), \qquad (38)
\end{aligned}
$$

since $\sum_{k=t}^{\infty}\theta_1^{k+1}u_{gt-k} = \theta_1^{t+1}\sum_{j=0}^{\infty}\theta_1^j u_{-j} = O_p\left(\left|\theta_1\right|^t\right)$. Next, observe that

$$
\begin{aligned}
\sum_{j=0}^{t}\theta_1^j U_{gt-j} &= \sum_{j=0}^{t}\theta_1^j\sum_{s=1}^{t-j}u_{gs} = \sum_{s=1}^{t}u_{gs}\sum_{j=0}^{t-s}\theta_1^j = \sum_{s=1}^{t}u_{gs}\frac{1-\theta_1^{t-s+1}}{1-\theta_1} \\
&= \frac{1}{1-\theta_1}\sum_{s=1}^{t}u_{gs} - \frac{1}{1-\theta_1}\sum_{s=1}^{t}u_{gs}\theta_1^{t-s+1} \\
&= \frac{1}{1-\theta_1}\sum_{s=1}^{t}u_{gs} - \frac{\theta_1}{1-\theta_1}\sum_{k=0}^{t-1}u_{gt-k}\theta_1^k \\
&= \frac{1}{1-\theta_1}U_{gt} - \frac{\theta_1}{1-\theta_1}\sum_{k=0}^{\infty}u_{gt-k}\theta_1^k + O_p\left(\left|\theta_1\right|^t\right), \qquad (39)
\end{aligned}
$$

since

$$
\sum_{k=t}^{t-1}u_{gt-k}\theta_1^{k+1} = \theta_1^{t+1}\sum_{j=0}^{\infty}u_{-j}\theta_1^j = O_p\left(\left|\theta_1\right|^t\right). \qquad (40)
$$

It follows that

$$\sum_{j=0}^{t}\theta_1^j\left[\theta_2\delta_r+\gamma_3\delta_c\right]'U_{gt-j}=\frac{\left[\theta_2\delta_r+\gamma_3\delta_c\right]'U_{gt}}{1-\theta_1}+v_t+O_p\left(\left|\theta_1\right|^t\right),$$

where

$$v_t=-\frac{\theta_1}{1-\theta_1}\sum_{k=0}^{\infty}\theta_1^k\left[\theta_2\delta_r+\gamma_3\delta_c\right]'u_{gt-k}, \tag{41}$$

thereby demonstrating that

$$\begin{aligned}\bar{T}_{t+1} &= \frac{\gamma_0+\theta_2\delta_{r0}+\gamma_3\delta_{c0}-\frac{\theta_1}{1-\theta_1}}{1-\theta_1}+\frac{\gamma_3}{1-\theta_1}t+\frac{\left[\theta_2\delta_r+\gamma_3\delta_c\right]'U_{gt}}{1-\theta_1}+v_t\\ &\quad+\gamma_3\sum_{j=0}^{t}\theta^j u_{ct-j}+O_p\left(\frac{1}{\sqrt{N}}+\sqrt{\frac{n}{N}}+t\left|\theta_1\right|^t\right)\\ &= \delta_{T0}^{\#}+\delta_{T1}t+\delta_T'U_{gt}+u_{Tt}^{\#},\end{aligned} \tag{42}$$

with

$$\begin{aligned}\delta_{T0}^{\#} &= \frac{\gamma_0+\theta_2\delta_{r0}+\gamma_3\delta_{c0}-\frac{\theta_1}{1-\theta_1}}{1-\theta_1},\\ \delta_{T1} &= \frac{\gamma_3\delta_{c1}}{1-\theta_1},\ \delta_T=\frac{\theta_2\delta_r+\gamma_3\delta_c}{1-\theta_1},\\ u_{Tt}^{\#} &= v_t+\gamma_3\sum_{j=0}^{\infty}\theta^j u_{ct-j}+O_p\left(\frac{1}{\sqrt{N}}+\sqrt{\frac{n}{N}}+t\left|\theta_1\right|^t\right),\end{aligned} \tag{43}$$

so that $u_{Tt}^{\#}$ is asymptotically stationary. This establishes the equation (37) above. Then

$$\begin{aligned}\bar{T}_t &= \delta_{T0}^{\#}+\delta_{T1}\left(t-1\right)+\delta_T'U_{gt-1}+u_{Tt-1}^{\#}=\delta_{T0}^{\#}+\delta_{T1}t+\delta_T'U_{gt}+u_{Tt-1}^{\#}-\left(\delta_{T1}+\delta_T'u_{gt}\right)\\ &= \delta_{T0}+\delta_{T1}t+\delta_T'U_{gt}+u_{Tt}^{+},\end{aligned}$$

where

$$\delta_{T0}=\delta_{T0}^{\#}-\delta_{T1}=\frac{\gamma_0+\theta_2\delta_{r0}+\gamma_3\delta_{c0}-\frac{\theta_1}{1-\theta_1}}{1-\theta_1}-\delta_{T1}, \tag{44}$$

and

$$u_{Tt}^{+}=u_{Tt-1}^{\#}-\delta_T'u_{gt}=v_{t-1}+\gamma_3\sum_{j=0}^{\infty}\theta_1^j u_{ct-1-j}-\delta_T'u_{gt}+O_p\left(\frac{1}{\sqrt{N}}+\sqrt{\frac{n}{N}}+t\left|\theta_1\right|^t\right). \tag{45}$$

Hence

$$\bar{T}_t=\delta_{T0}+\delta_{T1}t+\delta_T'U_{gt}+u_{Tt}^{+}, \tag{46}$$

with

$$u_{Tt}^{+} = u_{Tt} + O_p\left(\frac{1}{\sqrt{N}} + \sqrt{\frac{n}{N}} + t\left|\theta_1\right|^t\right), \tag{47}$$

where

$$u_{Tt} = v_{t-1} + \gamma_3 \sum_{j=0}^{\infty} \theta_1^j u_{ct-1-j} - \delta_T' u_{gt} \tag{48}$$

is stationary and u_{Tt}^{+} is asymptotically stationary, giving the stated result for $\bar{T}_t$. Combining (46) with (35) we have the stated system for W_t. ■

7.3 Proof of Theorem 2

The proof follows standard lines of derivation for FM-OLS asymptotics with possible degeneracies in the signal matrix of the regressors (Phillips, 1995). The present derivation deals with the additional complexity of a deterministic trend in one of the regressors as well as stochastic trends and intercepts in the equation and the effects of cross section aggregration. It is convenient to write the model (22) in the form

$$y_t = d_0 + d' x_t + \zeta_{d,t}, \tag{49}$$

and note that by definition of x_t

$$x_t = \left[\begin{array}{c} \delta_{r0} + \delta_r' U_{gt} + O_p\left(\sqrt{\frac{n}{N}}\right) \\ \delta_{c0} + \delta_{c1} t + \delta_c' U_{gt} + u_{ct} \end{array}\right] =: \delta_{x0} + \delta_{x1} t + D_x U_{gt} + u_{xt}. \tag{50}$$

with $\delta_{x0}' = (\delta_{r0}, \delta_{c0})$, $\delta_{x1}' = (0, \delta_{c1})$, $D_x = [\delta_r, \delta_c]'$, and $u_{xt} = \left(O_p\left(\sqrt{\frac{n}{N}}\right), u_{ct}\right)'$. Then, denoting demeaned observations using a tilde superscript and first differencing by the (unaffixed) operator Δ, we have

$$\tilde{x}_t = \delta_{x1}\tilde{t} + D_x \tilde{U}_{gt} + \tilde{u}_{xt}, \quad \widetilde{\Delta x}_t = D_x \tilde{u}_{gt} + \widetilde{\Delta u}_{xt} \tag{51}$$

The FM-OLS estimator of the coefficient vector d in (22) has the form

$$\hat{d}^{+} = \left(\tilde{X}'\tilde{X}\right)^{-1}\left(\tilde{X}'\widehat{\tilde{y}_t^{+}} - n\hat{\Delta}_{x\zeta}^{+}\right) \tag{52}$$

where $\tilde{X}$ is the $n \times 2$ matrix of demeaned observations $\tilde{X}_t = X_t - \bar{X}$ and $\widehat{\tilde{y}_t^{+}}$ is the vector of observations of the endogeneity corrected demeaned observations of the endogenous variable, i.e., $\widehat{\tilde{y}_t^{+}} = \tilde{y}_t - \hat{\Omega}_{\zeta x}\hat{\Omega}_{xx}^{-1}\widetilde{\Delta x}_t$, where $\left(\hat{\Omega}_{\zeta x}, \hat{\Omega}_{xx}\right)$ are consistent estimators of the long run covariance matrices $(\Omega_{\zeta x}, \Omega_{xx}) = \left(\sum_{h=-\infty}^{\infty} \mathbb{E}\left(\zeta_{d0} u_{gh}' D_x\right), \sum_{h=-\infty}^{\infty} D_x \mathbb{E}\left(u_{g0} u_{gh}'\right) D_x'\right)$. Note that even in the simple case of *iid* error components as in Assumptions A(i) and A(iv), these long run covariance matrices are needed because the error process $\zeta_t = u_{Tt}(1-\theta_1) - \gamma_3 u_{ct} +$

$o_p(1)$, and hence $\zeta_{dt} = \frac{1}{1-\theta_1}\zeta_t$, depend on the full history of the shocks (u_{ct}, u_{gt}) since, by definition in (14), $u_{Tt} = \gamma_3 \sum_{j=0}^{\infty} \theta_1^j u_{ct-1-j} - \frac{\theta_1}{1-\theta_1} \sum_{k=0}^{\infty} \theta_1^k \left[\theta_2 \delta_r + \gamma_3 \delta_c\right]' u_{gt-1-k} - \delta_T' u_{gt}$. The serial correlation correction in (52) depends on $\hat{\Delta}_{x\zeta}^{+} := \hat{\Delta}_{x0} - \hat{\Delta}_{xx}\hat{\Omega}_{xx}^{-1}\hat{\Omega}_{x\zeta}$, where $\left(\hat{\Delta}_{x0}, \hat{\Delta}_{xx}\right)$ are consistent estimates of the one-sided long run covariance matrices

$$(\Delta_{x0}, \Delta_{xx}) = \left(\sum_{h=0}^{\infty} D_x \mathbb{E}\left(u_{g,0}\zeta_{d,h}\right), \sum_{h=0}^{\infty} D_x \mathbb{E}\left(u_{g,0}u_{g,h}'\right) D_x' \right).$$

These consistent estimates are obtained in the usual manner by lag kernel methods applied to sample covariances of the residuals from a first stage consistent cointegrating regression on (49) by least squares.

As in the usual analysis of FM-OLS regression (Phillips and Hansen, 1990; Phillips, 1995), we construct the following augmented regression equation from (49)

$$\begin{aligned} y_t &= d_0 + d'x_t + \Omega_{\zeta x}, \Omega_{xx}^{-1}\Delta x_t + \zeta_{d\cdot x,t}, \text{ with } \zeta_{d\cdot x,t} = \zeta_{d,t} - \Omega_{\zeta x}, \Omega_{xx}^{-1}\Delta x_t, \qquad (53) \\ \Delta x_t &= \delta_{x1} + D_x u_{gt} + \Delta u_{xt}. \end{aligned}$$

Upon demeaning (53) and setting $E_{\zeta x} = \Omega_{\zeta x}\Omega_{xx}^{-1}$ we have

$$\tilde{y}_t = d'\tilde{x}_t + E_{\zeta x}\widetilde{\Delta x}_t + \tilde{\zeta}_{d\cdot x,t}$$

and we define the endogeneity corrected $\tilde{y}_t$ as $\tilde{y}_t^{+} = \tilde{y}_t - E_{\zeta x}\widetilde{\Delta x}_t$ so that $\tilde{y}_t^{+} = d'\tilde{x}_t + \tilde{\zeta}_{d\cdot x,t}$. Substituting consistent estimates of the long run matrices gives

$$\widehat{\tilde{y}_t^{+}} = \tilde{y}_t - \hat{E}_{\zeta x}\widetilde{\Delta x}_t = \tilde{y}_t^{+} - \left(\hat{E}_{\zeta x} - E_{\zeta x}\right)\widetilde{\Delta x}_t,$$

which we write in observation matrix form as $\widehat{\tilde{y}^{+}} = \tilde{y}^{+} - \widetilde{\Delta X}\left(\hat{E}_{\zeta x} - E_{\zeta x}\right)' = \tilde{X}d + \tilde{\zeta}_{d\cdot x,t} - \widetilde{\Delta X}\left(\hat{E}_{\zeta x} - E_{\zeta x}\right)'$. It follows that

$$\begin{aligned} \hat{d}^{+} &= \left(\tilde{X}'\tilde{X}\right)^{-1}\left(\tilde{X}'\widehat{\tilde{y}^{+}} - n\hat{\Delta}_{x\zeta}^{+}\right) = \left(\tilde{X}'\tilde{X}\right)^{-1}\left(\tilde{X}'\tilde{y}^{+} - n\hat{\Delta}_{x\zeta}^{+} - \widetilde{\Delta X}\left(\hat{E}_{\zeta x} - E_{\zeta x}\right)'\right) \\ &= d + \left(\tilde{X}'\tilde{X}\right)^{-1}\left\{\tilde{X}'\tilde{\zeta}_{d\cdot x} - n\hat{\Delta}_{x\zeta}^{+} - \tilde{X}'\widetilde{\Delta X}\left(\hat{E}_{\zeta x} - E_{\zeta x}\right)'\right\} \\ &= d + \left(\tilde{X}'\tilde{X}\right)^{-1}\left\{\tilde{X}'\tilde{\zeta}_{d\cdot x} - n\Delta_{x\zeta}^{+} + n\left(\Delta_{x\zeta}^{+} - \hat{\Delta}_{x\zeta}^{+}\right) - \tilde{X}'\widetilde{\Delta X}\left(\hat{E}_{\zeta x} - E_{\zeta x}\right)'\right\}. \qquad (54) \end{aligned}$$

The signal moment matrix $\tilde{X}'\tilde{X}$ has more complex asymptotics than usual in view of the presence of both a linear trend and stochastic trend in $\tilde{x}_t$, as evident in the definition (50) and (51). Using $D_x = [\delta_r, \delta_c]'$ in its partitioned form we have

$$\tilde{x}_t = \begin{bmatrix} \delta_r'\tilde{U}_{gt} + \tilde{u}_{xt} \\ \delta_{c1}\tilde{t} + \delta_c'\tilde{U}_{gt} + \tilde{u}_{xt} \end{bmatrix} =: \begin{bmatrix} \tilde{x}_{rt} \\ \tilde{x}_{ct} \end{bmatrix}, \qquad (55)$$

which separates the directions of (dominant) stochastic and deterministic trends and shows that these correspond exactly to those in $\tilde{x}_t$ because $\bar{R}_t$ does not have a deterministic trend and $\delta_{x1}' = (0, \delta_{c1})$. Denote $\tilde{U}_g(r) = U_g(r) - \int_0^1 U_g(s)\,ds$, $U_R(r) = \delta_r' U_g(r)$, and $\tilde{U}_R(r) = \delta_r' \tilde{U}_g(r)$. Using the partitioned notation of (55) in observation matrix form and projection matrix notation $Q_a = I - \tilde{X}_a \left(\tilde{X}_a' \tilde{X}_a\right)^{-1} \tilde{X}_a'$ for $a = \{r, c\}$, we have by standard manipulations with moment matrices of deterministic and stochastic trends (Phillips, 1988; Park and Phillips, 1988, 1989) that

$$\begin{aligned}
n^2 \left(\tilde{X}'\tilde{X}\right)^{-1} &= n^2 \begin{pmatrix} \tilde{X}_r'\tilde{X}_r & \tilde{X}_r'\tilde{X}_c \\ \tilde{X}_c'\tilde{X}_r & \tilde{X}_c'\tilde{X}_c \end{pmatrix}^{-1} \\
&= n^2 \begin{pmatrix} \left(\tilde{X}_r' Q_c \tilde{X}_r\right)^{-1} & -\left(\tilde{X}_r'\tilde{X}_r\right)^{-1} \tilde{X}_r'\tilde{X}_c \left(\tilde{X}_c' Q_r \tilde{X}_c\right)^{-1} \\ -\left(\tilde{X}_c' Q_r \tilde{X}_c\right)^{-1} \tilde{X}_c'\tilde{X}_r \left(\tilde{X}_r'\tilde{X}_r\right)^{-1} & \left(\tilde{X}_c' Q_r \tilde{X}_c\right)^{-1} \end{pmatrix} \quad (56) \\
&= \begin{pmatrix} \left(\frac{\tilde{X}_r' Q_c \tilde{X}_r}{n^2}\right)^{-1} & -\frac{1}{n^{1/2}} \left(\frac{\tilde{X}_r'\tilde{X}_r}{n^2}\right)^{-1} \frac{\tilde{X}_r'\tilde{X}_c}{n^{5/2}} \left(\frac{\tilde{X}_c' Q_r \tilde{X}_c}{n^3}\right)^{-1} \\ -\frac{1}{n^{1/2}} \left(\frac{\tilde{X}_r' Q_c \tilde{X}_r}{n^3}\right)^{-1} \frac{\tilde{X}_c'\tilde{X}_r}{n^{5/2}} \left(\frac{\tilde{X}_r'\tilde{X}_r}{n^2}\right)^{-1} & \frac{n^2}{n^3} \left(\frac{\tilde{X}_c' Q_r \tilde{X}_c}{n^3}\right)^{-1} \end{pmatrix} \\
&\Rightarrow \begin{pmatrix} \delta_r' \int_0^1 \underline{\tilde{U}}_g(r)\, \underline{\tilde{U}}_g(r)'\, dr \delta_r & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \int_0^1 \underline{\tilde{U}}_R(r)^2\, dr & 0 \\ 0 & 0 \end{pmatrix}. \quad (57)
\end{aligned}$$

Here, $\underline{\tilde{U}}_g(r) = \tilde{U}_g(r) - \left(\int_0^1 \tilde{U}_g(s)\,\tilde{r}(s)\,ds\right)\left(\int_0^1 \tilde{r}(s)^2\,ds\right)\tilde{r}(r)$ is the $L_2[0,1]$ projection residual of the functon $\tilde{U}_g(r)$ on the space spanned by $\tilde{r}(p) = p - \int_0^1 s ds$. Note also that

$$n^{-3}\tilde{X}'\tilde{X} \to_p \begin{pmatrix} 0 & 0 \\ 0 & \delta_{c1}^2 \int_0^1 \tilde{r}^2 dr \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & \frac{\delta_{c1}^2}{12} \end{pmatrix},$$

and

$$n^{-3}\tilde{X}_c' Q_r \tilde{X}_c = \frac{1}{n^3}\left(\tilde{X}_c'\tilde{X}_c\right) - \left(\frac{\tilde{X}_c'\tilde{X}_r}{n^{5/2}}\right)\left(\frac{\tilde{X}_r'\tilde{X}_r}{n^2}\right)^{-1}\left(\frac{\tilde{X}_r'\tilde{X}_c}{n^{5/2}}\right) \rightsquigarrow \delta_{c1}^2 \int_0^1 \tilde{r}_x(p)^2\,dp, \quad (58)$$

where $\tilde{r}_x(p) = \tilde{p} - \left(\int_0^1 s\tilde{U}_R(s)\,ds\right)\left(\int_0^1 \tilde{U}_R(s)^2\,ds\right)^{-1}\tilde{U}_R(p)$ is the $L_2[0,1]$ projection residual of the functon $\tilde{p}$ on the space spanned by the function $\tilde{U}_R(s)$.

Next note that

$$\begin{aligned}
(0,1)\left(\tilde{X}'\tilde{X}\right)^{-1} &= (0,1) \begin{pmatrix} \left(\tilde{X}_r' Q_c \tilde{X}_r\right)^{-1} & -\left(\tilde{X}_r'\tilde{X}_r\right)^{-1} \tilde{X}_r'\tilde{X}_c \left(\tilde{X}_c' Q_r \tilde{X}_c\right)^{-1} \\ -\left(\tilde{X}_c' Q_r \tilde{X}_c\right)^{-1} \tilde{X}_c'\tilde{X}_r \left(\tilde{X}_r'\tilde{X}_r\right)^{-1} & \left(\tilde{X}_c' Q_r \tilde{X}_c\right)^{-1} \end{pmatrix} \\
&= \left(-\left(\tilde{X}_c' Q_r \tilde{X}_c\right)^{-1} \tilde{X}_c'\tilde{X}_r \left(\tilde{X}_r'\tilde{X}_r\right)^{-1}, \left(\tilde{X}_c' Q_r \tilde{X}_c\right)^{-1}\right).
\end{aligned}$$

The scaled and centred estimator of the coefficient d_2 of $\ln(CO_{2t})$ in the cointegrating regression is

$$\begin{aligned}
& n^{3/2}\left(\hat{d}_2^+ - d_2\right) = n^{3/2}(0,1)\left(\hat{d}^+ - d\right) \\
= \; & n^{3/2}\left[-\left(\tilde{X}_c' Q_r \tilde{X}_c\right)^{-1} \tilde{X}_c' \tilde{X}_r \left(\tilde{X}_r' \tilde{X}_r\right)^{-1}, \left(\tilde{X}_c' Q_r \tilde{X}_c\right)^{-1}\right] \\
& \times \begin{bmatrix} \tilde{X}_r' \tilde{\zeta}_{d.x} - n\Delta^+_{x_r\zeta} + n\left(\Delta^+_{x_r\zeta} - \hat{\Delta}^+_{x_r\zeta}\right) - \tilde{X}_r' \widetilde{\Delta X}\left(\hat{E}_{\zeta x} - E_{\zeta x}\right)' \\ \tilde{X}_c' \tilde{\zeta}_{d.x} - n\Delta^+_{x_c\zeta} + n\left(\Delta^+_{x_c\zeta} - \hat{\Delta}^+_{x_c\zeta}\right) - \tilde{X}_c' \widetilde{\Delta X}\left(\hat{E}_{\zeta x} - E_{\zeta x}\right)' \end{bmatrix} \\
= \; & \left[-\left(\frac{\tilde{X}_c' Q_r \tilde{X}_c}{n^3}\right)^{-1} \left(\frac{\tilde{X}_c' \tilde{X}_r}{n^{5/2}}\right) \left(\frac{\tilde{X}_r' \tilde{X}_r}{n^2}\right)^{-1}, \left(\frac{\tilde{X}_c' Q_r \tilde{X}_c}{n^3}\right)^{-1}\right] \\
& \times \begin{bmatrix} \frac{1}{n}\left\{\tilde{X}_r' \tilde{\zeta}_{d.x} - n\Delta^+_{x_r\zeta} + n\left(\Delta^+_{x_r\zeta} - \hat{\Delta}^+_{x_r\zeta}\right) - \tilde{X}_r' \widetilde{\Delta X}\left(\hat{E}_{\zeta x} - E_{\zeta x}\right)'\right\} \\ \frac{1}{n^{3/2}}\left\{\tilde{X}_c' \tilde{\zeta}_{d.x} - n\Delta^+_{x_c\zeta} + n\left(\Delta^+_{x_c\zeta} - \hat{\Delta}^+_{x_c\zeta}\right) - \tilde{X}_c' \widetilde{\Delta X}\left(\hat{E}_{\zeta x} - E_{\zeta x}\right)'\right\} \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
\rightsquigarrow \; & \left[-\left(\int_0^1 \tilde{r}_x(p)^2\, dp\right)^{-1} \left(\int_0^1 \tilde{r}(p)\tilde{U}_R(p)\, dp\right) \left(\int_0^1 \tilde{U}_R(p)^2\, dp\right)^{-1}, \left(\int_0^1 \tilde{r}_x(p)^2\, dp\right)^{-1}\right] \\
& \times \begin{bmatrix} \int_0^1 \tilde{U}_R(p)\, dB_{\zeta\cdot x}(p) \\ \int_0^1 \tilde{r}(p)\, dB_{\zeta\cdot x}(p) \end{bmatrix} \\
= \; & \left(\int_0^1 \tilde{r}_x(p)^2\, dp\right)^{-1} \left(\int_0^1 \tilde{r}(p)\, dB_{\zeta\cdot x}\right) - \left(\int_0^1 \tilde{r}_x(p)^2\, dp\right)^{-1} \left(\int_0^1 \tilde{r}(p)\tilde{U}_R(p)'\, dp\right) \\
& \times \left(\int_0^1 \tilde{U}_R(p)^2\, dp\right)^{-1} \int_0^1 \tilde{U}_R(p)\, dB_{\zeta\cdot x}(p) \\
= \; & \left(\int_0^1 \tilde{r}_x(p)^2\, dp\right)^{-1} \left\{\int_0^1 \tilde{r}_x(p)\, dB_{\zeta\cdot x}\right\}, \text{ with } \tilde{r}_x(p) = \tilde{r}(p) - \left(\int_0^1 \tilde{r}\tilde{U}_R\right)\left(\int_0^1 \tilde{U}_R^2\right)^{-1} \tilde{U}_R(p) \\
= {}_d \; & \mathcal{MN}\left(0, \omega^2_{\zeta_d\cdot x}\left(\int_0^1 \tilde{r}_X(p)^2\, dp\right)\right),
\end{aligned}$$

giving the result as stated. In the above derivation the stochastic integral representation follows from standard limit theory on weak convergence to stochastic integrals (Phillips, 1988; Ibragimov and Phillips, 2008).

For coefficient d_1, we obtain in a similar manner

$$\begin{aligned} & n\left(\hat{d}_1^+ - d_1\right) = n\left(1,0\right)\left(\hat{d}^+ - d\right) \\ = \;& n\left[\left(\tilde{X}_r' Q_c \tilde{X}_r\right)^{-1}, -\left(\tilde{X}_r'\tilde{X}_r\right)^{-1}\tilde{X}_r'\tilde{X}_c\left(\tilde{X}_c' Q_r \tilde{X}_c\right)^{-1}\right] \\ & \times \begin{bmatrix} \tilde{X}_r'\tilde{\zeta}_{d.x} - n\Delta_{x_r\zeta}^+ + n\left(\Delta_{x_r\zeta}^+ - \hat{\Delta}_{x_r\zeta}^+\right) - \tilde{X}_r'\widetilde{\Delta X}\left(\hat{E}_{\zeta x} - E_{\zeta x}\right)' \\ \tilde{X}_c'\tilde{\zeta}_{d.x} - n\Delta_{x_c\zeta}^+ + n\left(\Delta_{x_c\zeta}^+ - \hat{\Delta}_{x_c\zeta}^+\right) - \tilde{X}_c'\widetilde{\Delta X}\left(\hat{E}_{\zeta x} - E_{\zeta x}\right)' \end{bmatrix} \end{aligned}$$

$$\begin{aligned} = \;& \left[\left(\frac{\tilde{X}_r' Q_c \tilde{X}_r}{n^2}\right)^{-1}, -\left(\frac{\tilde{X}_r'\tilde{X}_r}{n^2}\right)^{-1}\left(\frac{\tilde{X}_r'\tilde{X}_c}{n^{5/2}}\right)\left(\frac{\tilde{X}_c' Q_r \tilde{X}_c}{n^3}\right)^{-1}\right] \\ & \times \begin{bmatrix} \frac{1}{n}\left\{\tilde{X}_r'\tilde{\zeta}_{d.x} - n\Delta_{x_r\zeta}^+ + n\left(\Delta_{x_r\zeta}^+ - \hat{\Delta}_{x_r\zeta}^+\right) - \tilde{X}_r'\widetilde{\Delta X}\left(\hat{E}_{\zeta x} - E_{\zeta x}\right)'\right\} \\ \frac{1}{n^{3/2}}\left\{\tilde{X}_c'\tilde{\zeta}_{d.x} - n\Delta_{x_c\zeta}^+ + n\left(\Delta_{x_c\zeta}^+ - \hat{\Delta}_{x_c\zeta}^+\right) - \tilde{X}_c'\widetilde{\Delta X}\left(\hat{E}_{\zeta x} - E_{\zeta x}\right)'\right\} \end{bmatrix} \end{aligned}$$

$$\begin{aligned} \rightsquigarrow \;& \left[\left(\int_0^1 \underline{\tilde{U}}_R(r)^2\, dp\right)^{-1}, -\left(\int_0^1 \tilde{U}_R(p)^2\, dp\right)^{-1}\left(\int_0^1 \tilde{U}_R(p)\,\tilde{r}(p)\, dp\right)\left(\int_0^1 \tilde{r}(p)^2\, dp\right)^{-1}\right] \\ & \times \begin{bmatrix} \int_0^1 \tilde{U}_R(p)\, dB_{\zeta\cdot x}(p) \\ \int_0^1 \tilde{r}(p)\, dB_{\zeta\cdot x}(p) \end{bmatrix} \\ = \;& \left(\int_0^1 \underline{\tilde{U}}_R(p)^2\, dp\right)^{-1}\left\{\int_0^1 \underline{\tilde{U}}_R(p)\, dB_{\zeta\cdot x}\right\}, \text{ with } \underline{\tilde{U}}_R(p) = \tilde{U}_R(p) - \left(\int_0^1 \tilde{U}_R\tilde{r}\right)\left(\int_0^1 \tilde{r}^2\right)^{-1}\tilde{r}(p) \\ =_d \;& \mathcal{MN}\left(0, \omega_{\zeta_d\cdot x}^2\left(\int_0^1 \underline{\tilde{U}}_R(p)^2\, dp\right)^{-1}\right), \end{aligned}$$

as stated. ■

8 Appendix B: Additional Tables

Table B1: Residual Based Tests for Cointegration

	Statistic	10% Critical Value	1% Critical Value
Variables	R_t and $\ln(CO_{2,t})$		
C_ADF	-0.2279	-3.0890	-4.0245
Variables	$\bar{T}_t$ and $\ln(CO_{2,t})$		
C_ADF	-2.4897	-3.0890	-4.0245
Variables	$\bar{T}_t$, R_t and $\ln(CO_{2,t})$		
C_ADF	-5.0238***	-3.5292	-4.4969

Notes: Phillips and Ouliaris (1990) residual based augmented Dickey-Fuller (C_ADF) tests.

Table B2: Estimates of Transient Climate Sensitivity (TCS): Observational Data

TCS Parameter $\frac{\gamma_3}{1-\beta_1-\gamma_1} \times \ln(2)$ from Cointegrating Regression Equation (17)

Estimation Method	TCS	St.Error	95% Confidence Interval
OLS	2.8738	0.2422	(2.398, 3.348)
FM-OLS	2.8020	0.2253	(2.360, 3.243)
D-OLS(1,1)	2.9175	0.3587	(1.903, 3.932)
D-OLS(2,2)	2.5525	0.5507	(0.995, 4.109)
D-OLS(3,3)	3.1637	0.7506	(1.041, 5.286)
WG	2.7012	0.5666	(1.931, 3.471)

Notes: D-OLS(m,m): m leads and m lags in the DOLS regression

9 Appendix C

The empirical estimate of Transient Climate Sensitivity derived in this paper is based on land station observations (and hence land areas) only, for which we use the notation TCS_L. For comparison with previous papers it is of interest to obtain a global estimate, denoted here as TCS_G. To convert from TCS_L to TCS_G we use the conventional transformation:

$$\begin{aligned} TCS_G &= TCS_L \cdot \frac{A_L \cdot w_L + A_O \cdot w_O}{w_L} \\ &= TCS_L \cdot \left(A_L + A_O \cdot \frac{w_O}{w_L}\right) \\ &= TCS_L \cdot W_{trans} \end{aligned} \tag{59}$$

Earth's land area fraction (A_L) and ocean area fraction (A_O) are about 0.29 and 0.71, respectively. These fractions do not change over time or between observations, or among GCMs. The warming over land and ocean, w_L and w_O, do vary slightly over time and can vary quite substantially across GCMs. But the ratio $\frac{w_O}{w_L}$ is relatively similar across time and between observations and GCMs.

For the conversions in this paper we used observational land and ocean warming factors (in K) of $w_L = 1.005$, and $w_O = 0.62$. The resulting factor used to convert from TCS_L to TCS_G is

$$W_{trans} = A_L + A_O \cdot \frac{w_O}{w_L} = 0.728$$

and the observed ratio of warming over ocean to warming over land is :

$$W_{OL} = \frac{w_O}{w_L} = 0.617 \tag{60}$$

For GCMs we use the W_{OL} values reported in Table 1, and calculate the corresponding W_{trans} values. There is some uncertainty associated with the ratio in equation (60) in both observations and GCMs. The observational ratio will depend somewhat on the time period used to calculate it, and also which observational data sets are considered. For GCMs the measure will differ slightly among ensemble members of the same model and also manifest some time period dependence. We accounted for this uncertainty by adding an uncertainty bound of ± 0.05 to W_{OL}. If the ratio decreases by 0.05, the W_{trans} factor changes to

$$W^-_{trans} = A_L + A_O \cdot \frac{w_O}{w_L} \cdot (1 - 0.05) = 0.706;$$

and if it increases by 0.05, the W_{trans} factor changes to

$$W^+_{trans} = A_L + A_O \cdot \frac{w_O}{w_L} \cdot (1 + 0.05) = 0.750.$$

To obtain a confidence interval for TCS_G we simply multiply the bounds of the confidence interval (CI) for TCS_L by W_{trans}. To account for the uncertainty in W_{OL} we instead

multiply the lower bound of the CI by W^{-}_{trans} and the upper bound by W^{+}_{trans}. This adjustment leads to a slightly wider uncertainty range than the 95% confidence interval for TCS_G based on the transformation factor W_{trans} alone.

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.

Arellano, M. and S. Bond (1991). "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," *Review of Economic Studies* 58, 277–297.

Armour, K. (2017). "Energy budget constraints on climate sensitivity in light of inconstant climate feedbacks," *Nature Climate Change* 7, 331–335.

Blundell, R. W. and S. Bond (1998). "Initial conditions and moment restrictions in dynamic panel data models", Journal of Econometrics, 87, 115-143.

Breiman, L. (2001). "Random forests, " *Machine Learning*, 1, 5–32.

Eyring, V., S. Bony. G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer and K. E. Taylor (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, *Geosci. Model Dev.,* 9, 1937-1958.

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," *Bulletin of the American Meteorological Society,* 80, 831–850.

Hansen, J., L. Narazenko, R. Ruedy, M. Sato, J. Willis, A. Del Genio, D. Koch, A. Lacis, K. Lo, S. Menon, T. Novakov, J. Perlwitz, G. Russell, G. A. Schmidt, and N. Tausnev (2005). "Earth's energy imbalance: confirmation and implications," *Science,* 308, 1431–1435.

Hansen, J., M. Sato, P. Kharecha, and 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.

Hartmann, D.L., A.M.G. Klein Tank, M. Rusticucci, L.V. Alexander, S. Brönnimann, Y. Charabi, F.J. Dentener, E.J. Dlugokencky, D.R. Easterling, A. Kaplan, B.J. Soden, P.W. Thorne, M. Wild & P.M. Zhai, (2013) "Observations: Atmosphere and Surface. 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 [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. " *Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.*

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

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

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

Huang, Y. and M. Bani Shahabadi (2014). "Why logarithmic? A note on the dependence of radiative forcing on gas concentration, " *Journal of Geophysical Research: Atmospheres*, 119, 13,683â€"13.

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

Kaufmann R. K., H. Kauppi, M. L. Mann, and J. H. Stock (2011) "Reconciling anthropogenic climate change with observed temperature 1998–2008," *Proceedings of the National Academy of Sciences*, 108 (29), 11790–11793.

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

Kaufmann, R. K., H. Kauppi, and J. H. Stock (2006a) "Emissions, concentrations, and temperature: A time series analysis," *Climatic Change*, 77, 3â€"-4.

Kaufmann, R. K., H. Kauppi, and J. H. Stock (2006b) "The relationship between radiative forcing and temperature: what do statistical analyses of the instrumental temperature record measure," *Climatic Change*, 77, 249–278.

Kiehl, J. (2007) "Twentieth century climate model response and climate sensitivity," *Geophysical Research Letters*, 34, doi: 10.1029/GL2007031383.

Knutti, R., M. A. A. Rugenstein, and G. C. Hegerl (2017) "Beyond equilibrium climate sensitivity," *Nature Geoscience*, 10, 727–736.

Kostakis, A., A. Magdalinos and M. Stamatogiannis (2015). "Robust econometric inference for stock return predictability," *Review of Financial Studies*, 28(5), 1506-1553.

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., B. Melenberg and C. Muris (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. (1991). "Optimal Inference in Cointegrated Systems," *Econometrica* 59, 283–306.

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

Phillips, P. C. B. (2014). "Optimal estimation of cointegrated systems with irrelevant instruments," *Journal of Econometrics,* 178, 210-224.

Phillips, P. C. B. (2018). "Econometric Modeling of Climate Change," *Cowles Foundation Discussion Paper #2150, Yale University.*

Phillips, P. C. B. and S. N. Durlauf (1986). "Multiple Time Series Regression with Integrated Processes," *Review of Economic Studies* 53, 473–496.

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 T. Magdalinos (2007), "Limit Theory for Moderate Deviations from a Unit Root," *Journal of Econometrics,* 136, 115-130.

Phillips, P. C. B. and T. Magdalinos (2009). "Econometric Inference in the Vicinity of Unity," *Working Paper, Singapore Management University.*

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

Phillips, P. C. B. and S. Ouliaris (1990). Asymptotic Properties of Residual Based Tests for Cointegration," *Econometrica,* 58, 165-193.

Phillips, P. C. B. and V. Solo (1992). Asymptotics for Linear Processes," *Annals of Statistics,* 20, 971-1001.

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*, 9, No. 4, 286-289.

Storelvmo, T., U. K. Heede, T. Leirvik, P. C. B. Phillips, P. Arndt & M. Wild (2018). "Lethargic response to aerosol emissions in current climate models," *Geophysical Research Letters*, 45, No.15, 359-378.

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.

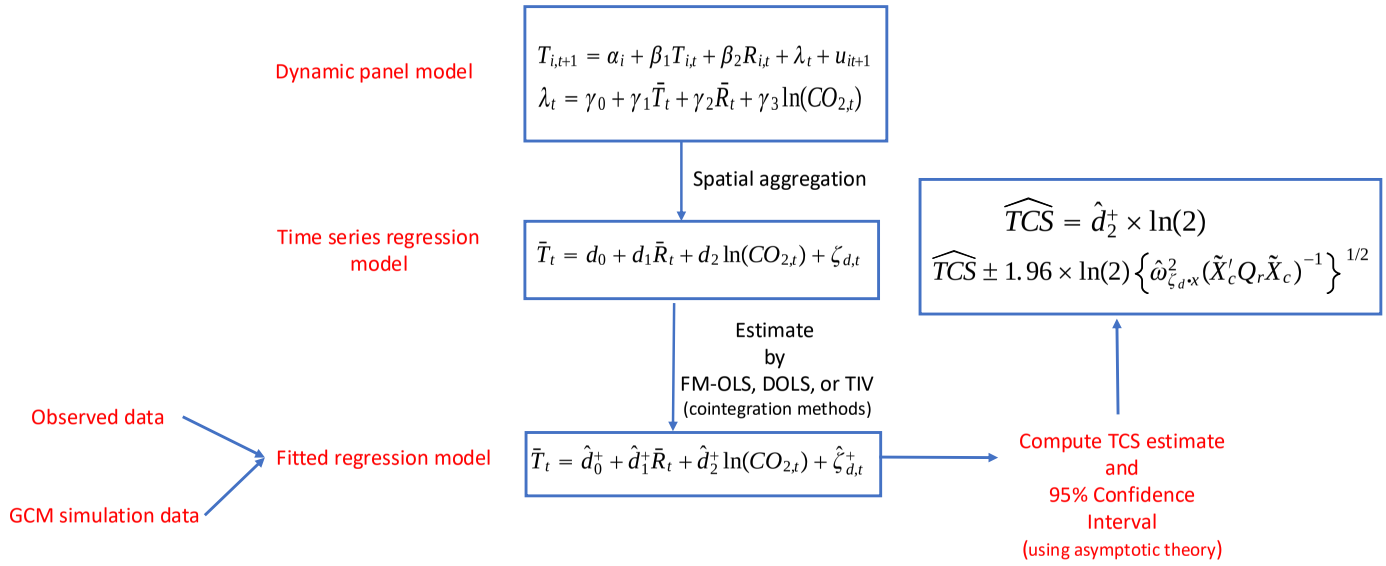

Figure 5: Roadmap for Computation: Econometric Estimation and Confidence Interval Construction