MONASH University

MONASH BUSINESS SCHOOL

DEPARTMENT OF ECONOMETRICS AND BUSINESS STATISTICS

ISSN 1440-771X

WORKING PAPER SERIES

Does Climate Sensitivity Differ Across Regions? A Varying-Coefficient Approach

Heather Anderson, Jiti Gao, Farshid Vahid, Wei Wei, and Yang Yang
Monash University, Melbourne Australia

Working Paper No. 07/23
May 2023

Does Climate Sensitivity Differ Across Regions ?

A Varying–Coefficient Approach

Heather Anderson, Jiti Gao, Farshid Vahid, Wei Wei, and Yang Yang

Monash University, Melbourne, Australia

May 4, 2023

Abstract

Global mean surface temperature has been increasing in response to growing greenhouse gas concentrations (IPCC, 2021). While Earth is getting warmer overall, regions that differ in local geographical features experience unequal increases in temperature. In this paper, we develop a dynamic varying-coefficient panel data model and use it to measure regional climate sensitivity, defined as the increase in temperature in that region, following a doubling of CO2 concentration. The inference method proposed in this paper is capable of accommodating heterogeneous co-integrating relationships between global and local variables, and it allows co-moving climate time series to possess both stochastic and deterministic trending components. Using observational data of mean surface temperatures, solar radiation, and carbon dioxide concentrations between 1959–2017, our model provides an estimate of a 3.7°C increase for average climate sensitivity. Moreover, our estimates indicate that high-latitude regions in the Northern Hemisphere are most vulnerable to global warming.

Keywords: Climate sensitivity; Dynamic panel; Varying-coefficient model; Cointegration.

JEL classification: C22, C53, Q54

1 Introduction

Earth's climate is getting warmer, and the dominant cause for this warming is anthropogenic greenhouse gas (GHG) emissions (IPCC, 2021, 2014). Measuring climate sensitivity, which refers to the magnitude of Earth's temperature response to the increase in atmospheric GHGs, is of critical importance to policy-makers who seek to mitigate future climate change with appropriate actions. Improvements in climate observations and climate models provide compelling evidence that global surface temperature during the decade spanning 2011 - 2020 was $1^{\circ}C$ higher than that observed over 1850 - 1900 (IPCC, 2021). This warming has not been homogeneous across the globe. It is well known that land surface has warmed more than the ocean surface, and that warming of the Arctic surface has been particularly rapid.(e.g. Screen and Simmonds, 2010). The purpose of this paper is to investigate the heterogeneity of climate sensitivity for land surfaces using a varying-coefficient panel data model of observational climate data.

The Earth's climate system relies on a balance of absorbed and reflected radiative forcing to sustain a steady state. Well-mixed GHGs, and in particular CO_2, absorb heat radiated from Earth's surface and act like a blanket that keeps heat from escaping. On the other hand, anthropogenic aerosol emissions, such as black carbon and sulfate from the combustion of fossil fuels, cool the Earth's climate by reducing the total solar radiation reaching the surface of the planet (Myhre et al., 2013). Aerosols contributed to a decrease in solar radiation at Earth's surface from about 1960 to 1985, but subsequent efforts to reduce the pollution associated with aerosols has led to increasing radiation since about 1990 (see, e.g., Power, 2003; Norris and Wild, 2007).

The warming effect of CO_2 and the cooling effect of aerosols work in opposite directions. Measuring CO_2-induced temperature changes thus requires accounting for the cooling effect of aerosol particles that offset the greenhouse effect by an unknown amount (e.g. Andreae et al., 2005). Magnus et al. (2011) pioneered the resolution of this problem by using panel econometric techniques and data from a large number of weather stations. Storelvmo et al. (2016) related panel model estimates to transient climate sensitivity (TCS) and their results reconciliated the different TCS values produced by observational studies versus global climate models.

The modelling of trends is particularly important when undertaking econometric analysis of climate data. Kaufmann et al. (2013) find that radiative forcing and surface temperature contain a common stochastic trend, while Gay-Garcia et al. (2009) argue that a deterministic trend with a break fits the data better. Chen et al. (2022) find a common trend between the global mean temperature and GHG concentration using a common feature approach that does not require

specifying the nature of trend. Phillips et al. (2020) provide trend specifications for local radiation and global GHG concentration that allow for both stochastic and deterministic trend, and they establish asymptotic theory to validate their estimated TCS and its confidence interval.

Although the existing econometric models are able to disentangle the temperature change due to GHG warming and aerosol cooling, they assume a constant temperature response to the change in CO_2 across regions. Climate research suggests that climate sensitivity depends on features such as clouds and snow cover which may be approximated by geographical variables such as latitude (see e.g. Boer and Yu, 2003; Bony et al., 2006). In this paper, we propose a novel varying-coefficient dynamic panel model which allows the relationship between temperature, radiation, and CO_2 to depend on time invariant variables. Our model yields station-level estimates of climate sensitivity that contribute towards an understanding of the regional heterogeneity of global warming.

We contribute to the panel data econometrics literature by providing an estimation method and associated asymptotic theory for our proposed model, which has interesting new features that can not be handled with existing methods. Our model follows Phillips et al. (2020) in that it allows for both deterministic and stochastic trends as well as cointegration, but it also allows for varying coefficients. Existing varying-coefficient models often assume strictly stationary regressors (see for example, Cai and Li (2008), Connor et al. (2012) and Feng et al. (2017)). Models that do allow varying-coefficients and non-stationary covariates are often cast in a time series setting without the cross-sectional dimension (see for example, Cai et al. (2009) and Sun et al. (2016)). Furthermore, our panel model is dynamic in that lagged dependent variables are also included as regressors. Dong et al. (2021) propose a varying-coefficient panel data model with non-stationary regressors, but they do not include lagged dependent variables as regressors.

The structure of this article is as follows. Section 2 introduces the varying-coefficient econometric framework for modeling climate observations collected over time at various spatial locations. Section 3 provides the estimation procedure and the asymptotic properties of the proposed estimator. In Section 4, we apply the proposed methodology to observational climate data related to those employed in Storelvmo et al. (2016) and Phillips et al. (2020). Using the estimated coefficient functions, we compute climate sensitivity at global and local levels. A summary and conclusions are presented in Section 5. Appendices A and B include some additional empirical analyses, Appendix C provides simulation results and Appendix D covers the mathematical derivations and proofs of the main results established in Section 3.

We use the following notation throughout the paper. The symbol "$\Rightarrow$" denotes weak convergence, "$\xrightarrow{a.s.}$" is almost sure convergence, and ":=" is definitional equivalence. For any matrix (function) A, we use A' and A'' to denote the first and second derivative, and $A^\top$ to denote the transpose. For a stochastic process $\mathcal{U}(r)$, integrals such as $\int_0^1 \mathcal{U}(r)dr$ are written as $\int \mathcal{U}$ and $\int_0^1 \mathcal{U}(r)d\mathcal{U}(r)$ as $\int \mathcal{U}d\mathcal{U}$. We use $\mathbb{E}_N$ to denote cross-sectional expectation.

2 A varying-coefficient climate econometric model

A standard panel model that relates temperature to solar radiation and CO_2 concentration, both at the individual-station level and at the global aggregate level, takes the following form:

$$\begin{aligned} T_{i,t+1} = \alpha_i + \beta_1 T_{i,t} + \beta_2 R_{i,t} + \gamma_1 \overline{T}_t + \gamma_2 \overline{R}_t \\ + \gamma_3 \ln(CO2_t) + \epsilon_{i,t+1}, \quad i = 1, \ldots, N \text{ and } t = 1, \ldots, n, \end{aligned} \tag{1}$$

where $T_{i,t}$ denotes the surface air temperature at a land-based station i in year t, R_{it} denotes the local downward shortwave radiation, and $\ln(CO2_t)$ denotes the logarithm of global atmosphere concentration of CO_2. Spatial averages are defined as $\overline{T}_t = \frac{1}{N}\sum_{i=1}^N T_{i,t}$ and $\overline{R}_t = \frac{1}{N}\sum_{i=1}^N R_{i,t}$. The error term $\epsilon_{i,t+1}$ captures idiosyncratic uncertainty. This model was proposed in Magnus et al. (2011) and used by Storelvmo et al. (2016) and Phillips et al. (2020) among others.

The econometric model in (1) is developed from an energy balance equation

$$c(T_{i,t+1} - T_{i,t}) = E_{i,t}^{in} - E_{i,t}^{out} + E_{i,t}^{exch}, \tag{2}$$

where c is the heat capacity of the Earth's surface, and $E_{i,t}^{in}$, $E_{i,t}^{out}$, and $E_{i,t}^{exch}$ denote the incoming, outflow, and net exchange of energy at location i, respectively.

The incoming energy, $E_{i,t}^{in}$, measures solar radiation that has reached and has been absorbed by Earth's surface, and this is modeled by

$$E_{i,t}^{in} = a_0 + a_1 \overline{R}_t + a_2(R_{i,t} - \overline{R}_t), \tag{3}$$

where $a_1 \geq a_2 \geq 0$, so that both the global and local effects of solar radiation (i.e. $a_1\overline{R}_t$ and $a_2(R_t - \overline{R}_t)$) increase energy arrival at station i.

The outflow of energy, $E_{i,t}^{out}$, measures the longwave radiation emitted from the atmosphere

to outer space, and this is modeled by

$$E_{i,t}^{out} = b_0 + b_1 \overline{T}_t + b_2 (T_{i,t} - \overline{T}_t) - b_3 \ln(CO2_t). \tag{4}$$

Econometric models for climate data generally assume that CO_2 concentration is the same for all locations in any particular year, since atmospheric CO_2 has limited spatial variability, (see, e.g., Magnus et al., 2011; Phillips et al., 2020). Given that an increase of global CO_2 concentration decreases the outflow $b_3 \geq 0$, but outflow increases with temperature which has both global and local positive effects (with $b_1 \geq b_2 \geq 0$).

The net exchange of energy, $E_{i,t}^{exch}$, depends on the difference between $T_{i,t}$ and $\overline{T}_t$ such that

$$E_{i,t}^{exch} = c_0 - c_1 (T_{i,t} - \overline{T}_t) \tag{5}$$

where $c_1 \geq 0$ reflects an outflow of energy from station i when $T_{i,t} > \overline{T}_t$.

Combining (2) (3) (4) and (5), and adding a station specific effect α_i as well as an idiosyncratic error term ϵ_{it}, leads to the model in (1). The β and γ parameters summarize the information in the latent structural parameters in (3) (4) and (5), but these latter parameters are not identifiable.

The energy equations (2) (3) (4) and (5) assume a constant relationship between the key climate variables across different locations. However, it is well known that these relationships vary according to geographical features. For example, water vapor acts like a greenhouse gas, but unlike CO_2, it is not evenly spread across the globe. When $T_{i,t}/\overline{T}_t$ increases, water vapor increases locally and this counters the effect of $T_{i,t}/\overline{T}_t$ on $E_{i,t}^{out}$ (see e.g. Bony et al., 2006; Pithan and Mauritsen, 2014). Hence, a location with greater potential for an increase in water vapor following a rise in temperature will have lower energy outflow than a location with little potential for water vapor. This implies that b_1/b_2 depends on a location's water reservoir, which can be approximated by its distance to coastline. Similarly, clouds have a strong impact on how $E_{i,t}^{out}$ responds to temperature, and large-scale atmospheric conditions can have different effects on temperatures at different latitudes (see e.g. Bony et al., 2006). In consequence, b_1/b_2 may also be a function of latitude.

Another important location-specific feature is surface albedo, which measures the amount of shortwave radiation reflected by a surface. Snow and ice have a large impact on land surface albedo (Henderson-Sellers and Wilson, 1983), and hence mid to high latitude areas in the northern hemisphere have different albedo than other regions. Albedo affects both (3) and (4). Since

$R_{i,t}/\overline{R}_t$ measures downward radiation which includes both absorbed and reflected radiation, the relationship between absorbed energy ($E^{in}_{i,t}$) and $R_{i,t}/\overline{R}_t$ depends on surface albedo. Furthermore, melting snow and ice in response to increased temperature decreases surface reflection and amplifies the rise in temperature (Screen and Simmonds, 2010). In consequence, a_1/a_2 and b_1/b_2 may depend on surface albedo which in turn depends on latitude.

The heat capacity in (2) varies across different types of Earth's surfaces. Land has lower heat capacity than ocean, and surface temperature increases are larger over land than over the ocean (IPCC, 2021). While our study only focuses on land surfaces, a region's distance to the coast may contribute to its local heat capacity, and hence affect the relationship between temperature, radiation and GHGs in this region.

We propose a varying-coefficient panel model to allow the relationship between radiation, temperature and CO_2 to vary with different location-specific features. Our model is

$$\begin{aligned} T_{i,t+1} = \alpha_i + \beta_1(z_i)T_{i,t} + \beta_2(z_i)R_{i,t} + \gamma_1(z_i)\overline{T}_t + \gamma_2(z_i)\overline{R}_t \\ + \gamma_3(z_i)\ln(CO2_t) + \epsilon_{i,t+1}, \quad i = 1,\ldots,N \text{ and } t = 1,\ldots,n, \end{aligned} \tag{6}$$

where $\{\beta_1(Z), \beta_2(Z), \gamma_1(Z), \gamma_2(Z), \gamma_3(Z)\}$ are unspecified twice continuously differentiable functions of a time invariant variable Z that has a support $\mathbb{Z}$ and a strictly positive and absolutely continuous density $f(z)$. We assume that

$$|\beta_1(z)| < 1, \text{for any } z \in \mathbb{Z}, \tag{7}$$

$$\int \frac{\gamma_1(z)}{1-\beta_1(z)} f(z)dz < 1. \tag{8}$$

Conditions (7) and (8) ensure that the local and global temperatures will eventually reach a steady state if R_{it} and $\ln(CO2_t)$ are held constant.

For simplicity we consider $\mathbb{Z} \subset \mathbb{R}^1$ but our methodology can be extended to higher dimensions. In our application, we use Z to describe geographical features of a location. For example, $Z = z_i$ could be the latitude of station i or its distance from the nearest coastline. The heterogeneity in the $\beta(Z)$ and $\gamma(Z)$ functions comes from the heterogeneity in the latent structural parameters such as a_1/a_2, b_1/b_2 and c, even though these latter parameters are not identifiable. The location specific fixed effect term α_i summarizes the effects of unobserved characteristics that are not captured by Z.

2.1 Global and local climate sensitivities

Following Storelvmo et al. (2016) and Phillips et al. (2020), we define climate sensitivity as the change in the steady state mean land surface temperature following a doubling of CO2. In our varying-coefficient model, different stations may reach steady states at different times. Suppose that all stations have reached their steady states at t_1, and let T^e_{i,t_1} denote the equilibrium/steady state temperature for station i. Then $T^e_{i,t_1} = T^e_{i,t_1+1}$ for all i leads to $N^{-1}\sum_{i=1}^{N} T^e_{i,t1} := \overline{T}^e_{t_1} = \overline{T}^e_{t_1+1}$. In other words, the global steady state is also reached at t_1. Assume that the global and local steady states, denoted by $\overline{T}^e_{t_2}$ and T^e_{i,t_2}, are reached at t_2 following a doubling of CO2. We show in Appendix A that global climate sensitivity (GCS), defined as $\overline{T}^e_{t_2} - \overline{T}^e_{t_1}$, is given by

$$GCS_N = \frac{\overline{\theta}_{3,N}}{1 - \overline{\theta}_{1,N}} \ln(2), \tag{9}$$

where

$$\overline{\theta}_{1,N} = \frac{1}{N}\sum_{i=1}^{N} \frac{\gamma_1(z_i)}{1 - \beta_1(z_i)} \text{ and}$$
$$\overline{\theta}_{3,N} = \frac{1}{N}\sum_{i=1}^{N} \frac{\gamma_3(z_i)}{1 - \beta_1(z_i)}.$$

For a station with characteristics $Z = z$, its local climate sensitivity (LCS) is given by

$$LCS_N(z) = \frac{\gamma_1(z)}{1 - \beta_1(z)} GCS_N + \frac{\gamma_3(z)}{1 - \beta_1(z)} \ln(2), \tag{10}$$

and we define GCS and LCS as the limiting values of GCS_N and LCS_N as $N \to \infty$. In contrast to the constant coefficient model in equation (1) where local climate sensitivity is constant across stations, LCS in our model varies with z.

2.2 General model and comparison to other varying-coefficient models

The model in (6) is a novel varying-coefficient dynamic panel model with a factor structure. A generalized version of (6) takes the form

$$y_{i,t} = \alpha_i + \beta^\top(z_i) x_{i,t} + \gamma^\top(z_i) f_t + \varepsilon_{i,t}, \tag{11}$$

where α_i is a fixed effect that is allowed to be arbitrarily correlated with the covariates $x_{i,t}$, f_t represents non-stationary common factors that are assumed to be fully observed, and z_i is exogenous with $\mathbb{E}(\varepsilon_{it}|z_i) = 0$. In this paper, we allow $x_{i,t}$ to be non-stationary and it may include lagged $y_{i,t}$. Furthermore, $x_{i,t}$ may be co-integrated or share the same deterministic trend with f_t.

Model (11) extends several existing varying-coefficient models in the literature and we briefly review some of them here. Cai and Li (2008) introduced the first varying-coefficient dynamic panel model together with an associated nonparametric generalized method of moments estimation procedure. Later, Connor et al. (2012) proposed a varying-coefficient factor model for asset pricing along with an efficient estimation procedure. Feng et al. (2017) included fixed effects in a varying-coefficient model with discrete covariates. All three of these studies assumed that the covariates were strictly stationary for each i. Meanwhile, Cai et al. (2009) and Sun et al. (2016) developed asymptotic theories for varying-coefficient models with non-stationary covariates in time series settings. More recently, Dong et al. (2021) have proposed a varying-coefficient panel data model with non-stationary $x_{i,t}$ and partially observed f_t, but their covariates are strictly exogenous and therefore do not allow lagged y_{it}. Further, they do not consider co-integration between x_{it} and f_t.

New and important features of (11) relative to the existing literature are that regressors can include lagged dependent variables, and they may have common non-stationary trends. Given these features, asymptotic theories for existing models do not directly apply. We present a kernel-based estimator and develop its asymptotic distribution in the next section.

3 Estimation and asymptotic theory

3.1 Estimation of the semi-parametric model

Suppose that we have a random sample $\{(z_i, T_{i,t}, R_{i,t}) : i = 1, \ldots, N; t = 1, \ldots, n\}$ and $\{CO2_t\}_{t=1}^{n}$. We first eliminate the station specific fixed effects by subtracting the time average of Eq. (6). For any variable A_{it}, we use the following notation:

$$\widetilde{A}_{i,t} = A_{i,t} - \frac{1}{n-1}\sum_{t=1}^{n-1} A_{i,t}, \quad \widetilde{\overline{A}}_t = \overline{A}_t - \frac{1}{(n-1)}\sum_{t=1}^{n-1} \overline{A}_t,$$

where $\overline{A}_t = \frac{1}{N}\sum_{i=1}^{N} A_{i,t}$. Similarly,

$$\widetilde{\ln(CO2_t)} = \ln(CO2_t) - \frac{1}{n-1}\sum_{t=1}^{n-1}\ln(CO2_t).$$

Let $\widetilde{\boldsymbol{X}}_{it} = \left(\widetilde{T}_{it}, \widetilde{R}_{it}, \widetilde{\overline{T}}_t, \widetilde{\overline{R}}_t, \widetilde{\ln(CO2_t)}\right)^\top$ and $\boldsymbol{\beta}(z_i) = (\beta_1(z_i), \beta_2(z_i), \gamma_1(z_i), \gamma_2(z_i), \gamma_3(z_i))^\top$. By subtracting the time average of Eq. (6), we have

$$\widetilde{T}_{i,t+1} = \widetilde{\boldsymbol{X}}_{it}^\top \boldsymbol{\beta}(z_i) + \widetilde{\epsilon}_{i,t+1}. \tag{12}$$

Varying-coefficient models are often estimated using a local constant or local linear method (see e.g. Fan and Zhang, 1999, among others). Approximating $\boldsymbol{\beta}(z_i)$ by a constant in the neighbourhood of z leads to the following minimization problem:

$$\min_{\boldsymbol{a}} \sum_{i=1}^{N}\sum_{t=1}^{n-1}\left(\widetilde{T}_{i,t+1} - \widetilde{\boldsymbol{X}}_{i,t}^\top \boldsymbol{a}\right)^2 K\left(\frac{z_i - z}{h}\right), \tag{13}$$

where $K(z)$ is a kernel function and h is a bandwidth parameter. The solution to Eq. (13) yields the local constant (Nadaraya-Watson) estimator:

$$\widehat{\boldsymbol{\beta}}(z) = \left(\sum_{i=1}^{N}\sum_{t=1}^{n-1}\widetilde{\boldsymbol{X}}_{i,t}\widetilde{\boldsymbol{X}}_{i,t}^\top K\left(\frac{z_i - z}{h}\right)\right)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{n-1}\widetilde{\boldsymbol{X}}_{i,t}\widetilde{T}_{i,t+1}K\left(\frac{z_i - z}{h}\right). \tag{14}$$

We consider both the local linear and the local constant estimators in Appendix B, and assess their performances using simulations. We find that they deliver similar squared errors for estimated coefficient functions given an appropriately chosen bandwidth. Hence, we focus on the asymptotic properties of the local constant estimator.

3.2 Model assumptions

The covariates in model (6) include two exogenous climate variables, R_{it} and $\ln(CO2_t)$. To derive the asymptotic distribution of the kernel estimator in Eq. (14), we need to characterize the properties of these variables. Empirical studies suggest that both $\ln(CO2_t)$ and R_{it} are non-stationary. We specify the type of trends in these variables in Assumption 1.

Assumption 1. The local radiation effects R_{it} and the logarithm of global CO_2 concentration

$\ln(CO2_t)$ are non-stationary time series generated by the following framework:

$$R_{i,t} = R^0_{i,t} + \delta_R^\top(z_i)\mathcal{U}_{g,t} + \mathcal{U}_{P,i,t},$$
$$\ln(CO2_t) = \delta_{c0} + \delta_{c1}t + \delta_c^\top \mathcal{U}_{g,t} + \mathrm{u}_{c,t}.$$

(i) The vector process $\mathcal{U}_{g,t}$ represents the stochastically non-stationary global trend in $\ln(CO2_t)$ and in $R_{i,t}$. It is a partial sum process defined by $\mathcal{U}_{g,t} = \sum_{s=1}^t \mathrm{u}_{g,s}$, where $\mathrm{u}_{g,s}$ is a vector linear process such that $\mathcal{U}_{g,t}$ satisfies the invariance principle. That is, $n^{-1/2}\mathcal{U}_{g,\lfloor nr \rfloor} \Rightarrow \mathcal{U}_g(r) \equiv BM(\Sigma_g)$, where $BM(\Sigma_g)$ denotes a vector Brownian motion with covariance matrix $\Sigma_g > 0$.

(ii) $\mathcal{U}_{P,i,t}$ represents a local non-stationary component that is independent from $\mathcal{U}_{g,t}$. It is a partial sum process defined by $\mathcal{U}_{P,i,t} = \sum_{s=1}^t \mathrm{u}_{P,i,s}$, where $\mathrm{u}_{P,i,s}$ is a linear process such that $\mathcal{U}_{P,i,t}$ satisfies the invariance principle: $n^{-1/2}\mathcal{U}_{P,i,\lfloor nr \rfloor} \Rightarrow \mathcal{U}_{P,i}(r) \equiv BM(\sigma^2_{i,P})$ for all i , and $BM(\sigma^2_{i,P})$ denotes a set of independent scalar Brownian motions with variances $\sigma^2_{i,P} > 0$.

(iii) $R^0_{i,t}$ represents the stationary component of the local radiation level.

(iv) The partial sum process $\mathcal{U}_{c,t} = \sum_{s=1}^t \mathrm{u}_{c,s}$ satisfies the invariance principle: $n^{-1/2}\mathcal{U}_{c,\lfloor nr \rfloor} \Rightarrow \mathcal{U}_c(r) \equiv BM(\sigma^2_c)$, with $BM(\sigma^2_c)$ denoting a scalar Brownian motion with variance $\sigma^2_c > 0$.

(v) The loading vector $\delta_R(z)$ is a twice continuously differentiable function over the support of Z.

Remark 1. Assumption 1 inherits the framework in Phillips et al. (2020) with the exception that we assume that the factor loading δ_R is a function of Z.

Phillips et al. (2020) derive conditions under which $T_{i,t}$, $\overline{T}_t$ and $\overline{R}_t$ in the constant coefficient model (1) are driven by the same trends as those in R_{it} and $\ln(CO2_t)$, should Assumption 1 hold. However, deriving similar conditions under which $T_{i,t}$, $\overline{T}_t$ and $\overline{R}_t$ in our varying-coefficient model (6) are driven by the same trends as those in R_{it} and $\ln(CO2_t)$ (given Assumption 1), requires the imposition of substantial conditions. To stay with the main ideas of this paper and avoid including additional technicalities, we impose the following assumption, and then briefly discuss some necessary conditions for this assumption to hold in Appendix C.1

Assumption 2. The station temperature $T_{i,t}$, the global average temperature $\overline{T}_t$ and global radiation $\overline{R}_t$ can be represented by

$$T_{i,t} = \delta_{T0,i} + \delta_{T1}(z_i)t + \delta_T^\top(z_i)\mathcal{U}_{g,t} + \delta_{TP}^\top(z_i)\mathcal{U}_{P,i,t} + \eta_{T,i,t},$$

$$\overline{T}_t = \overline{\delta}_{T0} + \overline{\delta}_{T1} t + \overline{\delta}_T^\top \mathcal{U}_{g,t} + \eta_{T,t},$$
$$\overline{R}_t = \overline{\delta}_{R0} + \overline{\delta}_R^\top \mathcal{U}_{g,t} + \eta_{R,t},$$

where $\delta_{T1}(Z)$, $\delta_T(Z)$, and $\delta_{Tp}(Z)$ are transformations of $\beta_1(Z), \beta_2(Z), \gamma_1(Z), \gamma_2(Z), \gamma_3(Z), \delta_R(Z)$ and the parameters in Assumption 1, and they are twice continuously differentiable functions over the support of Z. The parameters $\delta_{T0,i}$, $\overline{\delta}_{T0}$, $\overline{\delta}_{T1}$, $\overline{\delta}_T$, $\overline{\delta}_{r0}$, $\overline{\delta}_R$ are finite real scalars/vectors. The error process $\eta_{T,i,t}$ is stationary over t and independent across i, while $\eta_{r,t}$ and $\{\eta_{T,t}\}$ are stationary.

Remark 2. Assumption 2 ensures that $T_{i,t}$, $\overline{T}_t$, and $\overline{R}_t$ are driven by the same trends as $R_{i,t}$ and $\ln(CO2_t)$ and that the factor loadings and trends are multiplicatively separable.

Assumption 3. Further conditions for processes in the panel model (6) are

(i) The panel regression errors $\{\epsilon_{i,t}\} \overset{i.i.d.}{\sim} (0, \sigma_\epsilon^2)$ over i, and for each $i = 1, \dots N$, the errors $\{\epsilon_{i,t}\}$ are independent of $\{\mathsf{u}_{P,is}\}$, $\{\mathsf{u}_{c,s}\}$ and $\{\mathsf{u}_{g,s}\}$ for all t and s. Further,

$$\left(\frac{1}{\sqrt{n}} \sum_{s=1}^{[nr]} \epsilon_{i,s}, \frac{1}{\sqrt{n}} \sum_{s=1}^{[nr]} u_{p,i,s}, \frac{1}{\sqrt{n}} \sum_{s=1}^{[nr]} u_{g,s} \right) \Rightarrow \left(B_{\epsilon,i}(r), \mathcal{U}_{P,i}(r), \mathcal{U}_g(r) \right)$$

as $n \to \infty$ for each i, where $B_{\epsilon,i}(r) \equiv BM(\sigma_\epsilon^2)$ is independent across i.

(ii) The variable Z is independent of $\{\mathsf{u}_{P,i,t}\}$, $\{\mathsf{u}_{c,t}\}$, $\{\mathsf{u}_{g,t}\}$ and $\epsilon_{i,t}$ for all t. Further, the realisations of z_i, for $i = 1, \dots N$, are independently sampled from Z.

Assumption 4. Regularity conditions for establishing the asymptotic properties of $\widehat{\boldsymbol{\beta}}(z)$ are:

(i) $\boldsymbol{\beta}(z) = (\beta_1(z), \beta_2(z), \gamma_1(z), \gamma_2(z), \gamma_3(z))^\top$ is twice continuously differentiable for any $z \in \mathbb{Z}$. The density function $f(z)$ is strictly positive and absolutely continuous for any $z \in \mathbb{Z}$.

(ii) The kernel function $K(\cdot)$ is symmetric and bounded, satisfies the Lipschitz condition, and has a compact support, say $[-1, 1]$. Furthermore, $\int K(v)dv = 1$.

(iii) The bandwidth h satisfies $h \to 0$ and $Nh \to \infty$ as $N \to \infty$.

Remark 3.

(i) Assumption 4(i) is a typical smoothness constraint for varying-coefficient models.

(ii) Assumption 4(ii) can be easily satisfied by the commonly used kernel function such as the Epanechnikov kernel $K(z) = 0.75(1-z^2)\mathbb{1}(|z| \leq 1)$, where $\mathbb{1}(\cdot)$ is the indicator function.

(iii) Assumption 4(iii) is commonly used in nonparametric estimation (see, e.g., Li and Racine, 2007). The condition $Nh \to \infty$ ensures that asymptotically, infinitely many observations fall in the neighborhood of z with width h. As the overall sample size increases ($N \to \infty$), the interval width must shrink to zero in the limit ($h \to 0$) to retain a "local sample".

3.3 Asymptotic properties

Under Assumptions 1 and 2, the demeaned variables in $\widetilde{\boldsymbol{X}}_{i,t}$ are driven by the demeaned deterministic trend $\widetilde{t}$, the demeaned global stochastic trend $\widetilde{\mathcal{U}}_{g,t}$, and the demeaned local stochastic trend $\widetilde{\mathcal{U}}_{P,i,t}$ so that

$$\widetilde{\boldsymbol{X}}_{i,t} = \begin{pmatrix} \widetilde{T}_{i,t} \\ \widetilde{R}_{i,t} \\ \widetilde{\overline{T}}_t \\ \widetilde{\overline{R}}_t \\ \widetilde{\ln(CO2_t)} \end{pmatrix} = \begin{pmatrix} \delta_{T1}(z_i)\widetilde{t} + \delta_T^\top(z_i)\widetilde{\mathcal{U}}_{g,t} + \delta_{xp}(z_i)\widetilde{\mathcal{U}}_{P,i,t} + \widetilde{\eta}_{T,i,t} \\ \delta_R^\top(z_i)\widetilde{\mathcal{U}}_{g,t} + \widetilde{\mathcal{U}}_{P,i,t} + \widetilde{\eta}_{R,i,t} \\ \overline{\delta}_{T1}\widetilde{t} + \overline{\delta}_T^\top\widetilde{\mathcal{U}}_{g,t} + \widetilde{\eta}_{T,t} \\ \overline{\delta}_R^\top\widetilde{\mathcal{U}}_{g,t} + \widetilde{\eta}_{R,t} \\ \delta_{c1}\widetilde{t} + \delta_c^\top\widetilde{\mathcal{U}}_{g,t} + \widetilde{\eta}_{c,t} \end{pmatrix}$$
$$=: \delta_{x1}(z_i)\widetilde{t} + \delta_x(z_i)\widetilde{\mathcal{U}}_{g,t} + \delta_{xp}(z_i)\widetilde{\mathcal{U}}_{P,i,t} + \widetilde{\eta}_{x,i,t}. \tag{15}$$

For any Brownian motion $B(r)$, the demeaned process $\widetilde{B}(r) = B(r) - \int_0^1 B(s)ds$ and the demeaned partial sum converges to the corresponding demeaned Brownian motion as $n \to \infty$. For example, $n^{-1/2}\widetilde{\mathcal{U}}_{g,\lfloor nr \rfloor} \Rightarrow \widetilde{\mathcal{U}}_g(r)$. Similarly, $\widetilde{t}/n \to \tilde{r} = r - \int_0^1 r dr$.

The matrix $\sum_{t=1}^{n-1} \widetilde{X}_{i,t}\widetilde{X}_{i,t}^\top$ in equation (14) is asymptotically singular because of the common deterministic trend in $\widetilde{X}_{i,t}$. We use rotation to handle this singularity (see e.g. Park and Phillips, 1989; Phillips, 2020). Let $\widetilde{F}_t = \left(\widetilde{t}, \widetilde{U}_{g,t}^\top\right)^\top$, and rewrite (15) as

$$\widetilde{X}_{i,t} = \delta(z_i)\widetilde{F}_t + \delta_{xp}(z_i)\widetilde{\mathcal{U}}_{P,i,t} + \widetilde{\eta}_{x,i,t},$$

where $\delta(z_i) = \left(\delta_{x1}(z_i), \delta_x^\top(z_i)\right)$. As $\widetilde{\mathcal{U}}_{P,i,t}$ contains cross-sectionally independent Brownian motions, we need only to rotate the coordinates of $\delta(z_i)\widetilde{F}_t$. Specifically, for $z \in \mathbb{Z}$, we need a rotation matrix $H(z) \in \mathbb{R}^{5\times 5}$ such that

$$H(z)H(z)^\top = I_{5\times 5}, \tag{16}$$

and $H(z)^\top \delta(z) \widetilde{F}_t$ separates out the deterministic and the stochastic trend[1]. The choice of $H(z)$ is not unique. Suppose that we choose the direction of the deterministic trend as the first component of $H(z)$, i.e., we set $H(z) = (H_1(z), H_2(z))$ where

$$H_1(z) = \delta_{x1}(z) / \left(\delta_{x1}(z)^\top \delta_{x1}(z)\right)^{1/2}, \tag{17}$$

then $H_2(z) \in \mathbb{R}^{5\times 4}$ should satisfy

$$H_2(z)^\top \delta_{x1}(z) = 0_{4\times 1}. \tag{18}$$

A suitable $H_2(z)$ can be found using Eq. (18) and the orthonormal conditions. We then define

$$\begin{pmatrix} C_1(z), & C_2(z) \\ 0_{4\times 1}, & C_3(z) \end{pmatrix} := \begin{pmatrix} H_1(z)^\top \delta_{x1}(z), & H_1(z)^\top \delta_x^\top(z) \\ 0_{4\times 1}, & H_2(z)^\top \delta_x^\top(z) \end{pmatrix},$$

and

$$\begin{pmatrix} C_4(z) \\ C_5(z) \end{pmatrix} := \begin{pmatrix} H_1(z)^\top \delta_{xp}(z) \\ H_2(z)^\top \delta_{xp}(z) \end{pmatrix}.$$

Next, we use sequential limits to establish the asymptotic distribution of the local constant estimator. Specifically, we obtain intermediate limits by fixing N and letting $n \to \infty$, and then apply $N \to \infty$ to the intermediate limits. We denote this by $(n, N)_{seq} \to \infty$.

Lemma 1. Define $D_n = diag(n^{3/2}, nI_4)$ and suppose that $H(z)$ satisfies Eqs. (16), (17) and (18). Under Assumptions 1, 2, and 4, for any $z \in \mathbb{Z}$, as $(n, N)_{seq} \to \infty$,

$$\frac{1}{Nh} \sum_{i=1}^{N} \sum_{t=1}^{n} K\left(\frac{z_i - z}{h}\right) D_n^{-1} H(z)^\top \widetilde{X}_{i,t} \widetilde{X}_{i,t}^\top H(z) D_n^{-1} \Rightarrow f(z) S_1(z),$$

where

$$S_1(z) = \begin{pmatrix} C_1^2(z) \int_0^1 \widetilde{r}^2, & C_1(z) \int_0^1 \widetilde{r} \widetilde{U}_g^\top C_3(z)^\top \\ C_3(z) \left(\int_0^1 \widetilde{r} \widetilde{U}_g\right) C_1(z), & C_3(z) \left(\int_0^1 \widetilde{U}_g \widetilde{U}_g^\top\right) C_3(z)^\top + \frac{1}{6} C_5(z) C_5(z)^\top \sigma_P^2 \end{pmatrix}. \tag{19}$$

Proof. See Appendix C. □

[1] If the global stochastic trend $\widetilde{U}_{g,t}$ is one dimensional, $H(z)$ should separate the deterministic, the stochastic trend and the stationary component. We only present the results with one global stochastic trend for notational convenience.

Theorem 1. Suppose that $H(z)$ satisfies Eqs. (16) and (18). Under Assumptions 1, 2, and 4, for any given $z \in \mathbb{Z}$, as $(n, N)_{seq} \to \infty$,

$$\sqrt{Nh} D_n H(z)^\top \left(\widehat{\boldsymbol{\beta}}(z) - \boldsymbol{\beta}(z) - \kappa_2 h^2 B_\beta(z) \right) \Rightarrow MN \left(\mathbf{0}, f^{-1}(z) \nu_0 S_1^{-1}(z) \right),$$

where $B_\beta(z) = f^{-1}(z)\boldsymbol{\beta}'(z)f'(z) + \boldsymbol{\beta}''(z) + D_n^{-1} H(z)^\top S_1^{-1}(z) S_2(z) D_n H(z)^\top \beta'(z)$, $\kappa_2 = \int v^2 K(v) dv$, $\nu_0 = \int K^2(v) dv$, $S_1(z)$ is given in Eq. (19), $S_2(z)$ is relegated to the appendix, and MN denotes the mixed normal distribution.

Proof. See Appendix C. □

Proposition 1. Let $V \coloneqq f^{-1}(z) v_0 S_1^{-1}(z)$ and

$$V = \begin{pmatrix} V^{(1,1)}(z) & V^{(1,2)}(z) \\ & {\scriptstyle 4\times 1} \\ V^{(2,1)}(z) & V^{(2,2)}(z) \\ {\scriptstyle 1\times 4} & {\scriptstyle 4\times 4} \end{pmatrix}.$$

Suppose that the asymptotic distribution of the rotated $\widehat{\boldsymbol{\beta}}(z)$ is given by Theorem 1, Then

$$n\sqrt{Nh} \left(\widehat{\boldsymbol{\beta}}(z) - \boldsymbol{\beta}(z) - \kappa_2 h^2 B_\beta(z) \right) \Rightarrow MN \left(0, H_2(z) V^{(2,2)}(z) H_2(z)^\top \right).$$

Proof. See Appendix C. □

The asymptotic distributions established in Theorem 1 and Proposition 1 may not be useful for calculating confidence intervals in practice. This is due to the presence of stochastic covariance matrices and unknown rotation matrices in their construction. Therefore, we use a bootstrap method for such purposes, as described in Section 4.2.

4 Empirical analysis of climate sensitivity

This section uses observed climate data and the local constant estimator defined in (14) to estimate our proposed model (6). We consider two geographical input variables for the coefficient functions: (i) the latitude of a weather station; and (ii) each station's distance to the nearest coastline. We start this section by providing details of our data sources, and then compare our main results for each of the two varying-coefficient models with the constant coefficient model specified in (1). We find that each of our varying-coefficient models contribute to the study of the Earth's climate sensitivity at the global and regional levels.

4.1 Data description

This sub-section briefly discusses the sources and measurement of our three climate data series, as well as the two geographical variables that we use as inputs for our coefficient functions. We use three climate variables: near-surface air temperature measured in degrees Celsius ($^{\circ}C$), the amount of solar radiation that reaches the Earth's surface (measured in watts per square meter (Wm^2), and CO_2 concentration measured in parts per million by volume (ppmv). Data availability for CO_2 concentration limits our study to the period from 1959 to 2017.

4.1.1 Temperature data

The Climatic Research Unit (CRU) produces high-resolution grids of monthly temperature time series based on a large number of land weather stations around the globe (Harris et al., 2020). We consider the latest available data (version 4.04) covering the period from January 1901 to December 2019, which is available for download from the Center for Environmental Data Analysis Archive database (CEDA Archive, https://www.ceda.ac.uk/). The CEDA data base contains information on ten climate variables including mean near-surface (i.e 2 meters high) air temperatures which are provided for a 0.5° latitude by 0.5° longitude grid (equivalent to a spatial resolution of approximately 45×45 kilometers) over all land areas excluding Antarctica. We focus on grid cells that contain weather stations with adequate radiation data, and convert the relevant monthly time series for these cells into annual mean temperature series for each station. We discuss how we identify those weather stations with adequate radiation data below.

4.1.2 Solar radiation data

The Global Energy Balance Archive (GEBA, https://geba.ethz.ch/) stores data on instrumentally measured energy fluxes worldwide. This database provides quality controlled monthly means of downward shortwave radiation (global radiation) over the period 1950 – 2017 for almost 2,250 stations (see, e.g., Wild et al., 2017, for more information on the GEBA database). After excluding temporary stations (i.e. those on boats) and stations in the Arctic and Antarctic areas, we focus on 1,209 stations that report reliable observations (i.e., monthly mean radiation data without quality flags) for 24 consecutive months at least once during 1959 – 2017. Many of these stations report missing monthly observations that are unevenly distributed over time, and this hinders the computation of annual mean radiation time series. Figures 10 and 11 in Appendix D.1 provide a map of the chosen stations and an illustration of the "missingness" of

their radiation data, and Appendix D.1 also outlines how we fill in data gaps using information from neighboring stations and CRU climate variables not directly related to temperature. After imputing the missing values, we have a matrix of 1209×59 annual mean radiation observations.

4.1.3 CO_2 concentration data

We use annual CO_2 concentration data collected from the Mauna Loa Observatory (https://gml.noaa.gov/ccgg/trends/data.html) run by the National Oceanic and Atmospheric Administration (NOAA). Starting in 1959, this time series is longer than other broader measures of green house gas concentration (such as the Annual Greenhouse Gas Index at https://www.esrl.noaa.gov/gmd/aggi). Further, CO_2 provides a good proxy for all green house gases, as discussed in Hofmann et al. (2006), p. 617.

Figure 1 shows the station-averaged temperature, solar radiation, and atmospheric CO_2 concentration (in logarithmic scale) obtained from the balanced panel after filling in missing values. The observed $\ln(CO_2)$ between 1959–2017 displays an upward trend, which is consistent with a random walk plus drift as specified in Assumption 1. The station-averaged temperature $\overline{T}_t$ closely follows the trend in $\ln(CO_2)$. The aggregated solar radiation time series $\overline{R}_t$ is decreasing from the 1960s to the early 1980s, then starts to increase after that. This pattern is often described as global "dimming" and subsequent "brightening" in the literature (Norris and Wild, 2007), and it is consistent with a stochastic trend as specified in Assumption 1.

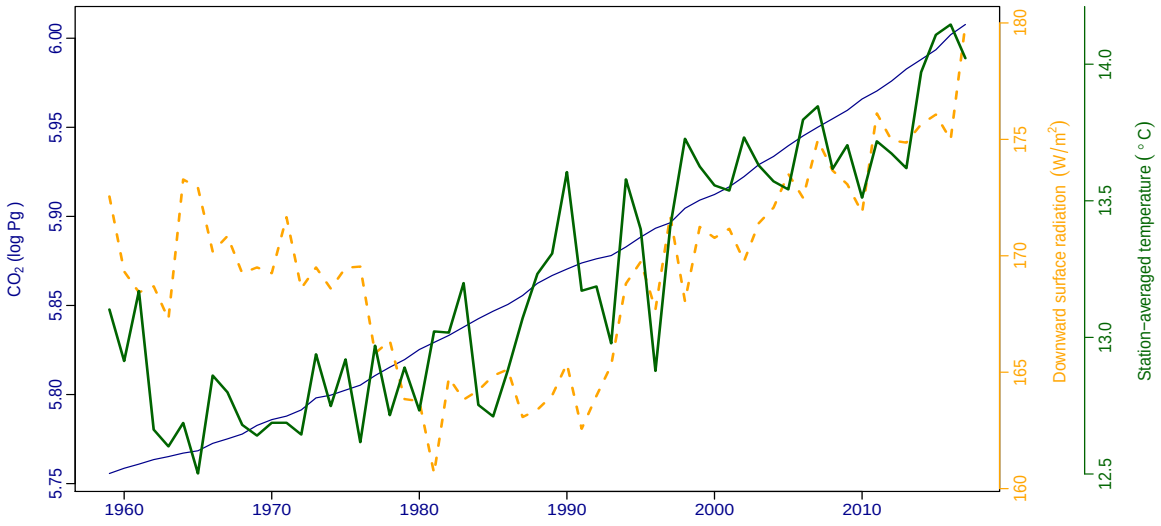

Figure 1: *Station-averaged temperature (° Celsius; green solid), solar radiation (Wm^2; orange dashed), and logarithms of CO_2 concentration (Pg: metric gigatons; blue solid) 1959 – 2017.*

4.1.4 Geographical variables

The latitude and longitude coordinates for each observation station are determined using the World Geodetic System 1984 (WGS84), which is based on the highly accurate International Terrestrial Reference Frame (see Kumar (1988) for details on WGS84). The distance to the coastline (DTC) for each station, is defined as the geodesic (i.e. the shortest path over the Earth's surface between two locations according to the standard reference ellipsoid model provided by WGS84), and we computed this using software written by Hijmans (2019). Figure 2 shows the spatial distribution of our 1209 weather stations over the continents, with the color scale of stations representing their distances (in meters) to the nearest coastline.

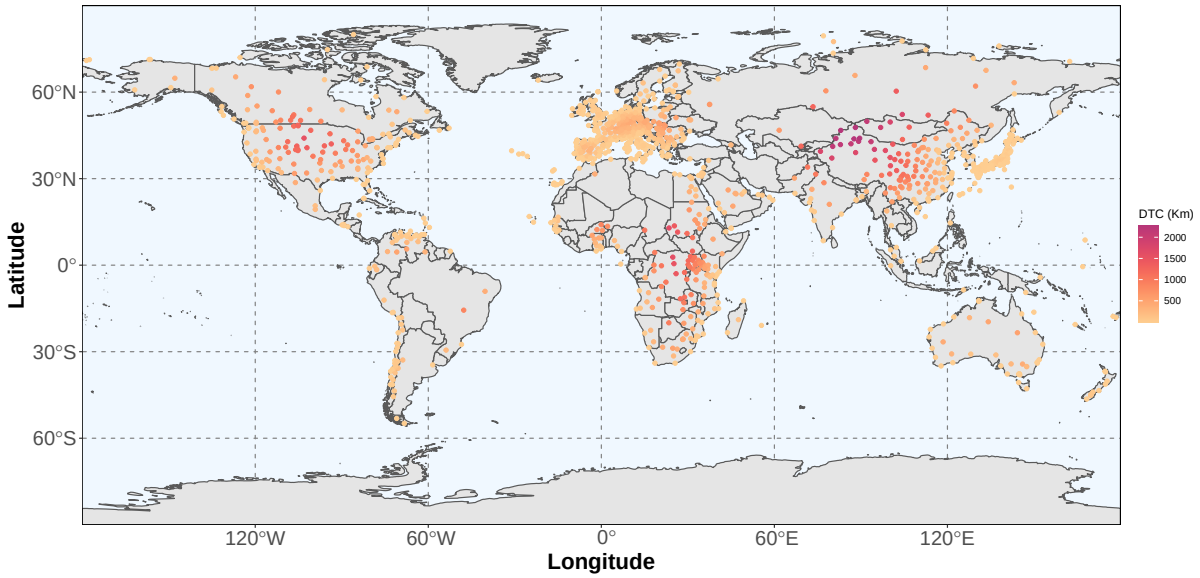

Figure 2: *Distribution of selected weather stations in the GEBA dataset together with their distances to the nearest coastline.*

4.2 Estimates of climate sensitivities that vary with latitude

We now present the empirical results of a latitude-varying panel data model. We first normalize the latitude values by

$$z_i = \frac{\text{Latitude}_i - \text{Latitude}_{min}}{\text{Latitude}_{max} - \text{Latitude}_{min}},$$

where Latitude_{max} and Latitude_{min} are the maximum and minimum values of latitudes, respectively. Substituting the z_is into (6), we obtain local constant estimates for the varying coefficient

functions using (14) with the bandwidth parameter selected via cross-validation. We use the dependent wild bootstrap in Yan et al. (2022) to obtain the confidence intervals. For a given level α (we use $\alpha = 0.05$), the bootstrap procedure is as follows:

1. For $\forall z \in (0,1)$, let $\widetilde{\boldsymbol{\beta}}(z)$ be defined in the same way as in (14) but with an over-smoothing bandwidth $\widetilde{h}$. Plugging $\widetilde{\boldsymbol{\beta}}(z_i)$ into (1) to obtain the estimated residuals given by $\widehat{\epsilon}_{i,t+1} = T_{i,t+1} - \boldsymbol{X}_{i,t}^{\top}\widetilde{\boldsymbol{\beta}}(z_i)$ for $i = 1,\ldots,N$ and $t = 1,\ldots,n-1$.

2. Collect the estimated residuals into vectors $\widehat{\epsilon}_t := (\widehat{\epsilon}_{1,t},\ldots,\widehat{\epsilon}_{N,t})^{\top}$, for $t = 1,\ldots,n-1$, form $\boldsymbol{T}_{t+1} := (T_{1,t+1},\ldots,T_{N,t+1})^{\top}$ and $\boldsymbol{X}_t := \left(\boldsymbol{X}_{1,t}^{\top},\ldots,\boldsymbol{X}_{N,t}^{\top}\right)^{\top}$ and then generate bootstrap samples using

$$\boldsymbol{T}_{t+1}^{*} = \boldsymbol{X}_t^{\top}\widetilde{\boldsymbol{\beta}}(z_i) + \boldsymbol{\epsilon}_t^{*},$$

   where $\boldsymbol{\epsilon}_t^{*} = \xi_t^{*}\widehat{\epsilon}_t$ with $\{\xi_t^{*}\} \overset{iid}{\sim} N(0,1)$.

3. Estimate $\widehat{\boldsymbol{\beta}}^{*}(z)$ $\forall z \in (0,1)$, following Eq. (13) but using the generated variable $\boldsymbol{T}_{t+1}^{*}$ as the response.

4. Repeat Steps 2 and 3 for J times. Let $\boldsymbol{q}_{\alpha}(z)$ be the α-quantile of J statistics $\widehat{\boldsymbol{\beta}}^{*}(z) - \widetilde{\boldsymbol{\beta}}(z)$. The $(1-\alpha)\times 100\%$ confidence interval of $\widehat{\boldsymbol{\beta}}(z)$ can be obtained as $\left[\widehat{\boldsymbol{\beta}}(z) - \boldsymbol{q}_{1-\alpha/2}(z), \widehat{\boldsymbol{\beta}}(z) - \boldsymbol{q}_{\alpha/2}(z)\right]$.

The estimated results for $\{\widehat{\beta}_1(z), \widehat{\beta}_2(z), \widehat{\gamma}_1(z), \widehat{\gamma}_2(z), \widehat{\gamma}_3(z)\}$ are illustrated in Figure 3, with the solid curves in black tracing out the estimated coefficient functions and the orange dotted lines providing their 95% confidence intervals. All estimated coefficient functions show significant variation with latitude.

The $\beta_1(z)$ function provides a measure of persistence in local temperature, and it shows a maximum occurring near the equator with decreases as latitudes rise in each hemisphere. This finding suggests that the areas near the equator may take longer to reach steady state following a change in GHGs or radiation. In addition, the $\beta_2(z)$ function suggests that local radiation effects on local temperature are strongest for areas that are north of the Arctic circle (and they increase temperature), while $\gamma_2(z)$ suggests that global radiation effects tend to depress local temperatures in the southern tropical areas.

The $\gamma_1(z)$ function contributes to measures of local and global climate sensitivity, and we observe that this function is mostly negative and relatively constant in the southern hemisphere, but statistically insignificant except for areas near the equator. In the northern hemisphere, $\gamma_1(z)$ increases with latitude, and it changes from negative to positive around the Northern Tropic

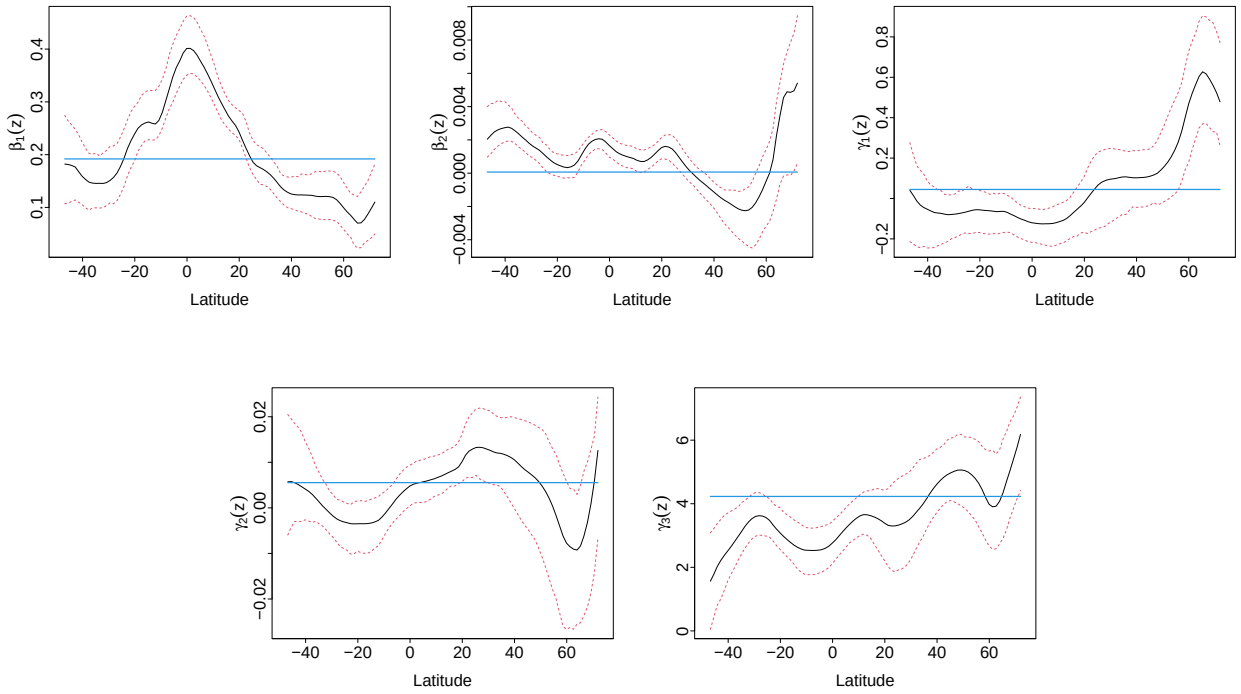

Figure 3: *Estimated coefficient functions of latitudes (solid black) together with their 95% confidence intervals (dotted orange) obtained by bootstrap. Within group estimates of the constant coefficient model (solid blue) are superimposed on the estimated functions.*

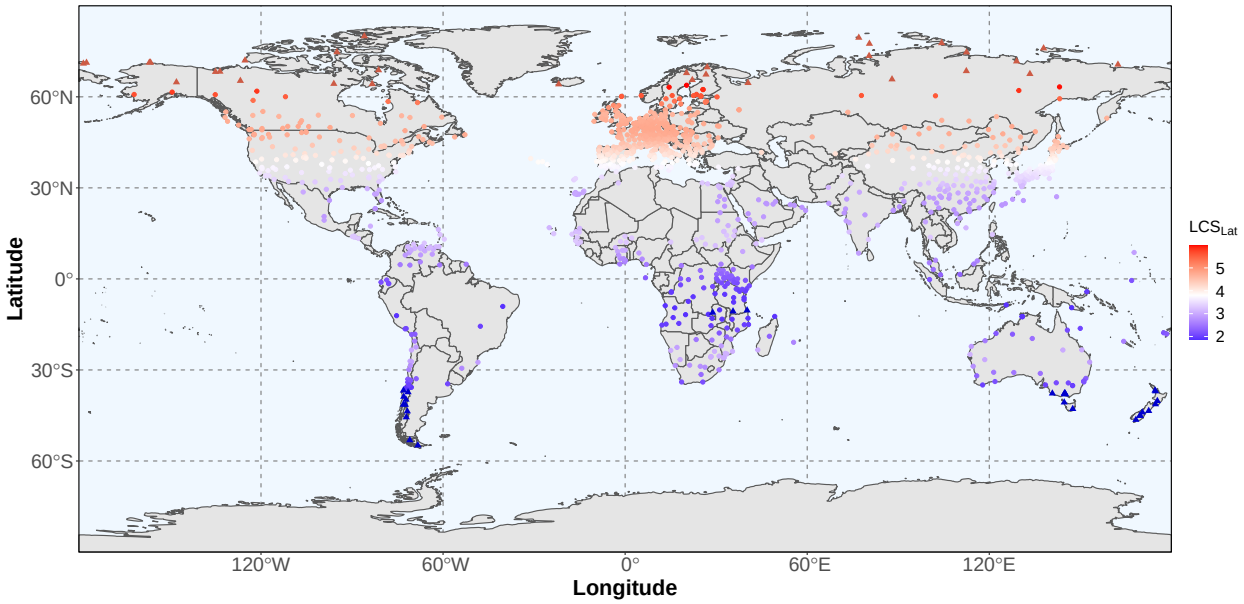

Figure 4: *Local climate sensitivity based on estimated coefficient functions that vary with latitude. Color indicates the LCS of weather stations with white indicating those stations for which LCS $\approx$ GCS. Stations with extreme climate sensitivities ($>$ 97.5% or $<$ 2.5% of LCS_{lat}) are highlighted with triangular symbols.*

(at 23° N). Recalling equation (10), this suggests that local climate sensitivity (LCS) is higher (lower) than global climate sensitivity (GCS), holding the effect of CO_2 on temperature (i.e. $\gamma_3(z)$)

constant. Hence, mid/high-latitude areas in the northern hemisphere are more likely to have higher LCS than the southern hemisphere and the equator.

The estimated $\gamma_3(z)$ function is mostly increasing from south to north, with three peaks at around 25°S, 10°N, and 50°N. Since LCS increases with $\gamma_3(z)$ (and hence with CO_2), the combined effects from $\hat{\gamma}_{11}(z)$ and $\hat{\gamma}_{31}(z)$ will lead to higher than average LCS in the northern hemisphere above the Tropic.

Figure 3 also presents estimates based on the constant coefficient model in (1) to facilitate comparison between the varying and constant coefficients. The constant coefficients are estimated using the within-group estimator in Phillips (2020), and they are shown in blue. They cross the (approximate) middle points of the ranges of the estimated coefficient functions, and they are outside the bootstrap confidence intervals for most latitudes and all coefficient functions. Significant variations of the estimated coefficient functions along latitude suggest that the varying-coefficient model is capable of revealing spatial patterns of temperature change in response to increasing atmospheric CO_2.

Figure 4 presents the estimated LCS for weather stations across all continents other than Antarctica. Here, the global climate sensitivity from the latitude-varying model is $\text{GCS}_{\text{Lat}} = 3.7°\text{C}$, and stations with $LCS \approx GCS$ are indicated by white points, whereas red and blue points indicate stations with LCS above and below the global average, respectively. Those stations in the top 2.5% of LCS are highlighted in dark red while those in the bottom 2.5% of LCS are highlighted in dark blue.

The GCS estimated from the latitude-varying model, $\text{GCS}_{\text{Lat}} = 3.7°\text{C}$, is very close to the GCS estimated from the constant coefficient model which is $\text{GCS} = 3.84°\text{C}$. Despite this, Figure 4 displays distinct spatial variations of LCS. In the northern hemisphere, the LCS increases monotonically in latitude, with areas near 50°N reaching LCS values of 5°C. The equator and the southern hemisphere have lower than average LCS, with areas near 30°S having LCS values of about 2.5°C.

Spatial patterns of the LCS suggest that the northern hemisphere generally has higher temperature changes than the southern hemisphere in response to the same CO_2 increase if solar radiation is stable. This pattern is consistent with patterns of observed temperature anomalies: global warming is ubiquitous, but high-latitude land regions in the northern hemisphere, especially those above 50° N, have experienced the largest temperature increases in recent decades (Hansen et al., 2006). The dependence of LCS on latitude may be attributed to surface albedo and

large scale circulation patterns, both of which are functions of latitude. Temperature increases in high-latitude regions drive the retreat of mountain and polar glaciers, leading to global sea level rise which threatens many coastal cities (IPCC, 2021, p.SPM-28). Our analysis shown in Figure 4 provides new econometric measures of the global and local effects of CO_2 emissions on temperature, which have potential to help policymakers of countries worldwide to develop optimal methods for mitigating future climate change.

4.3 Estimates of climate sensitivities that vary with distance to the coast

In this section, we investigate differences between coastal and inland climate sensitivities by incorporating our measures of distance to the coast (DTC) into the varying-coefficient panel data model. Figure 5 shows how the estimated coefficient functions in (6) vary with the natural logarithm of DTC. The maximum of $\beta_1(z)$ is located between 1 to 3 kilometers from the coastline (7–8 in $\ln(\text{DTC})$), indicating that the present temperature will be a better predictor of next year's temperature for regions along the coastline than for remote inland regions. The $\beta_2(z)$ function has positive (and statistically significant) estimates for regions within 20–50 kilometers to coast (10–11 in $\ln(\text{DTC})$), suggesting relatively strong local radiation effect in these regions. The $\gamma_3(z)$ function is relatively high for stations whose distances to coastline are greater than 150 kilometers (12 in $\ln(\text{DTC})$). The estimated $\gamma_1(z)$ and $\gamma_2(z)$ functions representing the aggregate temperature and radiation effects are close to their fixed counterparts obtained by within-group estimation.

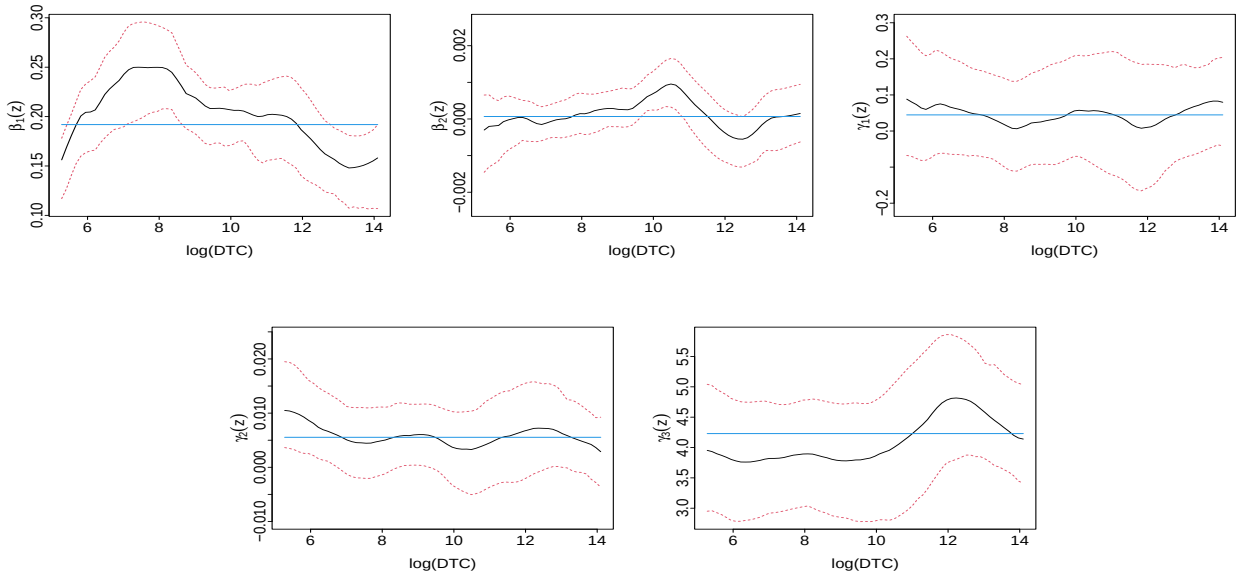

Figure 5: *Estimated coefficient functions of the (natural logarithm of) distance to the coast (solid black) together with 95% confidence intervals (dotted orange) obtained by bootstrap. Within group estimates of the constant coefficient model (solid blue) are superimposed on the estimated functions.*

Figure 6 displays the LCS at the station level using the DTC-varying model. The difference in LCS is less pronounced in the DTC-varying model than in the latitude-varying model, and the largest difference between coastal and inland LCS is less than 0.5°C. Nevertheless, stations located in coastal regions generally are estimated to have lower LCS than inland stations. The relatively stable coastal climates compared with inland climates is attributable to ocean's greater capacity to absorb heat relative to rock and soil in inland regions. In the areas that are farthest to coastlines (DTC greater than 1000km), LCS starts to decrease again. This may be due to limited water reservoirs in these areas. Water vapor is a powerful greenhouse gas. Following an increase in temperature due to CO_2, areas with more water reservoirs are more likely to produce water vapor which enhances the greenhouse effect.

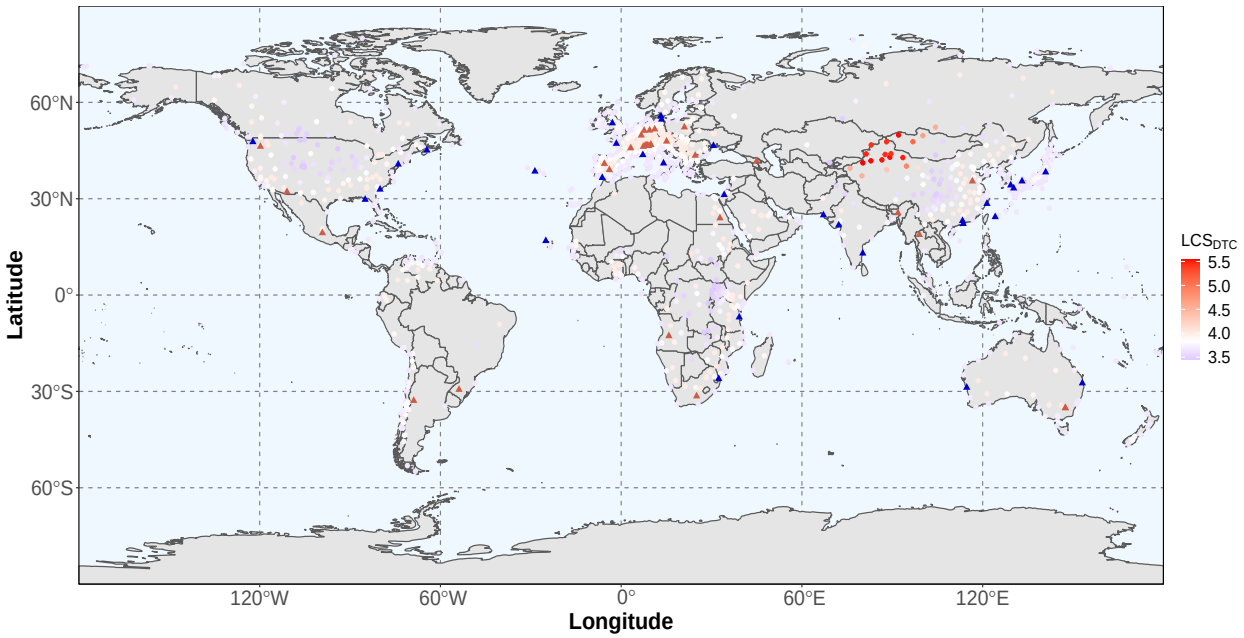

Figure 6: *Local climate sensitivity based on estimated coefficient functions that vary with the (natural logarithm of) distance to the coastline (DTC). Color indicates the LCS of weather stations, with white indicating those stations for which LCS $\approx$ GCS. Stations with extreme sensitivities ($>$ 97.5% or $<$ 2.5% of LCS_{DTC}) are highlighted with triangular symbols.*

We repeat the analysis in Sections 4.2 and 4.3 using the local linear estimator rather than the local constant estimator and find that both estimators deliver very similar results. Figures associated with the latter set of results are provided in Appendix E.

4.4 Out-of-sample forecasting comparison

Baltagi (2008) finds that in a variety of applications, a simple panel model with parsimonious parameters offers better out-of-sample forecasting performance than sophisticated models that allow heterogeneity. In this section, we compare the forecasting performance of our latitude-varying model with the constant coefficient (CC) model in Phillips et al. (2020). We estimate the model parameters using the temperature, CO_2 and radiation series from 1959 to 1999 (41 years). The fixed effect coefficients α_i are estimated using the average of $T_{i,t+1} - \boldsymbol{X}_{it}^\top \widehat{\boldsymbol{\beta}}$ or $T_{i,t+1} - \boldsymbol{X}_{it}^\top \widehat{\boldsymbol{\beta}}(z_i)$. We then forecast the temperature series from 2000 to 2017 (18 years) using the estimated parameters and the observed values of CO_2 and radiation in the forecasting period. In other words, we use the models to predict temperatures up to 18 years ahead conditional on the observed trajectories of CO_2 and radiation.

We let $\widehat{T}_{i,t+1}$ denote the temperature forecast and compute the root mean squared forecast error (RMSFE) across stations or time using

$$\text{RMSFE}_t = \sqrt{\frac{1}{N}\sum_i (T_{i,t+1} - \widehat{T}_{i,t+1})^2} \qquad \text{or} \qquad \text{RMSFE}_i = \sqrt{\frac{1}{18}\sum_t (T_{i,t+1} - \widehat{T}_{i,t+1})^2}.$$

For the latitude-varying model, we first select a bandwidth denoted by H using cross validation in the estimation window, then estimate the coefficients using nine different bandwidths between $H/2$ to $2H$. The RMSFE_t from all nine bandwidths are averaged to yield $\text{RMSFE}_{t,vc}$, and $\text{RMSFE}_{i,vc}$ is obtained in a similar fashion.

Figure 7 presents the difference between ($\text{RMSFE}_{t,cc}$ and $\text{RMSFE}_{t,vc}$) in the years from 2000 to 2017. The latitude-varying model outperforms the CC model in 15 out of 18 years. Figure 8 presents the difference between $\text{RMSFE}_{i,cc}$ and $\text{RMSFE}_{i,vc}$ across stations. For most of the stations that are located in the southern hemisphere or in the high-latitude regions in the northern hemisphere (>60° N), the latitude-varying model yields substantially lower RMSFE than the CC model. The CC model holds a slight edge in the region between the Northern Tropic and 60° N. Figures 7 and 8 also presents the range of RMSFE across the nine bandwidths used for the latitude-varying model, and the results suggest that the performance of our proposed model is robust to a wide range of bandwidths.

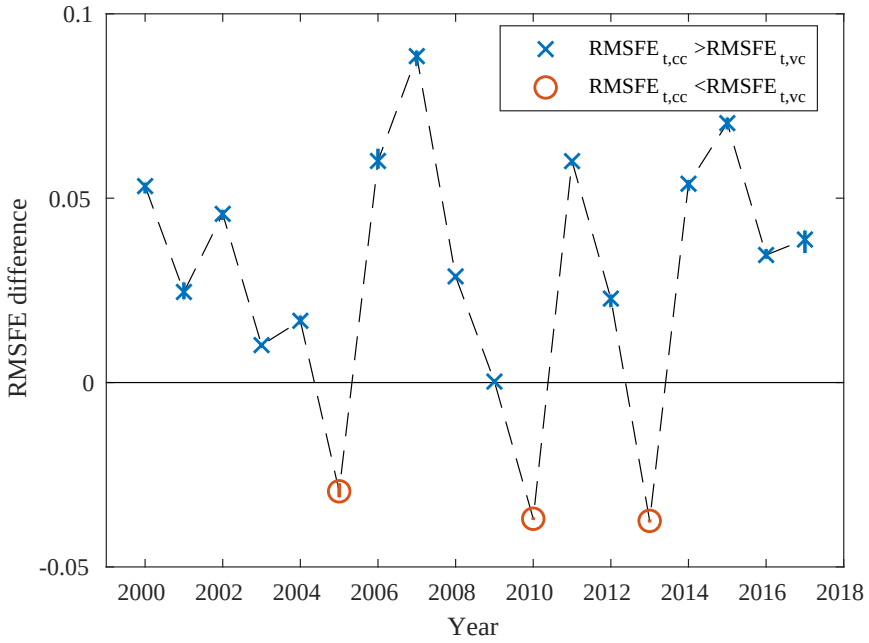

Figure 7: *RMSFE differences between the CC and the latitude-varying models in different years. This figure plots $RMSFE_{t,cc} - RMSFE_{t,vc}$ against the year. We use the blue crosses (red circles) to indicate that the latitude-varying model outperforms (underperforms) the CC model. The vertical lines represent the range of RMSFE differences across the nine bandwidths used for the VC model.*

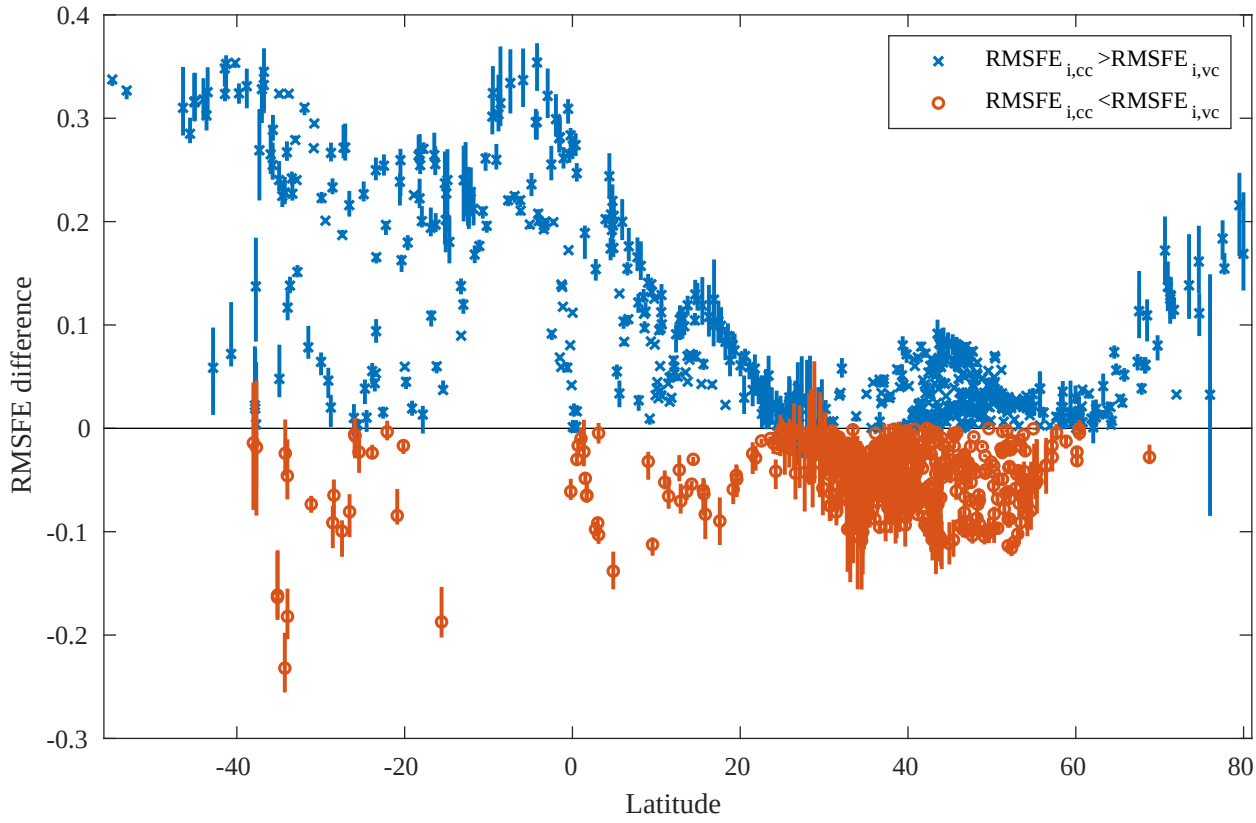

Figure 8: *RMSFE differences between the CC and the latitude-varying models in difference stations. This figure plots $RMSFE_{i,cc} - RMSFE_{i,vc}$ against the station latitude. We use the blue crosses (red circles) to indicate that the latitude-varying model outperforms (underperforms) the CC model. The vertical lines represent the range of RMSFE differences across the nine bandwidths used for the VC model.*

5 Conclusion

In this paper, we consider a varying-coefficient dynamic panel data model for investigating spatial differences of Earth's climate sensitivity to atmospheric greenhouse gases. This proposed spatial-temporal system uses functions of time invariant variables to account for the heterogeneity while accounting for cointegrating relationships between climate time series at the individual-station and global levels. The asymptotic theory and simulation results show that the proposed model and inferential methods are valid for studying responses of Earth's climate to non-stationary forcing variables driven by both deterministic and stochastic trends.

The empirical application using data from 1209 weather stations show that mid/high-latitude regions in the northern hemisphere are more sensitive to changes in GHGs than the equatorial area or the southern hemisphere, and that inland areas are more sensitive than coastal areas. Our latitude-varying model estimates suggest that global temperature would rise by 3.7°C following a doubling CO_2, with areas above 50°N rising by more than 5°C and areas near 30°S rising by 2.5°C. In the DTC-varying model, coastal areas have slightly lower estimated climate sensitivity than inland, but the largest difference between coastal and inland areas is less than 0.5°C.

In an out-of-sample forecasting exercise, we demonstrate that our latitude-varying model outperforms the parsimonious constant coefficient model in forecasting future temperatures. The gain in forecast performance is particularly prominent for regions where the LCS differs substantially from the GCS.

Our analysis shows the potential of varying-coefficient econometric models to provide evidence of ubiquitous global warming, and to pick out regions facing more serious consequences of climate change. The current research can be further extended to include a mixture of fixed and varying coefficients to represent different effects of global CO_2 forcing and local aerosol pollution. A time-varying version of the proposed dynamic panel data framework that allows the effects of forcing variables to shift with time may be useful for studying the changing speed of global warming in the past sixty years. We leave these topics for future research.

References

Andreae, M. O., Jones, C. D. and Cox, P. M. (2005), 'Strong present-day aerosol cooling implies a hot future', *Nature* **435**(7046), 1187–1190.

Baltagi, B. H. (2008), 'Forecasting with panel data', *Journal of forecasting* **27**(2), 153–173.

Boer, G. and Yu, B. (2003), 'Climate sensitivity and response', *Climate Dynamics* **20**(4), 415–429.

Bony, S., Colman, R., Kattsov, V. M., Allan, R. P., Bretherton, C. S., Dufresne, J.-L., Hall, A., Hallegatte, S., Holland, M. M., Ingram, W. et al. (2006), 'How well do we understand and evaluate climate change feedback processes?', *Journal of Climate* **19**(15), 3445–3482.

Breiman, L. (2001), 'Random forests', *Machine learning* **45**(1), 5–32.

Cai, Z. and Li, Q. (2008), 'Nonparametric estimation of varying coefficient dynamic panel data models', *Econometric Theory* **24**(5), 1321–1342.

Cai, Z., Li, Q. and Park, J. Y. (2009), 'Functional-coefficient models for nonstationary time series data', *Journal of Econometrics* **148**(2), 101–113.

Chen, L., Gao, J. and Vahid, F. (2022), 'Global temperatures and greenhouse gases: A common features approach', *Journal of Econometrics* **230**(2), 240–254.

Connor, G., Hagmann, M. and Linton, O. (2012), 'Efficient semiparametric estimation of the Fama–French model and extensions', *Econometrica* **80**(2), 713–754.

Croissant, Y. and Millo, G. (2008), 'Panel data econometrics in R: The plm package', *Journal of Statistical Software* **27**(2), 1–43.

Dong, C., Gao, J. and Peng, B. (2021), 'Varying-coefficient panel data models with nonstationarity and partially observed factor structure', *Journal of Business & Economic Statistics* **39**(3), 700–711.

Fan, J. and Zhang, W. (1999), 'Statistical estimation in varying coefficient models', *The Annals of Statistics* **27**(5), 1491–1518.

Feng, G., Gao, J., Peng, B. and Zhang, X. (2017), 'A varying-coefficient panel data model with fixed effects: theory and an application to US commercial banks', *Journal of Econometrics* **196**(1), 68–82.

Gay-Garcia, C., Estrada, F. and Sánchez, A. (2009), 'Global and hemispheric temperatures revisited', *Climatic Change* **94**(3), 333–349.

Hansen, J., Sato, M., Ruedy, R., Lo, K., Lea, D. W. and Medina-Elizade, M. (2006), 'Global temperature change', *Proceedings of the National Academy of Sciences* **103**(39), 14288–14293.

Harris, I., Jones, P. D., Osborn, T. J. and Lister, D. H. (2014), 'Updated high-resolution grids of monthly climatic observations–the CRU TS3. 10 Data set', *International Journal of Climatology* **34**(3), 623–642.

Harris, I., Osborn, T. J., Jones, P. and Lister, D. (2020), 'Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset', *Scientific Data* **7**(1), 1–18.

Henderson-Sellers, A. and Wilson, M. (1983), 'Surface albedo data for climatic modeling', *Reviews of Geophysics* **21**(8), 1743–1778.

Hijmans, R. J. (2019), *Geosphere: Spherical Trigonometry*. R package version 1.5-10.
URL: *https://CRAN.R-project.org/package=geosphere*

Hofmann, D. J., Butler, J. H., Dlugokencky, E. J., Elkins, J. W., Masarie, K., Montzka, S. A. and Tans, P. (2006), 'The role of carbon dioxide in climate forcing from 1979 to 2004: introduction of the annual greenhouse gas index', *Tellus B: Chemical and Physical Meteorology* **58**(5), 614–619.

IPCC (2014), *Climate change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, IPCC, Geneva, Switzerland.

IPCC (2018), *Global warming of* 1.5°*C. An IPCC Special Report on the impacts of global warming of* 1.5°*C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty*, Cambridge University Press, Cambridge.

IPCC (2021), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, IPCC, Geneva, Switzerland.

Kaufmann, R. K., Kauppi, H., Mann, M. L. and Stock, J. H. (2013), 'Does temperature contain a stochastic trend: linking statistical results to physical mechanisms', *Climatic Change* **118**(3), 729–743.

Kumar, M. (1988), 'World geodetic system 1984: A modern and accurate global reference frame', *Marine Geodesy* **12**(2), 117–126.

Li, Q. and Racine, J. S. (2007), *Nonparametric Econometrics: Theory and Practice*, Princeton University Press.

Magnus, J. R., Melenberg, B. and Muris, C. (2011), 'Global warming and local dimming: The statistical evidence', *Journal of the American Statistical Association* **106**(494), 452–464.

Mitchell, T. D. and Jones, P. D. (2005), 'An improved method of constructing a database of monthly climate observations and associated high-resolution grids', *International Journal of Climatology: A Journal of the Royal Meteorological Society* **25**(6), 693–712.

Myhre, G., Myhre, C. E., Samset, B. H. and Storelvmo, T. (2013), 'Aerosols and their relation to global climate and climate sensitivity', *Nature Education Knowledge* **4**.

Norris, J. R. and Wild, M. (2007), 'Trends in aerosol radiative effects over europe inferred from observed cloud cover, solar "dimming", and solar "brightening"', *Journal of Geophysical Research: Atmospheres* **112**(D8).

Park, J. Y. and Phillips, P. C. (1989), 'Statistical inference in regressions with integrated processes: Part 2', *Econometric Theory* **5**(1), 95–131.

Phillips, P. C. (2020), 'Dynamic panel modeling of climate change', *Econometrics* **8**(3), 30.

Phillips, P. C., Leirvik, T. and Storelvmo, T. (2020), 'Econometric estimates of Earth's transient climate sensitivity', *Journal of Econometrics* **214**(1), 6–32.

Pithan, F. and Mauritsen, T. (2014), 'Arctic amplification dominated by temperature feedbacks in contemporary climate models', *Nature geoscience* **7**(3), 181–184.

Power, H. (2003), 'Trends in solar radiation over Germany and an assessment of the role of aerosols and sunshine duration', *Theoretical and Applied Climatology* **76**(1-2), 47–63.

Screen, J. A. and Simmonds, I. (2010), 'The central role of diminishing sea ice in recent Arctic temperature amplification', *Nature* **464**(7293), 1334–1337.

Storelvmo, T., Leirvik, T., Lohmann, U., Phillips, P. C. and Wild, M. (2016), 'Disentangling greenhouse warming and aerosol cooling to reveal earth's climate sensitivity', *Nature Geoscience* **9**(4), 286–289.

Sun, Y., Cai, Z. and Li, Q. (2016), 'A consistent nonparametric test on semiparametric smooth coefficient models with integrated time series', *Econometric Theory* **32**(4), 988–1022.

Wild, M., Ohmura, A., Schär, C., Müller, G., Folini, D., Schwarz, M., Hakuba, M. Z. and Sanchez-Lorenzo, A. (2017), 'The global energy balance archive (GEBA) version 2017: A database for worldwide measured surface energy fluxes', *Earth System Science Data* **9**(2), 601–613.

Yan, Y., Gao, J. and Peng, B. (2022), *Asymptotics for time-varying vector MA(∞) process*. Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4011417.

Appendix A Derivation of Global and Local Climate Sensitivity

To find the steady state temperature, we plug $\overline{T}^e_{i,t} = \overline{T}^e_{i,t+1}$ into equation (6) to obtain

$$T^e_{i,t} = (1-\beta_1(z_i))^{-1}\left(\alpha_i + \beta_2(z_i)R_{i,t} + \gamma_1(z_i)\overline{T}^e_t + \gamma_2(z_i)\overline{R}_t + \gamma_3(z_i)\ln(CO2_t)\right) + o_p(1). \tag{A.1}$$

Taking the average of equation (A.1) across stations and rearranging yields

$$\begin{aligned}\left(1 - \frac{1}{N}\sum_{i=1}^{N}\frac{\gamma_1(z_i)}{1-\beta_1(z_i)}\right)\overline{T}^e_t = &\frac{1}{N}\sum_{i=1}^{N}\frac{\alpha_i}{1-\beta(z_i)} + \frac{1}{N}\sum_{i=1}^{N}\frac{\beta_2(z_i)}{1-\beta_1(z_i)}R_{i,t} \\ &+ \frac{1}{N}\sum_{i=1}^{N}\frac{\gamma_2(z_i)}{1-\beta_1(z_i)}\overline{R}_t + \frac{1}{N}\sum_{i=1}^{N}\frac{\gamma_3(z_i)}{1-\beta_1(z_i)}\ln(CO2_t) + o_p(1).\end{aligned} \tag{A.2}$$

Let $\overline{T}^e_{t_1}$ and $\overline{T}^e_{t_2}$ denote the global steady state temperature associated with CO_2 concentration that doubles between t_1 and t_2, i.e., $CO2_{t_2} = 2CO2_{t_1}$. Assume that the radiation level stays constant, i.e, $R_{i,t_1} = R_{i,t_2}$ for all i and $\overline{R}_{t_1} = \overline{R}_{t_2}$. Equation (A.2) applies to both t_1 and t_2, and the difference of Equation (A.2) at the two time periods yields

$$\overline{T}^e_{t_2} - \overline{T}^e_{t_1} = \left(1 - \frac{1}{N}\sum_{i=1}^{N}\frac{\gamma_1(z_i)}{1-\beta_1(z_i)}\right)^{-1}\frac{1}{N}\sum_{i=1}^{N}\frac{\gamma_3(z_i)}{1-\beta_1(z_i)}\ln(2), \tag{A.3}$$

which is the GCS in Eq. (9).

Given the regularity conditions in Assumption 4, $\overline{\theta}_{j,N} \xrightarrow{a.s.} \overline{\theta}_j := E_N\left(\frac{\gamma_j(z_i)}{1-\beta_1(z_i)}\right)$ as $N \to \infty$ for $j = \{1,3\}$. Hence, $GCS_N \xrightarrow{a.s.} GCS := \left(1-\overline{\theta}_1\right)^{-1}\overline{\theta}_3\ln(2)$.

For location i, the difference between equation (A.1) when $t = t_1$ and when $t = t_2$ is

$$T^e_{i,t_2} - T^e_{i,t_1} = (1-\beta_1(z_i))^{-1}\gamma_1(z_i)\left(\overline{T}^e_{t_2} - \overline{T}^e_{t_1}\right) + \gamma_3(z_i)\ln(2), \tag{A.4}$$

from which we obtain the LCS in Eq. (10). It is easy to see that $LCS_N(z) \xrightarrow{a.s.} LCS(z) := \frac{\gamma_1(z)}{1-\beta_1(z)}GCS + \frac{\gamma_3(z)}{1-\beta_1(z)}\ln(2)$ as $N \to \infty$.

Appendix B Simulation results

In this section, we carry out simulation studies to examine and compare the finite sample performance of the local constant and local level estimators in the context of our empirical setting. We consider the following data generating processes (DGPs) based on model (6) and the empirical characteristics of our data:

- Generate $\ln(CO2_t) = 5.6 + 0.0025t + 0.0025\mathcal{U}_{g1,t}$ where $\mathcal{U}_{g1,t}$ is a standard random walk with $\mathcal{U}_{g1,t} = \sum_{s=1}^{t} \mathrm{u}_{g1,s}$, and $\mathrm{u}_{g1,s} \overset{i.i.d.}{\sim} N(0,1)$.

- Randomly select N ($N < 1209$) stations from all 1209 stations, denote their annual radiation in year 1959 as $R_{i,0}$ for $i = 1, \ldots, N$, and then form the stationary component of radiation using $R^0_{i,t} = R_{i,0} + R^*_{i,t}$ where $R^*_{it} \overset{i.i.d.}{\sim} N(0, 25)$. We then generate radiation time series for station i from $R_{i,t} = R^0_{i,t} + \delta_R(z_i)\mathcal{U}_{g2i,t} + \mathcal{U}_{P,it}$ where the $\mathcal{U}_{g2i,t}$ are standard random walks that are independent from $\mathcal{U}_{g1,t}$, $\delta_R(z) = 2 + \cos(2\pi z)$, and $\mathcal{U}_{P,it}$ are standard random walks that are cross-sectionally independent.

- The station-specific effect is set to $\alpha_i \overset{i.i.d.}{\sim} N(-15, 7.4^2)$.

- For $t = 0$, set $\overline{T}_t$ and $T_{i,t}$ to the steady state associated with α_i, $\ln(CO2_0)$ and $R_{i,0}$ using Eqs. (A.2) and (A.1).

- For $t \in \{1, \ldots, n-1\}$, generate $T_{i,t+1} = \alpha_i + \beta_1(z_i)T_{i,t} + \beta_2(z_i)R_{i,t} + \gamma_1(z_i)\overline{T}_t + \gamma_2(z_i)\overline{R}_t + \gamma_3(z_i)\ln(CO2_t) + \epsilon_{i,t+1}$, where $\epsilon_{i,t+1} \overset{iid}{\sim} N(0, 0.5^2)$.

We select $z_i = i/N$ for $i = 1, \ldots, N$, and set the coefficient functions $\boldsymbol{\beta}(z)$ to fluctuate around the empirical values of the constant coefficient model. In particular, for $z \in [0, 1]$ we consider the following coefficient functions:

$$\beta_1(z) = 0.3 + 0.1\sin\left(\frac{3\pi z}{2}\right), \quad \beta_2(z) = -0.0001 + 0.01\cos(\pi z),$$
$$\gamma_1(z) = 0.15 - 0.2\cos(\pi z), \quad \gamma_2(z) = 0.025 - 0.01\sin(\frac{3\pi z}{2}), \quad \gamma_3(z) = 4.2 + 2\cos\left(\frac{3\pi z}{2} + \frac{\pi}{4}\right).$$

The two estimators that we wish to compare via simulation are the local constant estimator which is given in Equation (14) and the local linear estimator that we outline below. We assume that $\boldsymbol{\beta}(z)$ has continuous derivatives up to the second order, and then by a Taylor expansion we have

$$\boldsymbol{\beta}(z_i) = \boldsymbol{\beta}(z) + \boldsymbol{\beta}'(z)(z_i - z) + O\left[(z_i - z)^2\right], \tag{B.1}$$

where $\boldsymbol{\beta}'(z)$ is the first derivative of $\boldsymbol{\beta}(z)$. Then $\left(\boldsymbol{\beta}^\top(z), h[\boldsymbol{\beta}'(z)]^\top\right)^\top$ can be estimated by solving the minimization problem

$$\begin{pmatrix} \boldsymbol{\beta}(z) \\ h\boldsymbol{\beta}'(z) \end{pmatrix} = \arg\min_{\boldsymbol{a},\boldsymbol{b}} \sum_{i=1}^{N}\sum_{t=1}^{n-1} \left(\widetilde{T}_{i,t+1} - \widetilde{\boldsymbol{X}}^\top_{i,t}\left(\boldsymbol{a} + \boldsymbol{b}\left(\frac{z_i - z}{h}\right)\right)\right)^2 K\left(\frac{z_i - z}{h}\right), \tag{B.2}$$

where $\widetilde{T}_{i,t+1}$ and $\widetilde{X}_{i,t}$ are the demeaned values of $T_{i,t+1}$ and $X_{i,t}$, $K(z)$ is a kernel function and h is the bandwidth parameter.

Denote $\check{\boldsymbol{X}}_{i,t} = \left(\widetilde{\boldsymbol{X}}_{i,t}^{\top}, \frac{z_i - z}{h}\widetilde{\boldsymbol{X}}_{i,t}^{\top}\right)^{\top}$, and then the local linear estimator for $\boldsymbol{\beta}(z)$ is given by:

$$\widehat{\boldsymbol{\beta}}(z) = [\mathbf{I}_5, \mathbf{0}_{5\times 5}] \left(\sum_{i=1}^{N} \sum_{t=1}^{n-1} \check{\boldsymbol{X}}_{i,t} \check{\boldsymbol{X}}_{i,t}^{\top} K\left(\frac{z_i - z}{h}\right) \right)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{n-1} \check{\boldsymbol{X}}_{i,t} \widetilde{T}_{i,t+1} K\left(\frac{z_i - z}{h}\right), \tag{B.3}$$

where $\mathbf{I}_5$ is an identity matrix, and $\mathbf{0}_{5\times 5}$ is a 5×5 matrix of zeros.

B.1 Cross validation for bandwidth selection

We first investigate bandwidth selection using a leave-one-unit-out cross-validation method. For both the local constant and the local linear estimator, we choose the Epanechnikov kernel, $K(z) = 0.75(1 - z^2)_+$, and set the bandwidth to $h = cN^{-1/5}n^{-2/5}$ with $c \in [1, 10]$. We repeat the simulation for $B = 1000$ times. In the bth simulation, for $b = 1, \ldots, B$, we compute

$$\text{MSE}(b; c) = \frac{1}{Nn} \sum_{i=1}^{N} \sum_{t=1}^{n} \left(\widetilde{T}_{i,t+1}^{(b)} - \widehat{T}_{i,t+1}^{(b)} \right)^2,$$

where $\widehat{T}_{i,t+1}^{(b)} = \widetilde{\boldsymbol{X}}_{i,t}^{(b)\top} \widehat{\boldsymbol{\beta}}_{(-i)}^{(b)}(z_i)$, and $\widehat{\boldsymbol{\beta}}_{(-i)}^{(b)}(z)$ is the coefficient function estimated from leaving the unit i out. The average of $\text{MSE}(b; c)$ across B simulations is plotted against c in Figure 9 for $n = 50$, $N = 200$, and $n = 50$, $N = 1000$. From Figure 9, the local linear estimator is more stable at a larger bandwidth, but the minimum MSE is similar when we compare the local constant and the local linear methods. In the empirical section, we use cross-validation to select the bandwidth.

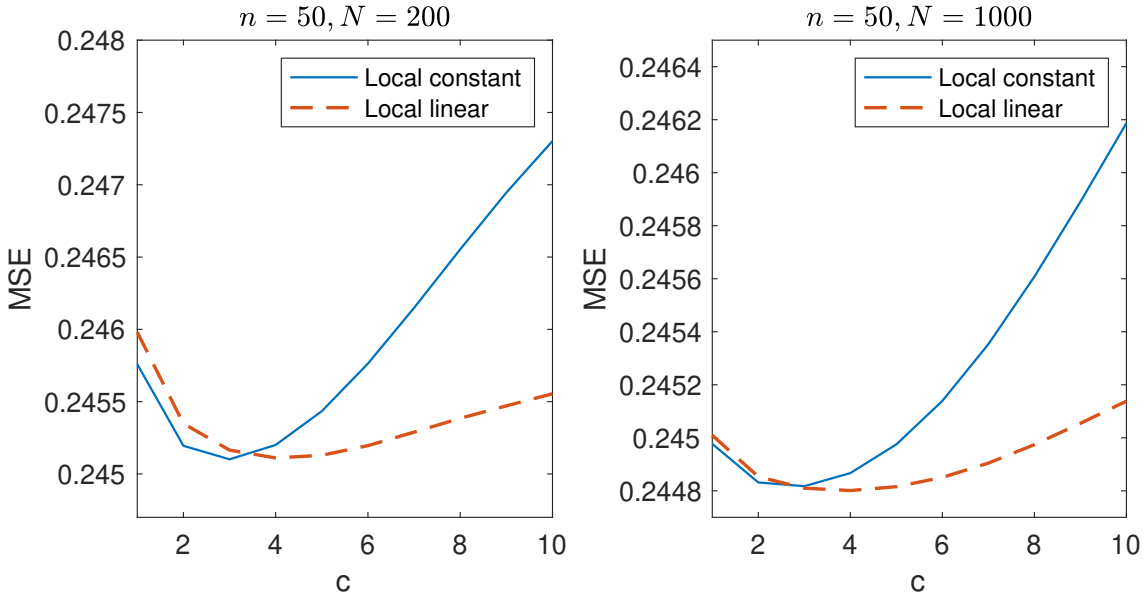

Figure 9: *Cross-validation results. This figures presents average MSE from 1000 simulations that use the leave-one-unit-out method. The bandwidth h is set to $cN^{-1/5}n^{-2/5}$ for $n = 50$, $N = 200$, and $n = 50$, $N = 1000$ with $c \in [1, 10]$.*

B.2 Numerical evaluations

In this section, we investigate the behavior of the estimator when the sample size is varied We consider $N \in \{50, 200, 1000\}$ and $n \in \{25, 50, 100, 200\}$ and the Epanechnikov kernel. We set the bandwidth $h = cN^{-1/5}n^{-2/5}$, and choose $c = 3$ for the local constant estimator and $c = 4$ for the local linear estimator given the cross-validation results in Figure 9.

Table 1: *Relative Squared Errors (RSEs) between the true coefficient function and its estimate*

	Local Constant					Local Linear				
	$N\backslash n$	25	50	100	200	$N\backslash n$	25	50	100	200
$\beta_1(z)$	50	0.224	0.131	0.086	0.083	50	0.272	0.162	0.102	0.068
		(0.036)	(0.013)	(0.005)	(0.004)		(0.045)	(0.016)	(0.007)	(0.003)
	200	0.193	0.099	0.055	0.047	200	0.226	0.122	0.069	0.043
		(0.017)	(0.005)	(0.002)	(0.001)		(0.019)	(0.007)	(0.003)	(0.001)
	1000	0.180	0.085	0.041	0.025	1000	0.199	0.098	0.052	0.029
		(0.007)	(0.002)	(0.001)	(0.000)		(0.008)	(0.003)	(0.001)	(0.000)
$\beta_2(z)$	50	0.583	0.425	0.340	0.316	50	0.643	0.434	0.303	0.200
		(0.304)	(0.145)	(0.093)	(0.091)		(0.341)	(0.147)	(0.064)	(0.025)
	200	0.321	0.245	0.195	0.183	200	0.345	0.243	0.161	0.110
		(0.082)	(0.044)	(0.032)	(0.035)		(0.093)	(0.042)	(0.017)	(0.007)
	1000	0.168	0.131	0.108	0.092	1000	0.171	0.118	0.082	0.053
		(0.020)	(0.013)	(0.010)	(0.009)		(0.020)	(0.009)	(0.004)	(0.001)
$\gamma_1(z)$	50	1.088	0.728	0.531	0.418	50	1.234	0.777	0.544	0.389
		(1.182)	(0.474)	(0.204)	(0.117)		(1.405)	(0.511)	(0.214)	(0.102)
	200	0.812	0.515	0.372	0.285	200	0.892	0.539	0.373	0.271
		(0.690)	(0.204)	(0.095)	(0.046)		(0.830)	(0.223)	(0.098)	(0.043)
	1000	0.509	0.305	0.208	0.158	1000	0.533	0.311	0.207	0.150
		(0.216)	(0.062)	(0.024)	(0.012)		(0.233)	(0.067)	(0.026)	(0.012)
$\gamma_2(z)$	50	0.466	0.314	0.224	0.174	50	0.545	0.347	0.232	0.158
		(0.194)	(0.080)	(0.035)	(0.018)		(0.270)	(0.098)	(0.040)	(0.017)
	200	0.288	0.198	0.147	0.117	200	0.321	0.210	0.144	0.105
		(0.072)	(0.029)	(0.014)	(0.008)		(0.088)	(0.033)	(0.014)	(0.007)
	1000	0.157	0.109	0.080	0.064	1000	0.167	0.108	0.075	0.056
		(0.020)	(0.008)	(0.003)	(0.002)		(0.024)	(0.008)	(0.003)	(0.002)
$\gamma_3(z)$	50	0.753	0.377	0.215	0.150	50	0.861	0.410	0.227	0.149
		(0.810)	(0.189)	(0.035)	(0.014)		(1.045)	(0.202)	(0.042)	(0.015)
	200	0.461	0.230	0.147	0.105	200	0.511	0.245	0.147	0.101
		(0.262)	(0.052)	(0.016)	(0.006)		(0.313)	(0.060)	(0.017)	(0.006)
	1000	0.254	0.131	0.081	0.059	1000	0.267	0.131	0.077	0.055
		(0.084)	(0.015)	(0.004)	(0.002)		(0.093)	(0.016)	(0.004)	(0.002)

The table reports the means and standard deviations (in parentheses) of RSEs across 1000 simulations. The RSE from each simulation is evaluated using 100 different values of z that are equal-spaced in $(0, 1)$.

Let $\beta(z)$ be a general notation for the coefficient functions such that $\beta(z) \in [\beta_1(z), \beta_2(z), \gamma_1(z), \gamma_2(z), \gamma_3(z)]$. The relative squared error (RSE) between the true function $\beta(z)$ and its estimator $\widehat{\beta}^{(b)}(z)$ obtained in the

bth simulation is calculated as follows:

$$\text{RSE}\left(\widehat{\beta}^{(b)}\right) = \frac{\int\left(\widehat{\beta}^{(b)}(z)-\beta(z)\right)f(z)dz}{\int\beta(z)^2 f(z)dz} \approx \frac{\sum_{i=1}^{N_0}\left(\widehat{\beta}^{(b)}(z_i)-\beta(z_i)\right)^2}{\sum_{i=1}^{N_0}\left(\beta(z_i)\right)^2},$$

with $N_0 = 100$ and $z_i = i/N_0$.

For each pair of $\{N, n\}$, the root mean and standard deviation (in parentheses) of RSEs for the estimates: $\{\widehat{\beta}_1^{(b)}(z), \widehat{\beta}_2^{(b)}(z), \widehat{\gamma}_1^{(b)}(z), \widehat{\gamma}_2^{(b)}(z), \widehat{\gamma}_3^{(b)}(z)\}$ over $B = 1000$ iterations are computed and presented in the following table. The RSE reported in the table improve substantially as either N or n increases, but it improves faster when n increases, which confirms the asymptotic results in Proposition 1.

Appendix C Technical derivations and proofs

C.1 Necessary conditions for Assumption 2

Using recursive substitution, Model (6) becomes

$$\begin{aligned} T_{i,t+1} =& \frac{1-\beta_1^{t+1}(z_i)}{1-\beta_1(z_i)}\alpha_i + \sum_{s=0}^{t}\beta_1^s(z_i)\beta_2(z_i)R_{i,t-s} + \sum_{s=0}^{t}\beta_1^s(z_i)\gamma_1(z_i)\overline{T}_{t-s} && \text{(C.1)}\\ &+ \sum_{s=0}^{t}\beta_1^s(z_i)\gamma_2(z_i)\overline{R}_{t-s} + \sum_{s=0}^{t}\beta_1^s(z_i)\gamma_3(z_i)\ln(CO2_t) && \text{(C.2)}\\ &+ \sum_{s=0}^{t}\beta_1^s(z_i)\varepsilon_{i,t+1-s} + \beta_1^{t+1}(z_i)T_{i,0}. && \text{(C.3)} \end{aligned}$$

If Assumption 2 holds, there exists a linear combination of $T_{i,t+1}$,$\overline{T}_t$, $R_{i,t}$, $\overline{R}_t$ and $\ln(CO2_t)$ such that this combination is stationary. Given the above equations (C.1), (C.2), and (C.3), one necessary condition for the existence of such a combination is that the items in line (C.3) are stationary when $t \to \infty$. Hence, the representation of $T_{i,t+1}$ in Assumption 2 requires

$$||\beta_1||_\infty < 1, \tag{C.4}$$

where $||\beta_1||_\infty$ denotes $\max_{z\in\mathbb{Z}}|\beta_1(z)|$ and $\mathbb{Z}$ is a compact set. In the constant coefficient model (1), $\beta_1(z) = \beta_1$, and this condition reduces to $|\beta_1| < 1$.

Averaging (C.1), (C.2), and (C.3) over i and using Assumption 1, we obtain

$$\begin{aligned} &\left(1 - \frac{1}{N}\sum_{i=1}^{N}\sum_{s=0}^{t}\beta_1^s(z_i)\gamma_1(z_i)L^s\right)\overline{T}_{t+1}\\ =&\frac{1}{N}\sum_{i=1}^{N}\frac{1-\beta_1^{t+1}(z_i)}{1-\beta_1(z_i)}\alpha_i \end{aligned} \tag{C.5}$$

$$+\frac{1}{N}\sum_{i=1}^{N}\sum_{s=0}^{t}\left(\beta_1^s(z_i)\left(\beta_2(z_i)+\gamma_2(z_i)\right)\delta_{R0}+\beta_1^s(z_i)\gamma_3(z_i)\delta_{c0}\right) \tag{C.6}$$

$$+\frac{1}{N}\sum_{i=1}^{N}\sum_{s=0}^{t}\beta_1^s(z_i)\gamma_3(z_i)\delta_{c1}\left(t-s\right) \tag{C.7}$$

$$+\frac{1}{N}\sum_{i=1}^{N}\sum_{s=0}^{t}\left(\beta_1^s(z_i)\left(\beta_2(z_i)+\gamma_2(z_i)\right)\overline{\delta}_R^{\top}+\beta_1^s(z_i)\gamma_3(z_i)\delta_c^{\top}\right)\mathcal{U}_{g,t-s} \tag{C.8}$$

$$+\frac{1}{N}\sum_{i=1}^{N}\sum_{s=0}^{t}\left(\beta_1^s(z_i)\beta_2(z_i)+\beta_1^s(z_i)\gamma_2(z_i)\right)\mathcal{U}_{P,i,t} \tag{C.9}$$

$$+\frac{1}{N}\sum_{i=1}^{N}\sum_{s=0}^{t}\beta_1^s(z_i)\varepsilon_{i,t+1-s}+\frac{1}{N}\sum_{i=1}^{N}\sum_{s=0}^{t}\beta_1^s(z_i)\gamma_3(z_i)\mathrm{u}_{c,t-s}+\frac{1}{N}\sum_{i=1}^{N}\beta_1^{t+1}(z_i)T_{i,0}, \tag{C.10}$$

where L denotes the lag operator. If $\overline{T}_{t+1}$ has the representation in Assumption 2, there exists a linear combination of $\overline{T}_{t+1}$, t, and $\mathcal{U}_{g,t}$ such that this combination is stationary. Let $\lambda_{N,t}$ denote the items on line (C.6). One necessary condition for the existence of such combination is that

$$\left(1-\frac{1}{N}\sum_{i=1}^{N}\sum_{s=0}^{t}\beta_1^s(z_i)\gamma_1(z_i)L^s\right)^{-1}\lambda_{N,t}<\infty \tag{C.11}$$

when $(t,N)\to\infty$. We first investigate $\lambda_{N,t}$. Let $g(z)$ denote a bounded function on $\mathbb{Z}$. Since

$$\sum_{s=0}^{t}\beta_1^s(z_i)g(z_i)=\frac{1-\beta_1^{t+1}(z_i)}{1-\beta_1(z_i)}g(z_i),$$

and we have imposed the condition that $||\beta_1||_\infty<1$ in (C.4), we have

$$\frac{1}{N}\sum_{i=1}^{N}\sum_{s=0}^{t}\beta_1^s(z_i)g(z_i)\xrightarrow{p}\mathbb{E}_N\left[\frac{g(z_i)}{1-\beta_1(z_i)}\right]<\infty$$

when $(t,N)\to\infty$. Hence, $\lambda_{N,t}$ converges to a finite constant, and we denoted this constant by λ_0. The necessary condition in (C.11) becomes

$$\left(1-\frac{1}{N}\sum_{i=1}^{N}\sum_{s=0}^{t}\beta_1^s(z_i)\gamma_1(z_i)\right)^{-1}\lambda_0<\infty$$

when $(t,N)\to\infty$, which implies that

$$\left|\mathbb{E}_N\left[\frac{\gamma_1(z_i)}{1-\beta_1(z_i)}\right]\right|<1. \tag{C.12}$$

In the constant coefficient case, this condition becomes $|\beta_1+\gamma_1|<1$. Moreover, the conditions (C.4) and (C.12) are sufficient for obtaining the representations of $\overline{T}_{t+1}$ and $T_{i,t+1}$ in Phillips et al. (2020).

In the varying-coefficient model, additional assumptions are required to represent terms such as $\left(1-\frac{1}{N}\sum_{i=1}^{N}\sum_{s=0}^{t}\beta_1^s(z_i)\gamma_1(z_i)L^s\right)^{-1}\times\left(\frac{1}{N}\sum_{i=1}^{N}\sum_{s=0}^{t}\beta_1^s(z_i)g(z_i)\mathcal{U}_{g,t-s}\right)$ by $\overline{\delta}_T^{\top}\mathcal{U}_{gt}$ and a stationary component. We impose Assumption 2 directly in the paper to avoid the inclusion of additional technicalities.

For the global average radiation, we assume that the local stationary component $R^0_{i,t}$ satisfies the cross-sectional strong LLN: $N^{-1}\sum_{i=1}^N R^0_{i,t} \xrightarrow{a.s.} R^0 := \lim_{N\to\infty}\left\{N^{-1}\sum_{i=1}^N \mathbb{E}\left[R^0_{i,t}\right]\right\}$ for each given t. Then,

$$\overline{R}_t = R^0 + \overline{\delta}_R^\top \mathcal{U}_{g,t} + O_p\left(\sqrt{\frac{n}{N}}\right),$$

since the station-specific idiosyncratic component satisfies

$$\begin{aligned} N^{-1}\sum_{i=1}^N \left(R^0_{i,t} + \mathcal{U}_{P,i,t}\right) &= N^{-1}\sum_{i=1}^N R^0_{i,t} + \sqrt{\frac{n}{N}} \times \frac{1}{\sqrt{N}}\sum_{i=1}^N \left(\frac{1}{\sqrt{n}}\mathcal{U}_{P,i,t}\right) \\ &= R^0 + O_p\left(\sqrt{\frac{n}{N}}\right). \end{aligned}$$

For the representation of $\overline{R}_t$ in Assumption 2 to hold, we require $\frac{n}{N} \to 0$ when $(n, N) \to (\infty, \infty)$. This condition ensures that the number of stations N dominate the sample size in the time direction n asymptotically. In our sample we have $N = 1209$ and $n = 58$, which makes this assumption reasonable for our application.

C.2 Proofs of the main results

To find the asymptotic distribution of $\widehat{\boldsymbol{\beta}}(z)$, we first write

$$\sqrt{Nh} D_n H(z)^\top \left(\widehat{\boldsymbol{\beta}}(z) - \boldsymbol{\beta}(z)\right) = Q^{-1}_{n,Nh}\left(\sqrt{Nh}\left(B_1 + B_2\right)\right), \tag{C.13}$$

where

$$\begin{aligned} Q_{n,Nh} &= \left(\frac{1}{Nh}\sum_{i=1}^N\sum_{t=1}^n K\left(\frac{z_i - z}{h}\right) D_n^{-1} H(z)^\top \widetilde{X}_{i,t}\widetilde{X}_{i,t}^\top H(z) D_n^{-1}\right), \\ B_1 &= \left(\frac{1}{Nh}\sum_{i=1}^N\sum_{t=1}^n K\left(\frac{z_i - z}{h}\right) D_n^{-1} H(z)^\top \widetilde{X}_{i,t}\widetilde{X}_{i,t}^\top \left(\boldsymbol{\beta}(z_i) - \boldsymbol{\beta}(z)\right)\right), \\ B_2 &= \left(\frac{1}{Nh}\sum_{i=1}^N\sum_{t=1}^n K\left(\frac{z_i - z}{h}\right) D_n^{-1} H(z)^\top \widetilde{X}_{i,t}\widetilde{\varepsilon}_{i,t+1}\right). \end{aligned} \tag{C.14}$$

Then Lemma 1 establishes the limit of (C.14), while Theorem 1 establishes the distribution of (C.13). We use sequential limits ($n \to \infty$ followed by $N \to \infty$, denoted by $(n, N)_{seq} \to \infty$), throughout the paper.

Proof of Lemma 1

First, using $z_i = z + vh$, we expand $\delta(z_i)$ and $\delta_{xp}(z_i)$ in (15) in the neighborhood of z: $\delta(z_i) = \delta(z) + \delta'(z)vh + O_p(h^2)$ and $\delta_{xp}(z_i) = \delta_{xp}(z) + \delta'_{xp}(z)vh + O_p(h^2)$. Then,

$$H(z)^\top \widetilde{X}_{i,t} =: W_{1,i}(z) + vhW_{2,i}(z) + h^2 W_{3,i}(z), \tag{C.15}$$

where

$$
\begin{aligned}
W_{1,i}(z) &= H(z)^\top \left(\delta(z)\widetilde{F}_t + \delta_{xp}(z)\widetilde{\mathcal{U}}_{P,it} + \widetilde{\eta}_{x,it}\right) \\
&= \begin{pmatrix} C_1(z)\widetilde{t} + C_2(z)\widetilde{\mathcal{U}}_{gt} + C_4(z)\widetilde{\mathcal{U}}_{P,it} + H_1(z)\widetilde{\eta}_{x,it} \\ C_3(z)\widetilde{\mathcal{U}}_{gt} + C_5(z)\widetilde{\mathcal{U}}_{P,it} + H_2(z)^\top\widetilde{\eta}_{x,it} \end{pmatrix}, \\
W_{2,i}(z) &= H(z)^\top \left(\delta'(z)\tilde{F}_t + \delta'_{xp}(z)\widetilde{\mathcal{U}}_{P,it}\right),
\end{aligned}
$$

and $W_{3,i}(z)$ denotes the remaining terms.

We first establish the limit when $n \to \infty$ while holding N fixed. From Eq. (C.15),

$$
\begin{aligned}
&\sum_{t=1}^{n} D_n^{-1} H(z)^\top \widetilde{X}_{i,t}\widetilde{X}_{i,t}^\top H(z) D_n^{-1} \\
&= \sum_{t=1}^{n} D_n^{-1}\left(W_{1,i}(z)W_{1,i}(z)^\top\right) D_n^{-1} + vh\sum_{t=1}^{n} D_n^{-1}\left(W_{1,i}(z)W_{2,i}(z)^\top + W_{1,i}(z)^\top W_{2,i}(z)\right) D_n^{-1} + O(h^2).
\end{aligned}
$$

Define

$$
S_{1,n,i} = \begin{pmatrix} S_{1,n,i}^{(1,1)} & S_{1,n,i}^{(1,2)} \\ S_{1,n,i}^{(2,1)} & S_{1,n,i}^{(2,2)} \end{pmatrix} = \sum_{t=1}^{n} D_n^{-1}\left(W_{1,i}(z)W_{1,i}(z)^\top\right) D_n^{-1}.
$$

Then

$$
\begin{aligned}
S_{1,n,i}^{(1,1)} =& C_1^2(z) n^{-3}\sum_t \widetilde{t}^2 + O_p(n^{-1/2}) \Rightarrow C_1^2(z)\int_0^1 \tilde{r}^2 =: S_{1,i}^{(1,1)}(z), \\
S_{1,n,i}^{(2,1)} =& C_3(z) n^{-\frac{5}{2}}\sum_t \widetilde{t}\widetilde{U}_{gt} C_1(z) + C_5(z) n^{-\frac{5}{2}}\sum_t \widetilde{t}\widetilde{\mathcal{U}}_{P,it} C_1(z) + O_P(n^{-1/2}) \\
\Rightarrow& C_3(z)\left(\int_0^1 \tilde{r}\widetilde{U}_g\right) C_1(z) + C_5(z)\left(\int_0^1 \tilde{r}\mathcal{U}_{iP}\right) C_1(z) =: S_{1,i}^{(2,1)}(z), \\
S_{1,n,i}^{(2,2)} =& C_3(z) n^{-2}\sum_t \widetilde{U}_{gt}\widetilde{U}_{gt}^\top C_3(z)^\top + C_3(z) n^{-2}\sum_t \widetilde{U}_{gt}\widetilde{\mathcal{U}}_{P,it} C_5(z)^\top \\
&+ C_5(z) n^{-2}\sum_t \widetilde{\mathcal{U}}_{P,it}\widetilde{U}_{gt}^\top C_3(z)^\top + C_5(z) n^{-2}\sum_t \widetilde{\mathcal{U}}_{P,it}^2 C_5(z)^\top \\
\Rightarrow& C_3(z)\left(\int_0^1 \widetilde{U}_g\widetilde{U}_g^\top\right) C_3(z)^\top + C_3(z)\left(\int_0^1 \widetilde{U}_g\widetilde{\mathcal{U}}_{iP}\right) C_5(z)^\top \\
&+ C_5(z)\left(\int_0^1 \widetilde{\mathcal{U}}_{iP}\widetilde{U}_g^\top\right) C_3(z)^\top + C_5(z)\left(\int_0^1 \widetilde{\mathcal{U}}_{iP}^2\right) C_5(z)^\top \\
=:& S_{1,i}^{(2,2)}(z),
\end{aligned}
$$

and $S_{1,i}^{(1,2)}$ is the transpose of $S_{1,n,i}^{(2,1)}$. Assume that

$$
S_{2,i}(z) = \lim_{n\to\infty}\sum_{t=1}^{n} D_n^{-1}\left(W_{1,i}(z)W_{2,i}(z)^\top + W_{1,i}(z)^\top W_{2,i}(z)\right) D_n^{-1} \tag{C.16}
$$

exists and may be singular. Then, holding N fixed, we have

$$\sum_{t=1}^{n} D_n^{-1} H(z)^\top \widetilde{X}_{i,t} \widetilde{X}_{i,t}^\top H(z) D_n^{-1} =_{\mathcal{D}} S_{1,i}(z) + vhS_{2,i}(z) + O_P(h^2) \tag{C.17}$$

as $n \to \infty$ for each i, where the symbol "$=_{\mathcal{D}}$" denotes that both sides are equivalent under weak convergence.

Next, we find the limit of (C.17) when $N \to \infty$. First, from Assumption 3(ii), (iv), (v), and using $f(z_i) = f(z) + f'(z)vh + O(h^2)$, we have

$$\begin{aligned}
\frac{1}{Nh}\sum_{i=1}^{N} K\left(\frac{z_i - z}{h}\right) S_{1,i}^{(1,1)}(z) &= \frac{1}{Nh}\sum_{i=1}^{N} K\left(\frac{z_i - z}{h}\right) C_1^2(z) \int_0^1 \widetilde{r}^2 \\
&= C_1^2(z) \int_0^1 \widetilde{r}^2 \mathbb{E}_N\left(\frac{1}{h} K\left(\frac{z_i - z}{h}\right)\right) + o_{(a.s.)}(1) \\
&= C_1^2(z) \int_0^1 \widetilde{r}^2 \int K(v) \left(f(z) + f'(z)vh + O_P(h^2)\right) dv \\
&= S_1^{(1,1)}(z) f(z) + O_P(h^2).
\end{aligned}$$

Second,

$$\begin{aligned}
&\frac{1}{Nh}\sum_{i=1}^{N} K\left(\frac{z_i - z}{h}\right) S_{1,i}^{(2,1)}(z) \\
=&C_3(z)\left(\int_0^1 \widetilde{r}\widetilde{U}_g\right) C_1(z) \mathbb{E}_N\left(\frac{1}{h} K\left(\frac{z_i - z}{h}\right)\right) + \mathbb{E}_N\left(\frac{1}{h} K\left(\frac{z_i - z}{h}\right) C_5(z)\left(\int_0^1 \widetilde{r}\mathcal{U}_{iP}\right) C_1(z)\right) + o_P(1) \qquad \text{(C.18)} \\
=&C_3(z)\left(\int_0^1 \widetilde{r}\widetilde{U}_g\right) C_1(z) f(z) + O_P(h^2) + \mathbb{E}_N\left(C_5(z)\left(\int_0^1 \widetilde{r}\widetilde{\mathcal{U}}_{iP}\right) C_1(z)\right) \mathbb{E}_N\left(\frac{1}{h} K\left(\frac{z_i - z}{h}\right)\right) \qquad \text{(C.19)} \\
=&S_1^{(2,1)}(z) f(z) + O_P(h^2), \qquad \text{(C.20)}
\end{aligned}$$

where the almost surely convergence in (C.18) comes from the independence of $\int_0^1 \widetilde{r}\mathcal{U}_{iP}$ and z_i across i, Eq. (C.19) comes from the independence between z_i and $\widetilde{\mathcal{U}}_{P,it}$, and Eq. (C.20) comes from $\mathbb{E}_N\left[\int_0^1 \widetilde{r}\widetilde{\mathcal{U}}_{iP}\right] = 0$. Similarly,

$$\begin{aligned}
&\frac{1}{Nh}\sum_{i=1}^{N} K\left(\frac{z_i - z}{h}\right) S_{1,i}^{(2,2)}(z) \\
=&C_3(z)\left(\int_0^1 \widetilde{U}_g \widetilde{U}_g^\top\right) C_3(z)^\top \mathbb{E}_N\left(\frac{1}{h} K\left(\frac{z_i - z}{h}\right)\right) o_P(1) \qquad \text{(C.21)} \\
&+ \mathbb{E}_N\left(\frac{1}{h} K\left(\frac{z_i - z}{h}\right)\right) \mathbb{E}_N\left(C_5(z)\left(\int_0^1 \widetilde{\mathcal{U}}_{iP}^2\right) C_5(z)^\top\right) \qquad \text{(C.22)} \\
=&C_3(z)\left(\int_0^1 \widetilde{r}\widetilde{U}_g\right) C_1(z) f(z) + \frac{1}{6} C_5(z) C_5(z)^\top \sigma_p^2 f(z) + O_P(h^2) \qquad \text{(C.23)} \\
=&S_1^{(2,2)}(z) f(z) + O(h^2), \qquad \text{(C.24)}
\end{aligned}$$

where (C.22) comes from Assumption 3(iii) and from which $\mathbb{E}_N\left(\int_0^1 \widetilde{\mathcal{U}}_{iP}\widetilde{\mathcal{U}}_g^\top\right) = 0$, while Eq. (C.23) comes from $\mathbb{E}_N\left(\int_0^1 \widetilde{\mathcal{U}}_{iP}^2\right) = \mathbb{E}_N\left(\int_0^1 \mathcal{U}_{iP}^2 - \int_0^1 \mathcal{U}_{iP}\int_0^1 \mathcal{U}_{iP}\right) = \frac{1}{6}\sigma_p^2$. Note that $S_{2,i}(z)$ is independent from z_i. Hence, we have

$$\frac{1}{Nh}\sum_{i=1}^{N} K\left(\frac{z_i - z}{h}\right)\left(S_{1,i}(z) + vhS_{2,i}(z) + O(h^2)\right) = f(z)S_1(z) + O_P(h^2) \tag{C.25}$$

when $N \to \infty$. Combining (C.17) and (C.25), we establish $Q_{n,Nh} \Rightarrow f(z)S_1(z)$ when $(n, N)_{seq} \to \infty$ as in Lemma 1.

Proof of Theorem 1

To prove Theorem 1, we first use Eq. (C.13) to find the bias term from B_1. Holding N fixed, we have

$$\begin{aligned}
&\sum_{t=1}^{n} D_n^{-1}H(z)^\top \widetilde{X}_{i,t}\widetilde{X}_{i,t}^\top\left(\boldsymbol{\beta}(z_i) - \boldsymbol{\beta}(z)\right) - \left(S_{1,i}(z) + vhS_{2,i}(z) + O(h^2)\right) D_nH(z)^\top\left(\boldsymbol{\beta}(z_i) - \boldsymbol{\beta}(z)\right)\\
=&\sum_{t=1}^{n} D_n^{-1}H(z)^\top \widetilde{X}_{i,t}\widetilde{X}_{i,t}^\top H(z)D_n^{-1}D_nH(z)^\top\left(\boldsymbol{\beta}(z_i) - \boldsymbol{\beta}(z)\right)\\
&- \left(S_{1,i}(z) + vhS_{2,i}(z) + O(h^2)\right) D_nH(z)^\top\left(\boldsymbol{\beta}(z_i) - \boldsymbol{\beta}(z)\right) \Rightarrow 0
\end{aligned} \tag{C.26}$$

when $n \to \infty$ for each i as in (C.17). Next, using $\boldsymbol{\beta}(z_i) = \boldsymbol{\beta}(z) + \boldsymbol{\beta}'(z)vh + \frac{1}{2}\boldsymbol{\beta}''(z)v^2h^2 + o_P(h^2)$ and similar arguments as (C.25), we have when $N \to \infty$

$$\begin{aligned}
&\frac{1}{Nh}\sum_{i=1}^{N} K\left(\frac{z_i - z}{h}\right)\left(S_{1,i}(z) + vhS_{2,i}(z) + O(h^2)\right) D_nH(z)^\top\left(\boldsymbol{\beta}(z_i) - \boldsymbol{\beta}(z)\right)\\
=&\kappa_2 h^2\left(S_1(z)D_nH(z)^\top\left(\boldsymbol{\beta}'(z)f'(z) + \boldsymbol{\beta}''(z)f(z)\right) + S_2(z)D_nH(z)^\top\boldsymbol{\beta}'(z)f(z) + D_nH(z)^\top o_P(1)\right),
\end{aligned} \tag{C.27}$$

where $S_2(z) := E_N(S_{2,i}(z))$. Combining (C.26) and (C.27), we have

$$\begin{aligned}
Q_{n,Nh}^{-1}B_1 \Rightarrow&\kappa_2 h^2 f^{-1}(z)S_1^{-1}(z)\left(S_1(z)D_nH(z)^\top\left(\boldsymbol{\beta}'(z)f'(z) + \boldsymbol{\beta}''(z)f(z)\right)\right.\\
&\left.+S_2(z)D_nH(z)^\top\boldsymbol{\beta}'(z)f(z) + D_nH(z)^\top o(1)\right),\\
=&D_nH(z)^\top\kappa_2 h^2\left(B_\beta(z) + o_P(1)\right).
\end{aligned} \tag{C.28}$$

Next we find the distribution of $\sqrt{Nh}B_2$. Note that

$$\sqrt{Nh}B_2 = \left(\frac{1}{\sqrt{Nh}}\sum_{i=1}^{N}\sum_{t=1}^{n} K\left(\frac{z_i - z}{h}\right) D_n^{-1}\left(W_{1,i}(z) + vhW_{2,i}(z) + v^2h^2W_{3,i}\right)\widetilde{\varepsilon}_{i,t+1}\right). \tag{C.29}$$

Holding N fixed, as $n \to \infty$, we have

$$\sum_{t=1}^{n} D_n^{-1} W_{1,i}(z)\widetilde{\varepsilon}_{i,t+1} = n^{-1/2} \sum_{t=1}^{n} \begin{pmatrix} C_1(z)\widetilde{\varepsilon}_{i,t+1}\widetilde{t}/n + O_p(\frac{1}{\sqrt{n}}) \\ C_3(z)\frac{\widetilde{\mathcal{U}}_{g,t}}{\sqrt{n}}\widetilde{\varepsilon}_{i,t+1} + C_5(z)\frac{\widetilde{\mathcal{U}}_{P,it}}{\sqrt{n}}\widetilde{\varepsilon}_{i,t+1} + O_p(\frac{1}{\sqrt{n}}) \end{pmatrix}$$
$$\Rightarrow \begin{pmatrix} C_1(z)\xi_{1i} \\ C_3(z)\xi_{2i} + C_5(z)\xi_{3i} \end{pmatrix} := \xi_i, \tag{C.30}$$

where $\xi_{1i} = \int_0^1 \widetilde{r} d\widetilde{B}_{\varepsilon,i}$, $\xi_{2i} = \int_0^1 \widetilde{\mathcal{U}}_g d\widetilde{B}_{\varepsilon,i}$, and $\xi_{3i} = \int_0^1 \widetilde{\mathcal{U}}_{P,i} d\widetilde{B}_{\varepsilon,i}$. From Assumption 3(iii), we have $\mathbb{E}_N[\xi_{ki}] = 0$, for $k = 1, 2, 3$. Furthermore,

$$\begin{aligned}
&E_N\left(\xi_{1i}^2\right) = \sigma_\epsilon^2 \int_0^1 \widetilde{r}^2 dr^2, \\
&E_N\left(\xi_{2i}\xi_{2i}^\top\right) = \sigma_\epsilon^2 \left(\int_0^1 \widetilde{\mathcal{U}}_g \widetilde{\mathcal{U}}_g^\top\right), \\
&E_N\left(\xi_{3i}^2\right) = \frac{1}{6}\sigma_\epsilon^2 \sigma_P^2, \\
&E_N\left(\xi_{1i}\xi_{2i}^\top\right) = \sigma_\epsilon^2 \int_0^1 \widetilde{r}\widetilde{\mathcal{U}}_g, \\
&E_N\left(\xi_{1i}\xi_{3i}^\top\right) = 0,\ E_N\left(\xi_{2i}\xi_{3i}^\top\right) = 0,
\end{aligned}$$

from which we obtain that $\mathbb{E}_N\left(\xi_i\xi_i^\top\right) = \sigma_\epsilon^2 S_1(z)$. Given the independence between ξ_{1i} and z_i in Assumption 3(v), we have

$$\mathbb{E}_N\left[\frac{1}{\sqrt{h}} K\left(\frac{z_i - z}{h}\right)\xi_i\right] = 0$$

and

$$\mathbb{E}_N\left[\left(\frac{1}{h}K^2\left(\frac{z_i - z}{h}\right)\xi_i\xi_i^\top\right)\right] = \mathbb{E}_N\left[\frac{1}{h}\int K^2\left(\frac{z_i - z}{h}\right) f(z_i) dz_i\right] \mathbb{E}_N\left(\xi_i\xi_i^\top\right)$$
$$= f(z)\nu_0 S_1(z)\sigma_\epsilon^2 + O(h^2).$$

Hence

$$\frac{1}{\sqrt{Nh}} \sum_{i=1}^{N} K\left(\frac{z_i - z}{h}\right)\xi_i \Rightarrow MN\left(0, f(z)\nu_0\sigma_\epsilon^2 S_1(z)\right) \tag{C.31}$$

as $N \to \infty$. Combining Eq. (C.30) and (C.31), we have

$$\frac{1}{\sqrt{Nh}} \sum_{i=1}^{N}\sum_{t=1}^{n} K\left(\frac{z_i - z}{h}\right) D_n^{-1} W_{1,i}(z)\widetilde{\varepsilon}_{i,t+1} \Rightarrow MN\left(0, f(z)\nu_0\sigma_\epsilon^2 S_1(z)\right). \tag{C.32}$$

The others terms in Eq. (C.29) are of smaller order. Combining Eq. (C.28), (C.32) and Lemma 1, we

establish Theorem (1) as:

$$\sqrt{Nh}D_nH(z)^\top\left(\widehat{\boldsymbol{\beta}}(z)-\boldsymbol{\beta}(z)-\kappa_2h^2B_\beta(z)\right)\Rightarrow f^{-1}(z)S_1^{-1}(z)MN\left(0,f(z)\nu_0\sigma_\epsilon^2S_1(z)\right)$$
$$=MN\left(0,f^{-1}(z)v_0S_1^{-1}(z)\right).$$

Proof of Proposition 1

From Theorem 1, we have

$$\sqrt{Nh}\begin{pmatrix}n^{3/2}H_1(z)^\top\\ nH_2(z)^\top\end{pmatrix}\left(\widehat{\boldsymbol{\beta}}(z)-\boldsymbol{\beta}(z)-\kappa_2h^2B_\beta(z)\right)$$
$$\Rightarrow MN\left(0,V(z)\right)=MN\left(0,\begin{pmatrix}V^{(1,1)}(z) & V^{(1,2)}(z)\\ V^{(2,1)}(z) & V^{(2,2)}(z)\end{pmatrix}\right).$$

Then

$$\begin{aligned}&\sqrt{Nhn}\left(\widehat{\boldsymbol{\beta}}(z)-\boldsymbol{\beta}(z)-B_\beta(z)\right)\\ =&\sqrt{Nhn}H(z)H(z)^\top\left(\widehat{\boldsymbol{\beta}}(z)-\boldsymbol{\beta}(z)-B_\beta(z)\right)\\ =&\sqrt{Nhn}H_1(z)H_1(z)^\top\left(\widehat{\boldsymbol{\beta}}(z)-\boldsymbol{\beta}(z)-B_\beta(z)\right)+\sqrt{Nhn}H_2(z)H_2(z)^\top\left(\widehat{\boldsymbol{\beta}}(z)-\boldsymbol{\beta}(z)-B_\beta(z)\right)\\ =&\sqrt{Nhn}H_2(z)H_2(z)^\top\left(\widehat{\boldsymbol{\beta}}(z)-\boldsymbol{\beta}(z)-B_\beta(z)\right)+o_p(1)\\ \Rightarrow&MN\left(0,H_2(z)V^{(2,2)}(z)H_2(z)^\top\right).\end{aligned}$$

Appendix D Imputation of missing data

D.1 Data description

GEBA reports an annual mean solar radiation observation for a station as a missing value if any of the relevant monthly radiation observations for that station and year are not available. Overall there are 43,660 missing annual mean radiation values for the 1,209 stations that we consider, and they are unevenly distributed over the period 1959–2017, as illustrated in Figure 10. Figure 11 provides a map of the 1,209 stations.

Magnus et al. (2011) treated data gaps in the GEBA solar radiation dataset as missing completely at random when they modelled changes in temperature. Specifically, they showed that their estimated model of first differenced temperature in their unbalanced panel over 1959–2002 was statistically the same as that

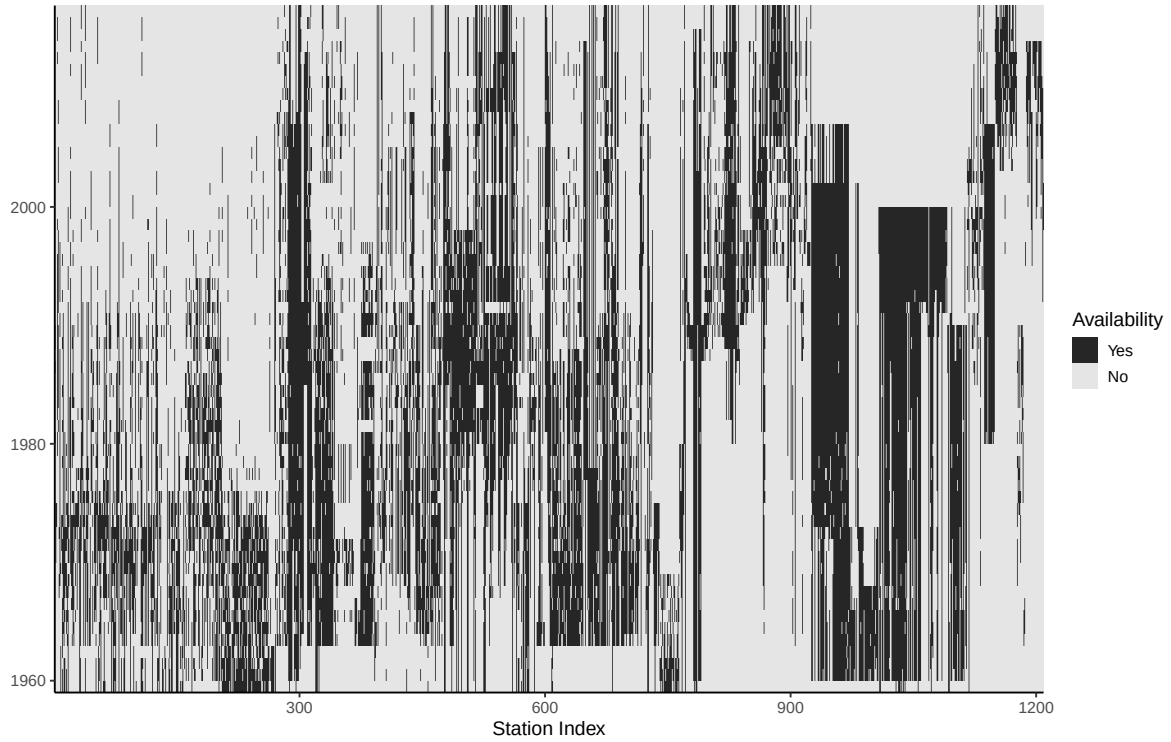

Figure 10: *Availability of annual mean solar radiation for the selected* 1,209 *GEBA stations, with available data highlighted in black. The station index numbers shown in the figure are obtained by rearranging the selected stations according to their original GEBA Station ID in an ascending order.*

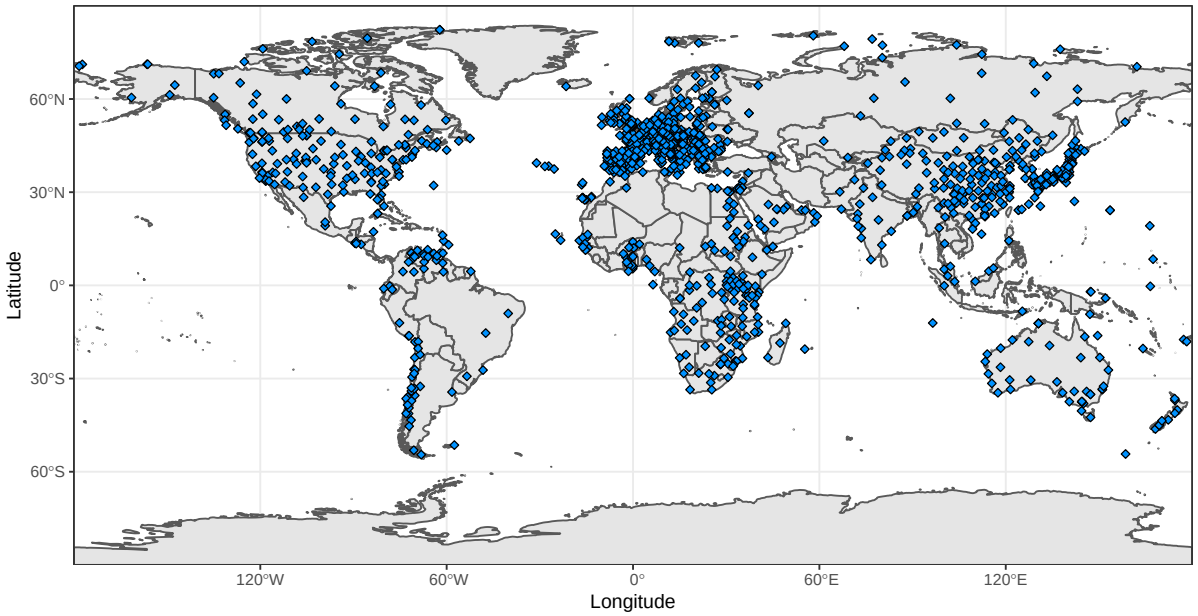

Figure 11: *Distribution of the 1,209 GEBA weather stations.*

for a balanced panel consisting of the subset of weather stations with continuous solar radiation records over the 43 years. Storelvmo et al. (2016) treated data gaps in their radiation data from GEBA data in the same way when studying temperature changes over the period from 1964 to 2010.

Phillips et al. (2020) filled in the data gaps in GEBA solar radiation data using the CRU climate data set and a "random forest" machine learning approach. This was possible because CRU records data on many climate variables (other than surface temperature) for the same stations over the period of interest. The implementation of random forest techniques is complex and involves choice with respect to input variables and algorithms (see, e.g., Breiman, 2001). Here, we use CRU data (and also radiation data from neighbouring stations) to impute missing radiation data, but we base our imputation technique on panel data models that can be more easily replicated. The following three subsections provide details relating to the ancillary data and imputation methodology that we use to complete our solar radiation data.

CRU climate variables

The CRU dataset contains ten climate variables. These include three temperature related variables (diurnal temperature range, and maximum and minimum temperatures) in addition to the mean near-surface temperature variable that we model, and six other variables. The latter variables (i.e., precipitation rate within a month (PRE), vapor pressure (VAP), wet days within a month (WET), cloud cover percentage (CLD), frost days within a month (FRS), and potential evapotranspiration within a day (PET)) are correlated with solar radiation level to various extents, and we use these to predict our missing annual mean solar radiation observations. We do not include the three other (temperature related) variables in our imputation models because they are directly related to the surface temperature in (6) and their inclusion might ultimately distort the interpretation of (6).

Information from neighbouring stations

A common strategy for filling in gaps in climate data is to use information from neighboring stations and spatial interpolation algorithms (Mitchell and Jones, 2005). In practice, the selection of a set of neighboring stations from which to construct a reference series can be a challenging problem. Given the sparse data coverage of GEBA observations, performing neighbor-based homogeneity tests on solar radiation variable is complicated and is not completely possible. Data from spatially close stations are more likely to be correlated than more-distant stations, and following Mitchell and Jones (2005) we select neighbors within a threshold distance from any considered station. Mitchell and Jones (2005) chose a threshold of $1,200$ km for the mean temperature in CRU database. Given that there are less available GEBA stations than CRU weather stations, we use a slightly larger distance threshold of $1,500$ km. Geodesic distances between the considered GEBA stations are computed using longitude and latitude coordinates of stations with reference to the World Geodetic System 1984 using the R package "geosphere" (Hijmans, 2019). Only 2 out of the $1,209$ GEBA stations that we consider do not have close neighbors according to the above distance-based requirements.

Method for filling in missing radiation data

Our algorithm for filling in missing values in the annual mean solar radiation time series $R_{i,t}$ uses the following steps:

1. **Construct time series $R_{i,t}^N$ for each of the i stations.** These time series are functions of climate data as well as observed values of $R_{i,t}$ from neighbouring stations.

   We let $J_{i,t}$ count the number of station i's neighbouring stations that have radiation observations available in year t ($J_{i,t}$ will differ across stations and across t), and let $s_{i,t} = \{1, 2, \cdots J_{i,t}\}$ be the set containing these neighbouring stations. For each station, we classify the set of $n-1$ years into two station-specific sets, $\mathcal{T}_{i,A}^N$ and $\mathcal{T}_{i,M}^N$, such that $t_i \in \mathcal{T}_{i,A}^N$ if $J_{i,t} \geq 1$ (at least one neighbouring station has a radiation observation in year t) and $t_i \in \mathcal{T}_{i,M}^N$ if $J_{i,t} = 0$ (no neighbouring station has a radiation observation in year t). These two sets are mutually exclusive, and $\mathcal{T}_{i,A}^N \cup \mathcal{T}_{i,M}^N = \{1, \cdots, n-1\}$.

   For each $i = 1, \ldots, N$ and $t_i \in \mathcal{T}_{i,A}^N$, we define $R_{i,t_i}^N = J_{i,t}^{-1} \sum_{j \in s_{i,t}} R_{j,t_i}$. We then fit a fixed-effect panel data model to R_{i,t_i}^N using the specification

$$R_{i,t_i}^N = \alpha_i^N + \boldsymbol{x}_{i,t_i}^{N\top} \boldsymbol{\beta}^N + e_{i,t_i}, \quad \text{for} \quad t_i \in \mathcal{T}_{i,A}^N, \tag{D.1}$$

   where $\boldsymbol{x}_{i,t}^N$ contains CRU climate variables consisting of $CLD_{i,t}^N$, $FRS_{i,t}^N$, $PET_{i,t}^N$, $PRE_{i,t}^N$ and $VAP_{i,t}^N$, that are statistically significant explanators at the 5% level. The superscript N on these climate variables indicates that their construction corresponds with that for R_{i,t_i}^N, but with the set $\mathcal{T}_{i,A}^N$ remaining the same as that for R_{i,t_i}^N.

   Model (D.1) is fitted using the R package "plm" (Croissant and Millo, 2008). We then use the estimated $\widehat{\boldsymbol{\beta}}^N$ and $\widehat{\alpha}_i^N$, to predict the missing values in $R_{i,t}^N$ as

$$\widehat{R}_{i,t_i}^N = \widehat{\alpha}_i^N + \boldsymbol{x}_{i,t_i}^{N\top} \widehat{\boldsymbol{\beta}}^N, \quad \text{for} \quad t_i \in \mathcal{T}_{i,M}^N. \tag{D.2}$$

   Equation (D.2) allows us to complete the construction of time series for $R_{i,t}^N$, by setting $R_{i,t_i}^N = \widehat{R}_{i,t_i}^N$ for $t_i \in \mathcal{T}_{i,M}^N$.

2. **Fill in the missing values of $R_{i,t}$ time series.**

   We first consider station i and classify the set of $n-1$ years into two mutually exclusive sets, $\mathcal{T}_{i,A}$ and $\mathcal{T}_{i,M}$, such that $t_i \in \mathcal{T}_{i,A}$ if $R_{i,t}$ is an available observation and $t_i \in \mathcal{T}_{i,M}$ if the observation for $R_{i,t}$ is missing. These two sets are complementary, so $\mathcal{T}_{i,A} \cup \mathcal{T}_{i,M} = \{1, \cdots, n-1\}$. We then use the $R_{i,t}^N$ values and another panel based two step procedure to fill in the missing $R_{i,t}$ data. Specifically, we estimate

$$R_{i,t_i} = \alpha_i + \boldsymbol{u}_{i,t_i}^\top \boldsymbol{\beta} + e_{i,t_i}, \quad t_i \in \mathcal{T}_{i,A}, \tag{D.3}$$

where $\boldsymbol{u}_{i,t} = \left(R^N_{i,t-1},\ R^N_{i,t},\ \text{PRE}_{i,t},\ \text{VAP}_{i,t},\ \text{WET}_{i,t}\right)^\top$. We use the original climate variables in (D.3) rather than the neighbourhood versions used in (D.1), and we find that the three climate variables $\text{PRE}_{i,t}$, $\text{VAP}_{i,t}$ and $\text{VAP}_{i,t}$ are statistically significant at the 5% level in (D.3), but $\text{CLD}_{i,t}$ is not. The CLD variable is derived from sunshine hours (see Harris et al., 2014, for more details), and its absence from (D.3) suggests that the constructed explanatory variable $R^N_{i,t}$ in (D.3) has already adequately accounted for cloud coverage information. We use the estimated $\widehat{\boldsymbol{\beta}}$ and $\widehat{\alpha}_i$ from (D.3), to predict the missing values in $R_{i,t}$ as

$$\widehat{R}_{i,t_i} = \widehat{\alpha}_i + \boldsymbol{u}^\top_{i,t_i}\widehat{\boldsymbol{\beta}}, \quad t_i \in \mathcal{T}_{i,M}, \tag{D.4}$$

and then complete the construction of time series for $R_{i,t}$, by setting $R_{i,t_i} = \widehat{R}_{i,t_i}$ for $t_i \in \mathcal{T}^N_{i,M}$.

Evaluation of constructed variables

We evaluate our constructed variables in a qualitative way by comparing our regional averages of imputed data with those that we calculated before we filled in missing values. Further, we compare them with the series in Phillips et al. (2020), that were constructed using different methodology and were based on different samples.

Figure 12 compares spatially averaged annual solar radiation data for each continent drawn from the completed panels (in blue), with those based on available (but incomplete) GEBA data (in black). The blue curves retain the general patterns of the raw data (i.e., a decreasing trend until the mid 1990s before slowly increasing for the northern hemisphere versus a gradually increasing tendency in the southern hemisphere), although changes are less abrupt.

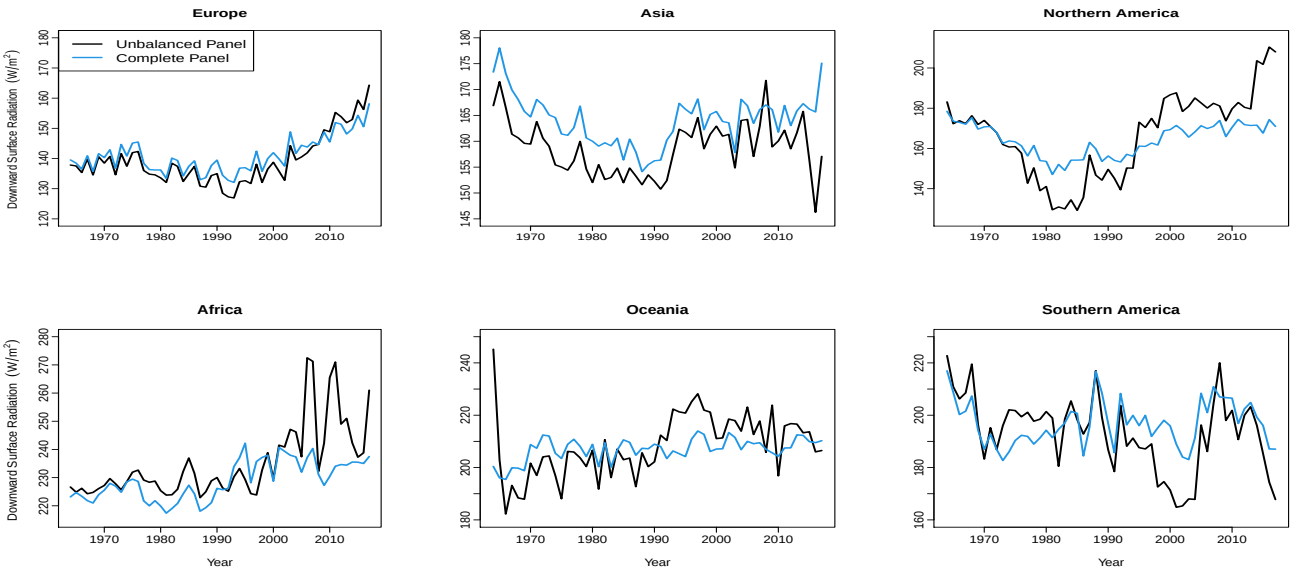

Figure 12: *Average of annual solar radiation by continents, 1959–2017.*

The U-shape seen in the top three panels illustrates the radiation "dimming" and "brightening" mentioned in Section 4.1.3. Extreme weather conditions probably explain why some weather stations have only a few years with complete data (Magnus et al., 2011). When local extreme weather conditions impact a cluster of closely located stations for a short period in a year, regional averages of annual solar radiation are computed without covering the entire area, and this can lead to surprisingly high or low values due to uneven coverage of land surface in a continent. The unusually high (black) spikes in average values for Africa and Oceania in Figure 12 might be caused by stations with low solar radiation intensity that is not measured because of weather events, and filling in the missing values can mitigate the impact of extreme weather conditions on solar radiation records, as shown in the blue plots based on completed panels. In contrast, many European weather stations record continuous observations as shown in Figure 11, and the difference between the averages of the imputed and raw data in Europe are small compared to other continents. This suggests that filling in missing values has minimal impact on high quality data.

Completing the panel by filling in the missing radiation data, allows us to use all temperature data when estimating our model, and we compare plots of spatially averaged temperatures based on all stations with those based on the subset of stations for which radiation data is available in Figure 13.

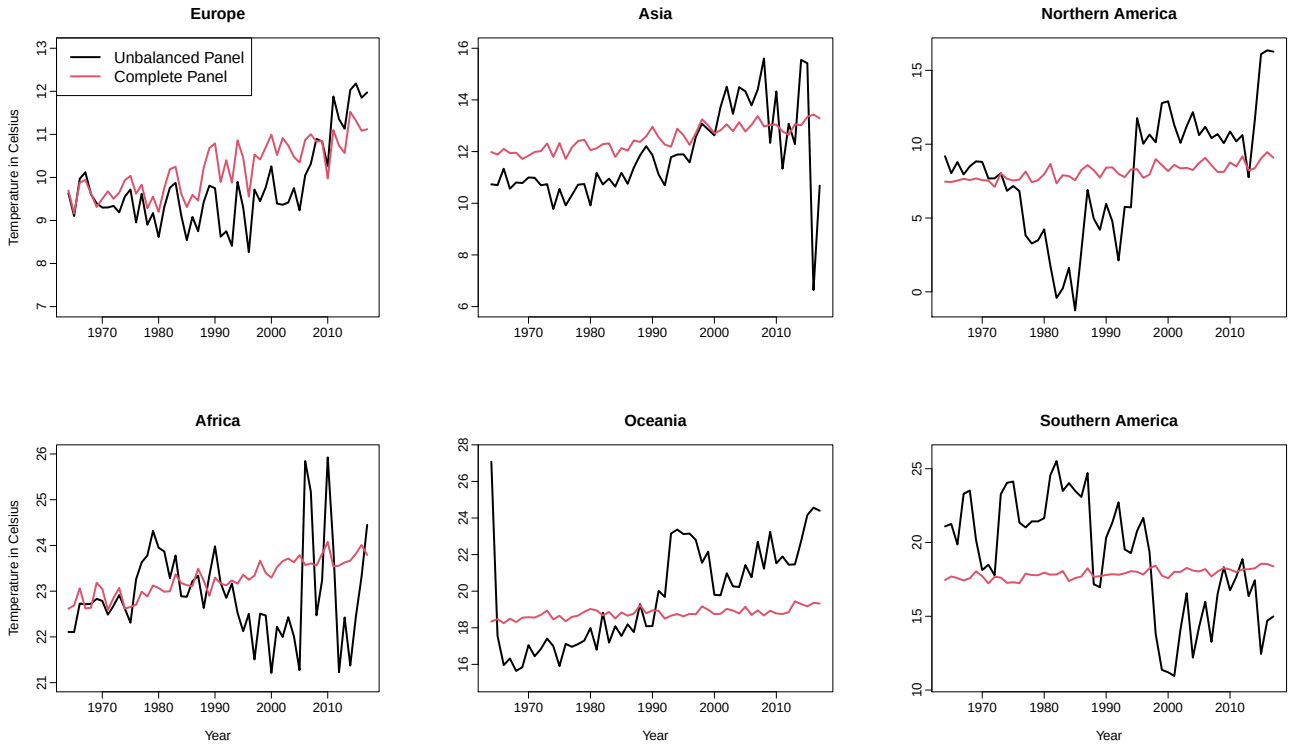

Figure 13: *Average of annual temperature by continents, 1959–2017.*

The black curves in Figure 13, display considerable fluctuations, which are partially due to variation in the number and locations of stations that report measurements over time. The red curves in Figure 13 show that the averages of annual temperature computed from the complete panel gradually increase over the 59 years with modest variations, which is closely aligned with estimates produced by IPCC (IPCC,

2018). Compared with the raw observations with missing values, the complete panel with less volatile average annual temperature curves are more appropriate for our analysis. The plots based on complete panels indicate that spatially averaged temperatures in each of the six continents have risen since 1959.

We also compare our completed data series with those used in Phillips et al. (2020). The latter series use temperature data from a different source and the associated analysis is based on more stations (1484) over a shorter time span (1964 - 2005). Further, these authors used random forest algorithms to fill in missing data values. Their resulting final series are not publicly available, but Figure 1 in Phillips et al. (2020) plots the global average of each of their annual time series. Our plots of our versions of these time series shown in our Figure 1 are very similar over 1964-2005, with respect to the patterns in the series and their numerical values. Hence, we conclude that our missing data imputation method based on panel regressions delivers series that are in line with those used by others in climate sensitivity analyses.

Appendix E Results based on local linear estimators

In this section, we present additional empirical results using the local linear estimator. Figures 14 and 15 can be compared with Figures 3 and 4 in the text, and Figures 16 and 17 can be compared with Figures 5 and 6. In all four cases the results are very similar to those obtained using the local constant estimators.

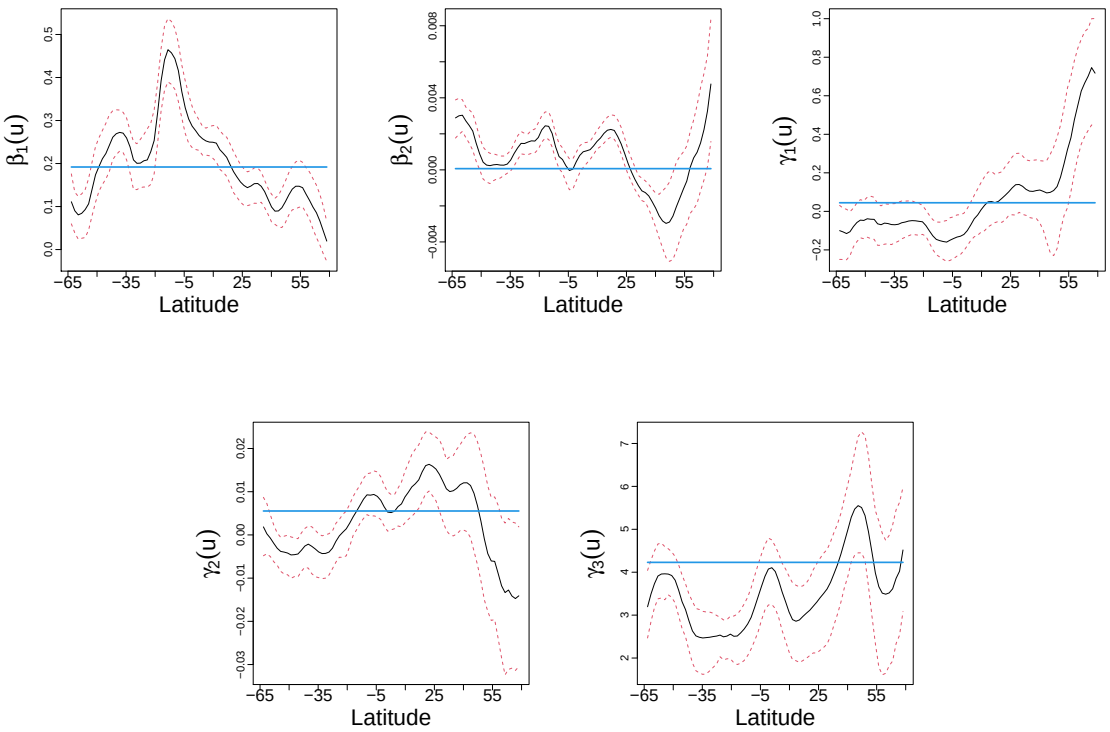

Figure 14: *Estimated coefficient functions of latitude (solid black) together with their 95% confidence intervals (dotted orange) obtained by bootstrap. Within group estimates (solid blue) are superimposed on the estimated functions.*

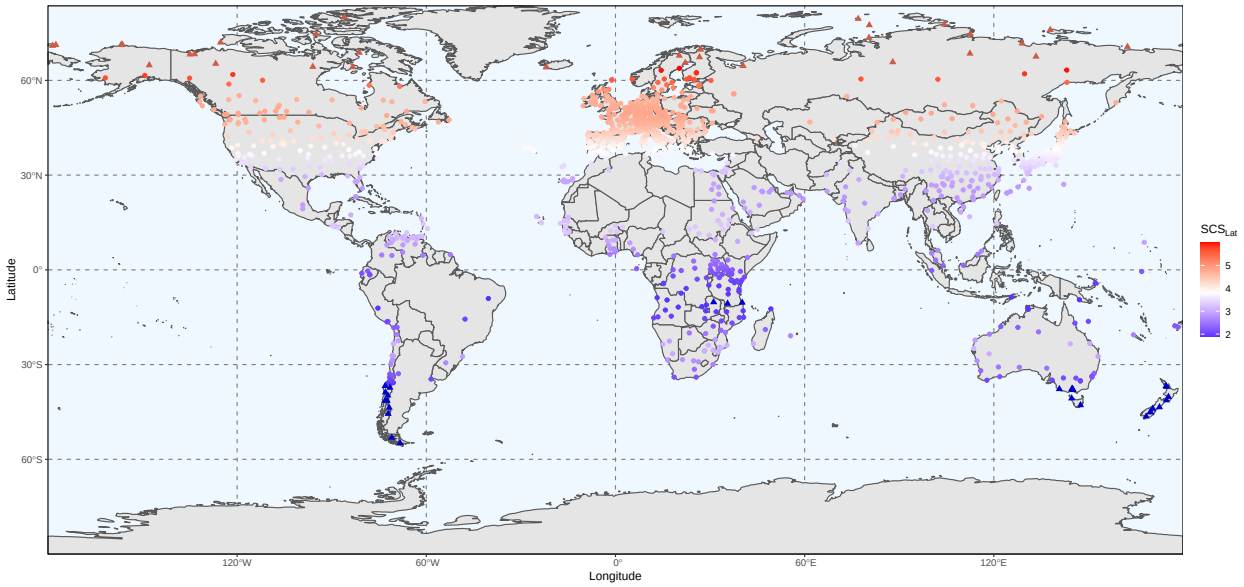

Figure 15: *Local climate sensitivity based on estimated coefficient functions that vary over latitude. Color indicates the LCS of weather stations with white indicating those stations for which LCS $\approx$ GCS. Stations with extreme climate sensitivities ($>$ 97.5% or $<$ 2.5% of LCS_{lat}) are highlighted with triangular symbols.*

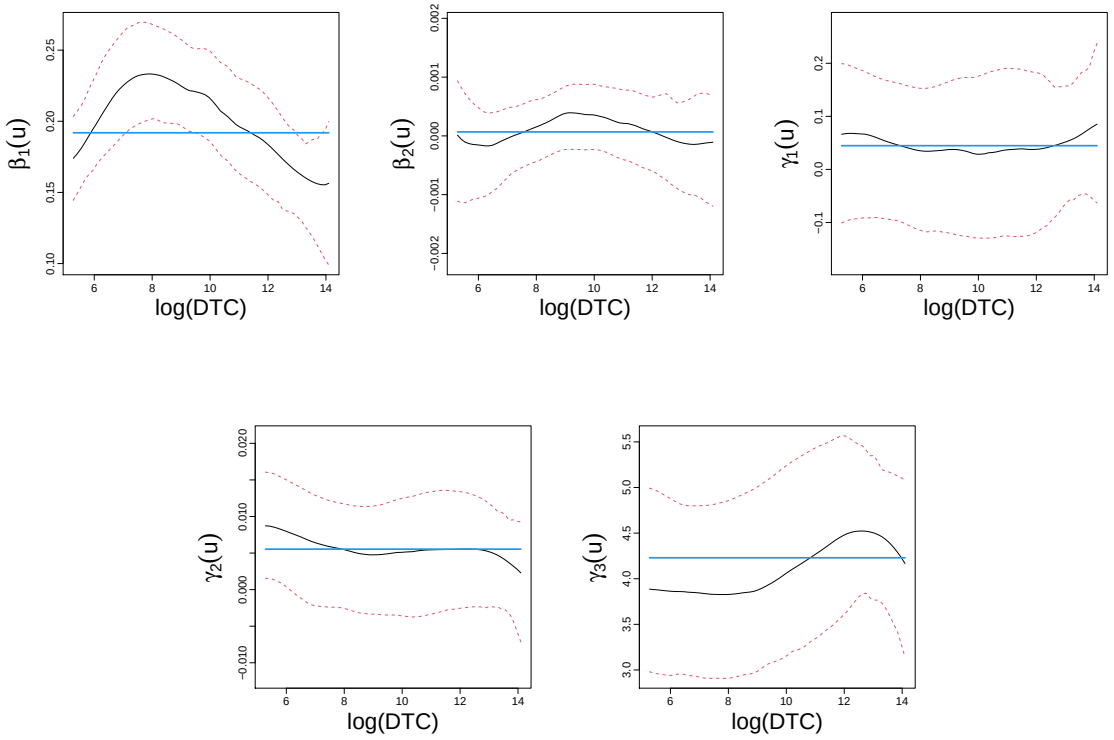

Figure 16: *Estimated coefficient functions of (the natural logarithms of) distance to the coast (DTC) (solid black) with their 95% confidence intervals (dotted orange) obtained by bootstrap. Within group estimates (solid blue) are superimposed on the estimated functions.*

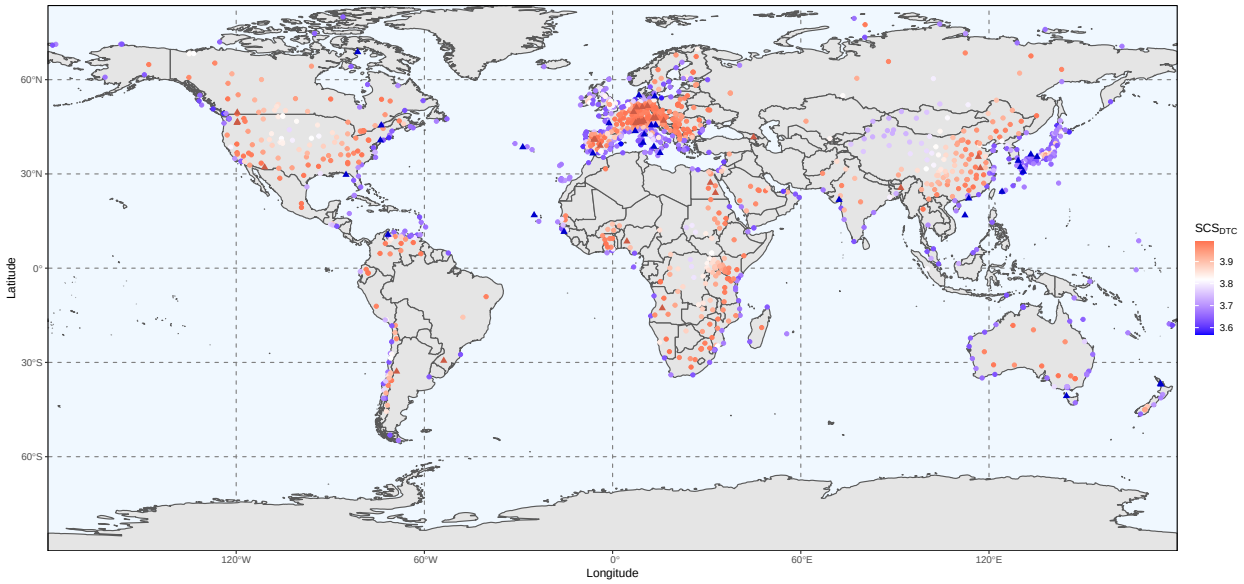

Figure 17: *Local climate sensitivity based on estimated coefficient functions that vary with (the natural logarithm of) distance to the coastline (DTC). Color indicates the LCS of weather stations, with white indicating those stations for which LCS $\approx$ GCS. Stations with extreme climate sensitivities ($>$ 97.5% or $<$ 2.5% of LCS_{DTC}) are highlighted with triangular symbols.*