

## Article

# On Spurious Causality, CO<sub>2</sub>, and Global Temperature

Philippe Goulet Coulombe <sup>1,\*</sup> and Maximilian Göbel <sup>2</sup><sup>1</sup> Département des Sciences Économiques, Université du Québec à Montréal, Montréal, QC H2L 2C4, Canada<sup>2</sup> Lisbon School of Economics and Management, Universidade de Lisboa, 1200-781 Lisboa, Portugal; maximilian.goebel@phd.iseg.ulisboa.pt

\* Correspondence: p.gouletcoulombe@gmail.com or goulet\_coulombe.philippe@uqam.ca

**Abstract:** Stips et al. (2016) use information flows (Liang (2008, 2014)) to establish causality from various forcings to global temperature. We show that the formulas being used hinge on a simplifying assumption that is nearly always rejected by the data. We propose the well-known forecast error variance decomposition based on a Vector Autoregression as an adequate measure of information flow, and find that most results in Stips et al. (2016) cannot be corroborated. Then, we discuss which modeling choices (e.g., the choice of CO<sub>2</sub> series and assumptions about simultaneous relationships) may help in extracting credible estimates of causal flows and the transient climate response simply by looking at the joint dynamics of two climatic time series.

**Keywords:** information flows; vector autoregressions; global warming; climate econometrics



**Citation:** Goulet Coulombe, Philippe, and Maximilian Göbel. 2021. On Spurious Causality, CO<sub>2</sub>, and Global Temperature. *Econometrics* 9: 33. <https://doi.org/10.3390/econometrics9030033>

Academic Editor: Claudio Morana

Received: 30 July 2021

Accepted: 26 August 2021

Published: 7 September 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Causality is fundamental to science. While causal statements can reasonably be made without hardship in controlled environments, things are far less straightforward when only observational data are available. Answering the question of what happens to  $Y$  if one intervenes on  $X$  is compromised by the simple fact that  $X$  and  $Y$  were not generated by exogenously modulating  $X$ , but (often) by endogenous interactions between the two variables.

Yet, for many scientific interrogations of capital importance (like the relationship between greenhouse gases and global temperature), only observational data are available. To what extent do different forcings cause global mean surface temperature anomalies (GMTA) Stips et al. (2016)? Stips, Macias, Coughlan, Garcia-Gorriz, and Liang (2016) (henceforth, SMCGGL) set up to answer the question using *information flow* (IF) from one variable to another for bivariate stochastic dynamical systems—a methodology developed in Liang (2008, 2014, 2015, 2016). The authors make grand claims about the technique being able to extract “true” and “rigorous” causality, which clashes with the usual somber tone of causal analysis.

We show that the IF methodology will *not* extract causality for the near-universe of bivariate stochastic dynamic systems estimated on real data—with those of SMCGGL included. In a nutshell, this occurs because the Liang (2008) formulas that are used throughout assume that a certain statistical quantity is zero when it is not. This can be tested and it is rejected almost all the time. More precisely, the Liang (2014) formula assumes that when conditioning on the past state of the system ( $X_{t-1}$  and  $Y_{t-1}$ ), the remaining variation in  $X_t$  and  $Y_t$  (the innovations driving the system) are *uncorrelated*. In SMCGGL, where the time unit  $t$  is one year, this implies that forcings and GMTA are uncorrelated *within the same year*, conditional on the last year's values of both. That is unlikely to happen, and the data will testify to that. The assumption is presented as “linearity” (Liang 2014), but it has little to do with it. Rather, it assumes that the system is identified, meaning that the joint dynamics of the time series data fall on the knife edge case where causality can indeed be claimed from the data without any external assumption (Sims 1980). The problem

is that for discretely-sampled time series, the assumption is almost always rejected by lower-frequency data.<sup>1</sup> In short, SMCGL assume there to be no identification problem (in the sense of Sims (1980)), and find no identification problem.

In this note, we detail the consequences of this high-standing omission on IF measures and SMCGL's results. First, in simulations (using data-generating processes proposed by Liang (2014)) of which we know the true causality structure, it is shown that IF will often conclude that  $X_t$  is largely causing  $Y_t$ , when in fact the reverse is true. Second, we reconsider a key part of SMCGL's study where the authors investigate the causal structure between different forcings and GMTA. Using an appropriate methodology that accounts for the correlated innovations, it is found that, in most instances, the data by themselves cannot back SMCGL's claims. That is, unlike what the authors have put forward, it is not possible (within this framework) to claim that many forcings are causing GMTA's increase as a direct implication of the data. In other words, from the data *alone*, it is impossible to discriminate between certain forcings mostly causing GMTA or the reverse. External assumptions based on physical knowledge could remedy that, or different data. We show that, in the case of CO<sub>2</sub>, results in accordance with the scientific consensus can be obtained when using concentration directly rather than its radiative content.

The alternative methodology that we use is structural Vector Autoregressions (SVAR), which are simultaneous dynamic systems of equations. They can characterize a linear dynamic system in discrete time. The methodology was introduced to macroeconomics by Sims (1980) and has since then entered various fields, ranging from neuroscience (Chen et al. 2011) to climate (Goulet Coulombe and Göbel 2021). To document its reliability for this application, we report implied transient climate response (TCR) estimates of our alternative methodology. We find that those (i) largely depend on the necessary assumptions made about the simultaneous (with one year) impact of CO<sub>2</sub> forcing on GMTA but (ii) are in the range of recent estimates (Montamat and Stock 2020; Otto et al. 2013) if one assumes simultaneous causality running from CO<sub>2</sub> forcing to GMTA.

The remainder of this article is structured as follows: Section 2 briefly reviews IFs and explains where the problematic assumption occurs. It also discusses relevant notions of Vector Autoregressions (VARs, Sims 1980) as a comprehensive framework to think about causality in multivariate time series systems. Section 3 displays IFs possibly spurious behavior using simulated data where the true causality is known. Section 4 revisits the question of causality between different forcings and GMTA using appropriate tools. Section 5 concludes the paper.

## 2. Information Flows, Non-Innocuous Assumptions and VARs

In this section, we review the basics of IFs as applied in empirical work, pin down the problematic assumption, explain why it is harmful through the lenses of a VAR, and propose an alternative measure of IF based on the VAR. Liang (2008) initiate the derivation of IF by considering the following data generating process (DGP):

$$d\mathbf{X} = \mathbf{F}(\mathbf{X}, t)dt + \mathbf{B}(\mathbf{x}, t)d\mathbf{W}, \quad (1)$$

where  $\mathbf{X} = (X_1, X_2) \in \mathbb{R}^2$  are the state variables and  $\mathbf{F} = (F_1, F_2)$ .  $\mathbf{W} = (W_1, W_2)$  is a standard Wiener process, with  $\Delta\mathbf{W} = \frac{d\mathbf{W}}{dt}$  and  $E(\Delta W_i) = 0$  and  $E(\Delta W_i)^2 = \Delta t$ .  $\mathbf{B}$  is a  $2 \times 2$  matrix and its entries  $b_{ij}$  govern how perturbations instantaneously impact the system. At this point, the only non-innocuous assumption is that of a bivariate system. The information flow from  $X_2$  to  $X_1$ ,  $T_{2 \rightarrow 1}$ , is then defined as

$$T_{2 \rightarrow 1} = \frac{dH_1}{dt} - \frac{dH_{1|2}}{dt} \quad (2)$$

where  $\frac{dH_1}{dt}$  is the evolution of the marginal entropy of  $X_1$  and  $\frac{dH_{1|2}}{dt}$  denotes  $\frac{dH_1}{dt}$  where the spillovers from  $X_2$  are excluded. After lengthy derivations and reasonable assumptions,

such as Euler–Bernstein approximations, Taylor-series expansions or that the components of  $\mathbf{W}$ —which are soon to be called *structural shocks*—are uncorrelated, it is obtained that

$$T_{j \rightarrow i} = -E \left[ \frac{1}{\rho_i} \frac{\partial (F_i \rho_i)}{\partial X_i} \right] + \frac{1}{2} E \left[ \frac{1}{\rho_i} \frac{\partial^2 (g_{ii} \rho_i)}{\partial X_i^2} \right], \quad (3)$$

where  $\rho$  is the joint probability density of variables  $X_i$  and  $X_j$ , and  $\rho_i$  denotes the marginal density of series  $X_i$ . Given that the model in (1) is not readily identified from the data, this is not yet operational. After further derivations and assuming  $\mathbf{B} = \begin{bmatrix} b_{ii} & 0 \\ 0 & b_{jj} \end{bmatrix}$ , Liang (2014) transforms (3) into a workable formula made of empirical moments

$$T_{j \rightarrow i} = \frac{\sigma_{i,i} \sigma_{i,j} \sigma_{j,\Delta i} - (\sigma_{i,j})^2 \sigma_{i,\Delta i}}{(\sigma_{i,i})^2 \sigma_{j,j} - \sigma_{i,i} (\sigma_{i,j})^2} \quad (4)$$

where  $\sigma_{i,i}$  is the variance of  $i$ ,  $\sigma_{i,j}$  the covariance between  $i$  and  $j$ , and  $\sigma_{i,\Delta i}$  is the covariance between  $i$  and the  $k^{\text{th}}$  difference of  $i$  (we follow Liang (2015) and set  $k = 1$ ). However, the validity of this appealing formula rests on the seemingly technical assumption of a diagonal  $\mathbf{B}$ . Our point is that this assumption is far from merely technical and very frequently wrong. Motivating the diagonal  $\mathbf{B}$  among other assumptions, Liang (2014) states

*Since the dynamics are unknown, we first need to choose a model. As always, a linear model is the natural choice, at least at the initial stage of development.*

The problem is that (i) of course, we must choose an empirical model, but we will try to avoid those that the data blatantly reject, (ii)  $\mathbf{B}$  being diagonal has nothing to do with linearity, and (iii) there is no further potential “development” possible without this assumption, which, in effect, assumes the causal problem away. As a result, whenever the diagonal  $\mathbf{B}$  is violated by the data, IFs—as currently used in empirical studies—provide spurious causality.

We now use a very popular framework from macroeconometrics to think more clearly about  $\mathbf{B}$ . Liang (2008)’s flow of simplifications and assumptions make his once sophisticated (1) collapse to that of a bivariate VAR with one lag

$$\begin{bmatrix} X_{1,t} \\ X_{2,t} \end{bmatrix} = \underbrace{\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}}_c + \underbrace{\begin{bmatrix} a_{11}^1 & a_{12}^1 \\ a_{21}^1 & a_{22}^1 \end{bmatrix}}_A \begin{bmatrix} X_{1,t-1} \\ X_{2,t-1} \end{bmatrix} + \underbrace{\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}}_B \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}. \quad (5)$$

The seemingly innocuous assumption is much less so within a statistical framework relating assumptions directly to observable quantities. Indeed, the uncorrelatedness of  $W$ ’s, which here translates to that of  $\varepsilon$ , combined with  $b_{21} = b_{12} = 0$ , have a very stark implication. Let the number of endogenous variables be  $M$  and the number of lags  $P$ . Provided imposing  $M = 2$  and  $P = 1$  is reasonable, (4) is valid if and only if regression residuals  $\begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$  are *not* cross-correlated. This is easily testable: one needs to estimate Equations (1) and (2) separately by least squares, collect the residuals, and calculate their correlation  $\rho_u$ . If the latter is different from zero (and this could be formally tested with a t-test), then Liang (2008)’s simple formula does not apply. While  $\rho_u = 0$  might be plausible in continuous time (or anything near it), this is a monumental stretch for data sampled at the yearly frequency (like in SMCG). Fortunately, unlike the true structural causality, the data can directly inform us on whether Liang (2014)’s formula is valid or not for specific pairs of time series.

## 2.1. Acknowledging the Identification Problem: Vector Autoregressions

This exposition closely follows [Goulet Coulombe and Göbel \(2021\)](#). In time series analysis, the "identification" problem originates from simultaneity in the data. We can learn whether

$$X_{t-1} \rightarrow Y_t \quad \text{or} \quad Y_{t-1} \rightarrow X_t$$

is more plausible. This is predictive causality in the sense of [Granger \(1969\)](#). However, the data themselves cannot discriminate between

$$X_t \rightarrow Y_t \quad \text{and} \quad X_t \leftarrow Y_t.$$

In essence, a correlation between  $X_t$  and  $Y_t$  can be generated by two different causal structures. [Liang \(2008\)](#)'s solution is to assume such relationships do not exist—yet, they do. Within a VAR, the problem boils down to the need for identifying  $C$  in

$$C\mathbf{y}_t = \Psi_0 + \sum_{p=1}^P \Psi_p \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t, \quad (6)$$

where  $\mathbf{y}_t$  is an  $M$  by one vector—meaning the dynamic system incorporates  $M$  variables.  $\Psi_p$ 's parameterizes how each of these variables is predicted by its own lags and lags of the  $M - 1$  remaining variables.  $P$  is the number of lags being included. The matrix  $C$  characterizes how the  $M$  different variables interact contemporaneously—e.g., how total forcing affects GMTA within the *same* year (a time unit  $t$  in SMCG). Finally, the *structural* disturbances are mutually *uncorrelated* with mean zero:

$$\boldsymbol{\varepsilon}_t = [\varepsilon_{1,t}, \dots, \varepsilon_{M,t}] \sim N(0, I_M).$$

Equation (6) is the so-called structural form of the VAR, which *cannot* be estimated because  $C$  is not identified by the data. SMCG uses formula (4) that implicitly assumes a constrained version of (6) with  $M = 2$ ,  $P = 1$ , and, most importantly,  $C$  being a diagonal matrix. The validity of their analysis hinges upon those constraints not being rejected by the data. In Section 4, we find that the data disagrees with at least two of them.

Equation (6) is a structural model that can be used to answer causal questions directly. However, the elements of  $C$  are not plain regression coefficients and cannot be estimated as such—they would be biased. It does not mean that they do not exist. The implications of their existence can best be understood by looking at an estimable "reduced-form" VAR

$$\mathbf{y}_t = \mathbf{c} + \sum_{p=1}^P \Phi_p \mathbf{y}_{t-p} + \mathbf{u}_t, \quad (7)$$

where  $\mathbf{c} = C^{-1}\Psi_0$  and  $\Phi_p = C^{-1}\Psi_p$  are both regression coefficients obtained by running least squares separately on each equation.  $\mathbf{u}_t$  are now regression residuals

$$\mathbf{u}_t = [u_{1,t}, \dots, u_{M,t}] \sim N(0, \Sigma_u)$$

which will be cross-correlated if the true  $C$  is *not* diagonal. As mentioned earlier, [Liang \(2014\)](#)'s simplifying assumptions translate into  $\Sigma_u = C^{-1'}C^{-1}$  being diagonal, which is often at odds with the data. All of the parameters of (7) can be estimated with traditional methods, but the model is not "structural" and cannot be used for causal inference. Structural VARs, which aim at uncovering "structural" causality (instead of predictive causality à la [Granger 1969](#)) acknowledge  $\Sigma_u$  being non-diagonal and provide ways to obtain  $C$ . As a by-product, they can procure valid measures of information flow.

The raw material of causal measures are  $\boldsymbol{\varepsilon}_t$ , the structural disturbances entering the systems. However, those like structural causality are not directly extractable from the data: we only have  $\mathbf{u}_t$  and translating those back to  $\boldsymbol{\varepsilon}_t$  necessitates  $C$ . The latter is not directly

attainable, but can be retrieved using the mapping  $\Sigma_u = C^{-1'}C^{-1}$ . In words, this means that the covariance matrix of regression residuals from (7) can be used as raw material to retrieve the “structural”  $C$ . Mechanically, the identification problem emerges because there are many  $C$ ’s satisfying  $\hat{\Sigma}_u = C^{-1'}C^{-1}$ —numerous causal structures deliver the same residuals’ cross-correlations.

The strategy we opt for is the traditional Choleski decomposition of  $\hat{\Sigma}_u$ . This is one of many identification strategies for the VAR (Kilian and Lütkepohl 2017). However, among the catalog of methodologies, the Choleski decomposition is certainly popular (if not the most popular) in applied work and is simple to implement. Mechanically, it provides a lower-triangular matrix  $C$ , satisfying  $\hat{\Sigma}_u = \mathbf{B}'\mathbf{B}$  where  $\mathbf{B}$  is the same as from Equation (5), but with dimensions  $M \times M$ . Its purpose is to transform cross-correlated regression residuals  $u_t$  (Equation (7)) into uncorrelated structural shocks  $\varepsilon_t$  (Equation (6)). This is done by reversing the relationship  $u_t = C\varepsilon_t$ .

Uncorrelatedness is essential to study how GMTA responds to a given forcing, *keeping everything else constant*. Such a causal claim would be impossible when considering an impulse from correlated residuals  $u_t$  as those always co-move. A Choleski decomposition of  $\Sigma_u$  is one way to transform the observed  $u_t$  into the unobserved fundamental shocks  $\varepsilon_t$ . The assumption underlying such an approach to orthogonalization is a causal ordering of shocks. The ordering restricts how variables interact with each other *within the same year*, conditional on the previous state of the system. In our bivariate setup, ordering forcing  $j$  after GMTA implies that forcing cannot impact GMTA within the same year. Ordering GMTA after a given forcing implies that GMTA cannot impact forcing  $j$  within the same year. SMCGL implicitly assumes both restrictions at the same time, which results in a rejected over-identified model. In contrast, when the model is just identified (when only one restriction is imposed), this choice cannot be validated by the data itself as it does not alter the likelihood.

When revisiting SMCGL’s empirical work, we consider both orderings and document how sensitive conclusions are to that necessary choice.<sup>2</sup> There are cases where the sign of net IF (a qualitative notion) between  $j$  and GMTA depends on the ordering choice, and cases where it does not. For instance, we will find that whatever is assumed about  $C$  in a proper VAR system, total forcing appears to be causing GMTA much more than the reverse. In other cases, like CO<sub>2</sub>-induced radiative forcing, this cannot be simply ruled out by the data.

## 2.2. An Adequate Measure of IF Based on the VAR

In a complete multivariate system like a VAR, the errors of the  $h$ -steps ahead forecast  $y_{t+h,m}$  can be related back to structural shocks—that is, the anomalies driving the dynamics of the system. For instance, we can compute the share of the forecast error variance of GMTA 10 years from now that is attributable to CO<sub>2</sub> anomalies. Intuitively, if CO<sub>2</sub> is causing GMTA to increase, its exogenous impulses should be an important driver of GMTA’s variance rather than GMTA anomalies themselves—a high information flow/transfer between the two variables. Accordingly, VAR forecast errors are

$$u_{t+h} = y_{t+h} - \hat{y}_{t+h} = \sum_{h'=0}^{h-1} \Theta_{h'} \varepsilon_{t+h-h'}$$

where  $\Theta_{h'}$  is a function of  $\Phi_p$ ’s and  $C$ . See Kilian and Lütkepohl (2017) for further details. The forecast error variance decomposition (FEVD) of the whole system at horizon  $h$  can be analytically calculated using the entries of matrix  $\Theta_{h'}$ . Precisely,

$$\text{MSPE}_h = E(u_{t+h} u_{t+h}') = \sum_{h'=0}^{h-1} \Theta_{h'} \Theta_{h'}'.$$



An information flow “share” of  $i$  for  $j$  can be characterized by the share of forecast error variance of variable  $j$  attributable to structural shocks of  $i$ .<sup>3</sup> While those measures can be assessed for any horizon  $h$  (which can contain useful information), we focus on the cumulative sum since it is a measure of total flow. For non-stationary VARs (like those of the empirical section), we use a horizon of  $h = 15$  years. In the case of stationary VAR process, FEVD measures quickly converge to their long-run value as  $h$  increases. Moreover, in that case, our FEVD-based IF measures can be more directly motivated from the Wold representation of a VAR process (Kilian and Lütkepohl 2017). Finally, note that whenever the  $b_{12} = b_{21} = 0$  assumption is approximately true, the FEVD approach and IFs give qualitatively similar answers.

### 3. Simulations

This section showcases how IF can lead to a pretense of causal knowledge, with conclusions that are sometimes the exact opposite of the truth. We consider four data-generating processes (DGPs) where the true  $P$  is one. The first two correspond to what is proposed in Liang (2014). The last two are generic VAR positively-autocorrelated processes with differing degrees of persistence. Following the notation above (Equation (5)):

$$DGP(1) := \left\{ A = \begin{bmatrix} 0.5 & 0.5 \\ 0 & 0.6 \end{bmatrix}, \quad c = \begin{bmatrix} 0.1 \\ 0.7 \end{bmatrix} \right\} \quad (8a)$$

$$DGP(2) := \left\{ A = \begin{bmatrix} -0.5 & 0.9 \\ -0.2 & 0.5 \end{bmatrix}, \quad c = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\} \quad (8b)$$

$$DGP(3) := \left\{ A = \begin{bmatrix} 0.5 & -0.2 \\ -0.5 & 0.25 \end{bmatrix}, \quad c = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\} \quad (8c)$$

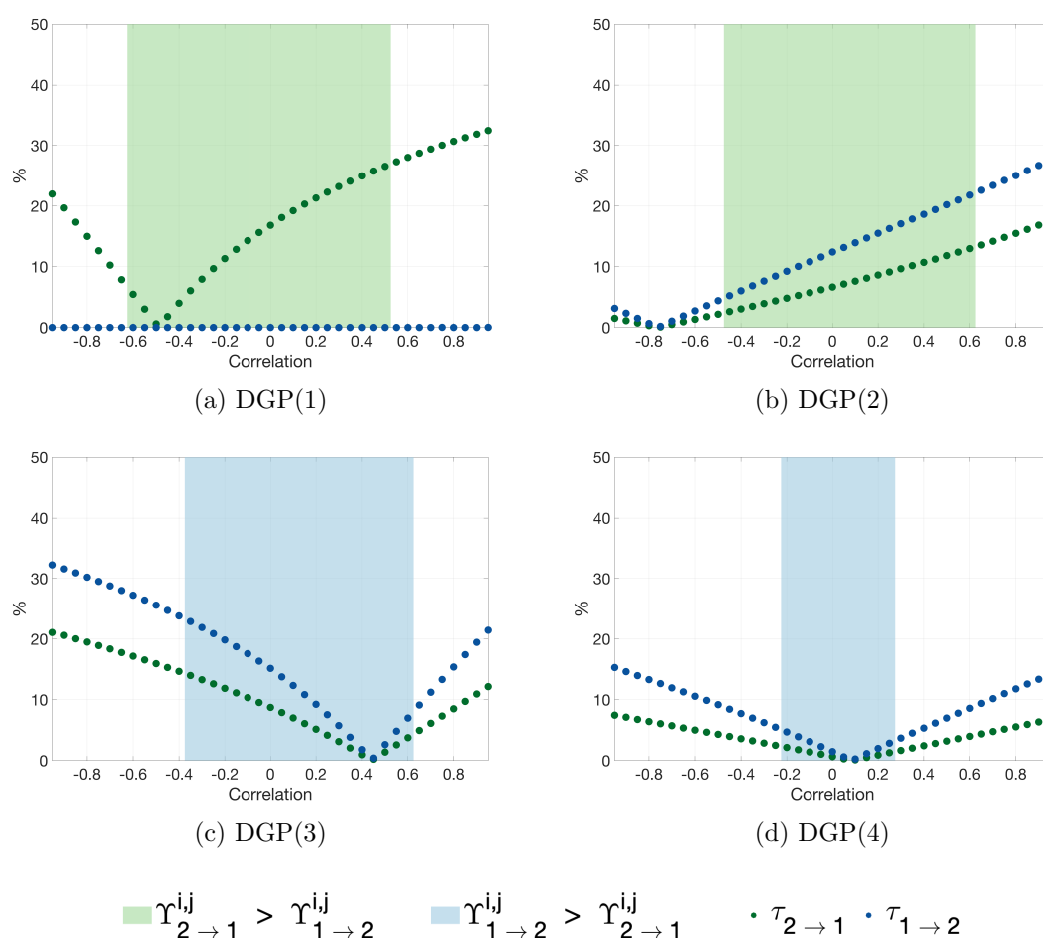
$$DGP(4) := \left\{ A = \begin{bmatrix} 0.25 & -0.1 \\ -0.2 & 0.1 \end{bmatrix}, \quad c = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\} \quad (8d)$$

As highlighted in the previous section, IFs are calculated assuming  $b_{ij} = 0$ . However, as will be reported and formally tested in Section 4, this is frequently not the case for most time series, especially those sampled at low frequencies. Using a controlled simulation environment, we study how IFs behave for values of  $\rho_u := \text{corr}(u_t^{X_1}, u_t^{X_2}) \in [-1, 1]$ .<sup>4</sup> Note that IFs are invariant to  $\rho_u \neq 0$  emerging from  $b_{12} = 0$  or  $b_{21} = 0$ , which is obviously problematic given the causal content of  $b_{ij}$ .<sup>5</sup>

As our indicator of true underlying information flow, taking into account  $\rho_u \neq 0$ , we use the FEVD-based measure described earlier. Note that here, unlike the application to real data in Section 4, we know what are the true  $b_{ji}$ 's. When  $\rho_u \neq 0$ , the correlation must be attributed to either  $b_{12}$  or  $b_{21}$ , or a combination of both. In a bivariate setup, this amounts to setting  $b_{ij} = 0$  and attributing  $\rho \neq 0$  to  $b_{ji}$ . Hence, it is possible to tell when standard IFs conclude falsehoods, because in simulations  $b_{12}$  and  $b_{21}$  are known.

In terms of notation,  $Y_{i \rightarrow j}^{i,j}$  means the FEVD share at horizon  $h = 10$  with  $i = 1, 2$  and  $i \neq j$ . The superscripts in  $Y_{i \rightarrow j}^{i,j}$  determine the true ordering, hence  $i$  ordered before  $j$ . The subscripts indicate that we plot the contribution of variable  $i$  to the forecast error variance of variable  $j$  at horizon  $h = 10$ . The simulations have shown that  $h = 10$  is sufficient for convergence.

Figure 1 plots the absolute normalized information flows (NIF),  $\tau_{i \rightarrow j}$  for DGP(1) through DGP(4), with varying  $\rho_u$ . A first observation is that  $\tau_{i \rightarrow j}$  often varies significantly with  $\rho_u$ , even when the very formula underlying it assumes  $\rho_u = 0$ . The relative strength of  $\tau_{1 \rightarrow 2}$  and  $\tau_{2 \rightarrow 1}$  can easily collapse to 0 or be much higher than what IFs report for  $\rho_u = 0$ .



**Figure 1.** Ranking of NIFs ( $\tau_{i \rightarrow j}$ ) vs. Ranking of FEVDs ( $Y_{i \rightarrow j}^{i,j}$ ); Different Levels of Correlation; Horizon  $h = 10$ . **Notes:** Color specification: the *light blue* area shows the region in which FEVDs suggest  $Y_{1 \rightarrow 2}^{i,j} > Y_{2 \rightarrow 1}^{i,j}$ , regardless of the ordering. Similarly, the *light green* area shows the region in which FEVDs suggest  $Y_{2 \rightarrow 1}^{i,j} > Y_{1 \rightarrow 2}^{i,j}$ —independent of the ordering. The *white/non-colored* areas show those regions for which the ordering of  $X_1$  and  $X_2$  does matter. An unambiguous determination of  $Y_{1 \rightarrow 2}^{i,j} > Y_{2 \rightarrow 1}^{i,j}$ , respectively,  $Y_{2 \rightarrow 1}^{i,j} > Y_{1 \rightarrow 2}^{i,j}$ , is not possible.

The fact that IFs vary with  $\rho_u$  could give false hopes that they account for simultaneous relationships. We conduct a simple exercise to show they do not. An interesting empirical question is whether  $i$  is causing  $j$  more than  $j$  is causing  $i$ , which is the appealing promise of IFs. This moves beyond testing for Granger Causality and aims at quantifying the flow of information. The shaded green region corresponds to the values of  $\rho_u$  for which either the true underlying causality is that  $X_2$  causes  $X_1$  more than the reverse *irrespective* of whether it emerges from either  $b_{12} = 0$  or  $b_{21} = 0$ . For those values of  $\rho_u$ , when the true ordering is unknown, the *qualitative* conclusion about the sign of the net causality flow does not hinge on knowledge of  $b_{ij}$ . Analogously, the blue shade represents values of  $\rho_u$  for which  $X_1$  causing more  $X_2$  is unanimous among  $b_{ij}$  configurations. White regions are values of  $\rho_u$  where conclusions about the sign of net causality flow cannot be determined from the data, i.e., the unknown  $b_{ij}$  is necessary to settle. In other words, given the data available to the modeler,  $X_2$  causing more  $X_1$  and vice versa are both equally likely, and sorting it out decisively implies making a successful guess on the true value of  $b_{ij}$ .

If IFs were correctly calibrated, the green dotted line should only be above the blue in the shaded green regions, and the dotted blue line above the green one in blue regions. Figure 1 clearly demonstrates this not to be the case. For every DGP, there is a substantial range of values (white spaces) for which IFs clearly conclude  $\tau_{i \rightarrow j} > \tau_{j \rightarrow i}$  when there is one pair of  $(b_{ij}, b_{ji})$  out of two for which the reverse is true. To address concerns about a horizon

mismatch  $h$  between IFs and FEVD, Figure A1 (Appendix B) shows results for  $h = 2$ . As it turns out, things worsen for IFs (the blue and green regions shrink) since the simultaneity problem is less diluted in dynamics at short horizons.

In sum, those simulations (largely inspired by Liang (2015) DGPs) show two things. First, the tractable IFs (from formula (4)) are functions of  $\rho_u$  even though they assume it to be zero. This compromises any statement on the strength of causal links. Second, for any DGP, there is an underlying pair  $(b_{ij}, b_{ji})$  for which the IFs' conclusion about the net causal flow is the opposite of reality. This is not only an artifact of large  $|\rho_u|$ , as exemplified by DGP(4).

#### 4. GMTA and Radiative Forcing Revisited

In this section, we first revisit SMCGL's application of IFs to GMTA and forcings. Second, we look at what lies behind opaque FEVD measurements by reporting impulse response functions and computing the transient climate response implied by simplistic bivariate VARs.

Global warming generated by man-made forcing is the prevalent generalization of the notion *climate change*. Numerous researchers have dedicated their works to this very relationship between anthropogenic forcing and constituents of global climate (Andrews et al. 2010; Hansen et al. 2006; Li et al. 2013; Notz and Stroeve 2016) and the socio-economic consequences thereof (Nordhaus 2014; Shukla et al. 2019). Despite overwhelming evidence for anthropogenic forcing being the main driver of global climate change (Masson-Delmotte et al. 2018), scientists have also observed that especially since the turn of the millennium, global temperature has plateaued despite ever-rising greenhouse gases and contrary to projections from key climate models.<sup>6</sup>

IFs hinge exactly on these bivariate relationships, which are attractive in their clarity, but are certainly a stark oversimplification of a complex system. Nevertheless, to give empirical content to our critique of the methodology, we study the relationship between GMTA and seven forcings from SMCGL using both IFs and our FEVD-based remedy. The sample of annual means ranges from 1850 to 2005.<sup>7</sup> We follow SMCGL and take our data from therein referenced data providers. Table 1 summarizes the results.<sup>8</sup>

**Table 1.** Empirical Results for the Bivariate Relationship—Between Various Forcings and GMTA.

	Correlation	Normalized IF (IF $\times$ 100)		Lags ( $P$ )	Correlation of Residuals ( $\rho_u$ )	FEVD			
		$i \rightarrow$ GMTA	GMTA $\rightarrow i$			Ordering: $i, \text{GMTA}$		Ordering: GMTA, $i$	
						$i \rightarrow$ GMTA	GMTA $\rightarrow i$	$i \rightarrow$ GMTA	GMTA $\rightarrow i$
Total Forcing	0.73	30.6 (15.3)	20.8 (11.1)	4 1	0.23 *** 0.29 ***	47.4 51.4	13.0 9.6	28.0 27.6	25.4 28.7
Anthropogenic	0.86	39.8 (35.7)	−20.0 (−0.6)	4 1	−0.19 ** −0.19 **	6.5 5.0	3.7 5.8	3.8 2.2	13.4 17.1
CO <sub>2</sub> -ERF (W/m <sup>2</sup> ) SMCGL	0.86	39.6 (35.1)	−15.2 (−0.4)	4 1	−0.14 * −0.15 *	6.5 2.8	8.4 4.7	5.6 1.1	17.4 12.8
Aerosol	−0.82	35.9 (24.3)	−24.5 (−0.4)	4 1	−0.19 ** −0.10	2.9 3.5	0.6 4.0	2.1 1.8	1.6 1.2
Solar	0.49	13.5 (3.8)	6.7 (2.3)	8 1	0.05 0.08	8.5 16.6	1.6 4.2	6.6 12.3	2.5 6.8
Volcanic	0.09	0.9 (0.2)	−0.5 (−0.4)	4 1	0.18 ** 0.20 **	10.9 7.1	0.8 1.4	3.1 0.6	3.6 3.7
PDO	0.17	−1.2 (−0.2)	−0.6 (−0.5)	4 1	0.35 *** 0.34 ***	31.1 9.1	0.9 0.5	6.3 0.2	10.3 10.7
CO <sub>2</sub> (Mt/yr)	0.82	37.1 (27.0)	−4.3 (−0.0)	2 1	−0.10 −0.05	8.9 4.2	2.1 0.0	10.7 4.4	0.6 0.4
CO <sub>2</sub> (W/m <sup>2</sup> )	0.86	39.5 (34.5)	−14.0 (−0.3)	4 1	0.23 *** 0.07	5.2 1.6	16.8 4.1	2.9 0.9	4.7 1.8

**Notes:**  $i$  corresponds to the type of radiative forcing, listed in the left most column. The second column ("Correlation") documents the correlation between GMTA and variable  $i$ . FEVD values are taken at horizon  $h = 15$ , which translates into the contribution of variable  $i$  in the variance of the forecast error of variable  $j$  a decade and a half after the in-sample end date. Numbers in bold underline the highest absolute causal flow among a  $(i, \text{GMTA})$  pair for a given measure. "\*\*\*", "\*\*", and "\*" means that the null of the residuals cross-correlation of residuals is rejected at the 10%, 5%, and 1% level, respectively.

In Table 1, we report the estimated correlation between residuals of the bivariate VAR(1) implied by IFs. When forcing  $P = 1$ , in five cases out of seven, the null that



$\hat{\rho}_u^{i,GMTA} = 0$  is rejected at least at the 10% level. When choosing  $P$  with Bayesian Information Criterion (BIC), only  $\hat{\rho}_u^{Solar,GMTA} = 0$  cannot be rejected. Hence, as repeatedly mentioned in the text, IFs assume something that can be and is rejected by the data. Moreover, in the light of simulations carried out earlier, the qualitative and quantitative insights from IFs are often spurious under such conditions.

Naturally, we concentrate on FEVD results, which correctly account for  $\rho_u \neq 0$ . However, setting  $\rho_u = 0$  is only one of the empirical shortcomings of the empirical IF formula—it also sets  $P$ , the number of lags of each  $X_t$  as one. Clearly, it is empirically plausible that  $X_{i,t-2}$  or  $X_{j,t-4}$  may have an impact on  $X_{i,t}$  beyond what is channeled by single year lags ( $X_{i,t-1}$  and  $X_{j,t-1}$ ). In other words,  $P = 1$  is extremely restrictive on climatic dynamics. When choosing  $P$  with BIC, results align better with prior scientific knowledge. Nonetheless, for the sake of completeness, we report both results ( $P = 1$  and  $P = P^*$ , with  $P^*$  being BIC's choice).

With  $P = P^*$ , the fact that total forcing causes GMTA more than the reverse is without appeal. Nevertheless, the quantitative answer is, again, highly dependent on the ordering choice. After 15 years, total forcing anomalies are responsible for explaining between 28.0% and 47.4% of that of GMTA—depending on the preferred ordering. The net causal flow being higher from aerosol and solar to GMTA is also unanimous, but much smaller. Indecisive results are reported for Anthropogenic, Volcanic, and PDO. Overall, Table 1 suggests that the data themselves do not support the strong qualitative conclusions of SMCGL for their CO<sub>2</sub> measure and Anthropogenic.

When  $P$  is forced to one, as in IFs, inconclusive results are reported for total forcing, aerosol, volcanic, and PDO.  $P = 1$  specifications, irrespective of the ordering<sup>9</sup>, it can be concluded that GMTA is causing more CO<sub>2</sub> and Anthropogenic than the reverse. Only solar forcing results are unanimous and in line with what climatic common wisdom suggests. Note that it also agrees with results from original IFs, which is not surprising given that  $\hat{\rho}_{Solar,GMTA}$  is in the close vicinity of zero. However, choosing a proper  $P$  nearly doubles the share of forecast errors attributable to solar forcing.

#### 4.1. What's Up with CO<sub>2</sub>?

Results for CO<sub>2</sub> are rather surprising. Irrespective of the ordering, GMTA is reported to cause more CO<sub>2</sub> than the reverse, a finding contradicting SMCGL's results and common wisdom.<sup>10</sup> However, SMCGL's CO<sub>2</sub> measure exhibits ill dynamic behavior that cannot possibly be that of a natural quantity. This becomes obvious when plotting their CO<sub>2</sub>-ERF measure in first difference: it evolves according to a series of dichotomous jumps.<sup>11</sup> This CO<sub>2</sub>-ERF series, in which ERF stands for effective radiative forcing, is a concept presented in Myhre et al. (2013). The series itself is based on Etminan et al. (2016).

Given the strange jumping behavior of the CO<sub>2</sub>-ERF series, and to make our calculations comparable to other findings in the literature (Bruns et al. 2020; Montamat and Stock 2020), we derive RF from CO<sub>2</sub> concentration as follows: we use the well-established Meinshausen et al. (2017) data set on annual global means of CO<sub>2</sub> concentration, measured in parts per million (*ppm*). We follow Myhre et al. (2013) and transform the increase in CO<sub>2</sub> concentration in year  $t$  measured in *ppm* relative to the concentration in a given base year,  $CO_{2,base}$ , into radiative forcing,  $RF_t^{CO_2}$ —measured in  $W/m^2$ —as follows:  $RF_t^{CO_2} = 5.35 \times \ln(CO_{2,t}/CO_{2,base})$ . We use 1850 as the base year following Bruns et al. (2020).

As it turns out, considering this less contentious CO<sub>2</sub> series does not resolve the apparently counterintuitive finding that GMTA explains a larger portion of the forecast error variance of CO<sub>2</sub> than vice versa. Such a finding has also been reported using a different methodology in Koutsoyiannis and Kundzewicz (2020). We explore a last avenue, that of using annual CO<sub>2</sub> emissions rather than  $RF_t^{CO_2}$ . This last attempt is successful in reconciling the FEVD approach with the traditional wisdom that CO<sub>2</sub> is causing GMTA “more” than the reverse. This finding is independent of the ordering choice.

In sum, *based on this particular time series evidence*, the causal link between certain forcings and GMTA remains disputable. What are less disputable are the effects of total forcing, CO<sub>2</sub> emissions, and solar forcing, which all explain an important share of GMTA anomalies independently of arbitrary ordering preferences. Nonetheless, we see such analyses as rather primitive and potentially misleading. For instance, X causing “more” Y does not mean that the reverse causality is not quantitatively important, or climatologically relevant. We now turn to a more promising way to extract meaning out of selected bivariate VARs.

#### 4.2. Impulse Response Functions

To open the black box of those rather abstract measurements, we report in Figure 3 impulse response functions (IRF) for the bivariate models of CO<sub>2</sub>-GMTA and total forcing-GMTA. Since Sims (1980), the dominant approach for studying the properties of the VAR around its deterministic path has been IRFs to *structural shocks*. Their dynamic effect can be analyzed as that of a randomly assigned treatment because those have been transformed to be uncorrelated, which provides the “keeping everything else constant” interpretation.

The IRF of a variable  $i$  to a one standard deviation shock of  $\varepsilon_{j,t}$  is defined as

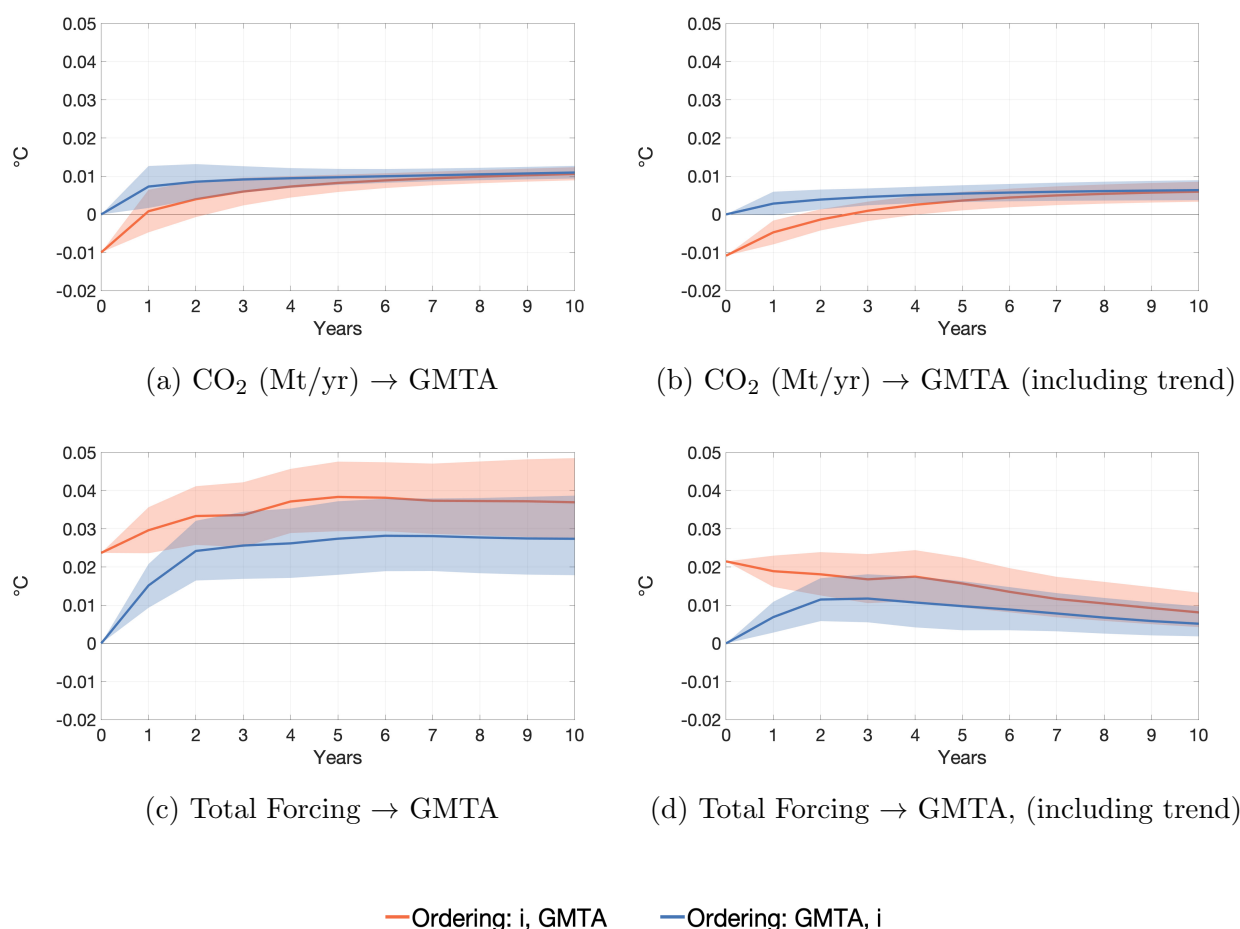
$$IRF(j \rightarrow i, h) = E(y_{i,t+h} | \mathbf{y}_t, \varepsilon_{t,j} = \sigma_{\varepsilon_j}) - E(y_{i,t+h} | \mathbf{y}_t, \varepsilon_{t,j} = 0). \quad (9)$$

Thus, it is the expected difference,  $h$  months after “impact”, between a bivariate system that responded to an unexpected CO<sub>2</sub> increase, and the same system where no such increase occurred. In a linear VAR, the above takes a closed-form solution in terms of the matrices from (6). Models are now estimated with Bayesian methods, optimizing the hyperparameters of a standard Minnesota prior, and choosing the number of lags as reported in the gray-shaded rows of Table 1. The primary motivation is to obtain a valid inference even in the presence of non-stationarity. Point estimates are nearly identical to that of OLS. For details on such choices in the context of climate data and a more thorough (yet introductory) treatment of IRFs, see Goulet Coulombe and Göbel (2021).

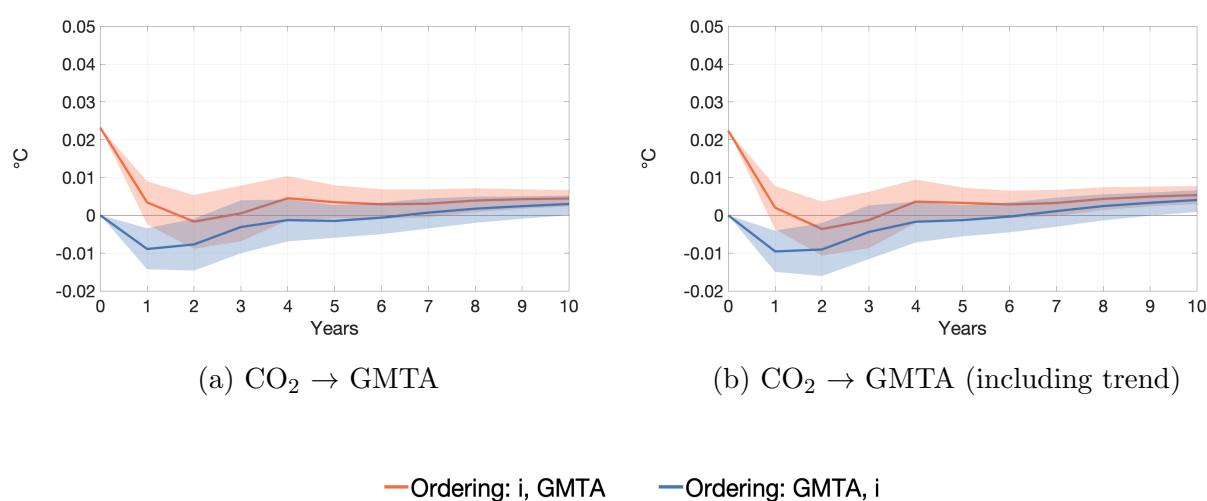
In Figure 2, we show the effect of an unexpected increase in *annual* emissions on GMTA, while Figure 3 shows the response of GMTA radiative forcing generated by an unexpected rise in *cumulative* emissions.

Qualitatively, the impact of annual emissions and cumulative CO<sub>2</sub>-induced forcing *shocks* on global temperature is vastly similar. In accordance with findings in Goulet Coulombe and Göbel (2021) for the effect of CO<sub>2</sub> on Arctic sea ice extent, the impact of total forcing and CO<sub>2</sub> shocks is highly durable. This time, it is on GMTA rather than sea ice extent. As reflected in Figures 2 and 3, this result is qualitatively independent of the ordering choice. In both cases (and for both forcing variables), the effect of forcing takes about two years to completely settle in. However, it is clear that the reported short-run impact strongly depends on the identification assumptions, which SMCGCL completely abstract from.

Whether the slightly negative short-run response of GMTA in the first panel of Figure 2 favors the ordering  $\{GMTA, CO_2\}$  over  $\{CO_2, GMTA\}$  is debatable: Forster et al. (2020) find the reduction in global CO<sub>2</sub> emissions during the COVID-19 pandemic to have resulted in a short-run rise of global temperature. The key mechanism is a decline in the cooling-effect of aerosols as a result of less SO<sub>2</sub> emissions. The authors project a rise in global temperature over the first 24 months following the pandemic-induced reduction in global nitrogen oxide (NO<sub>x</sub>) emissions.<sup>12</sup>



**Figure 2.** IRFs: Annual Emissions and Global Temperature. **Notes:** We show impulse response functions from bivariate Vector Autoregressions of GMTA and annual  $\text{CO}_2$  emissions. The right column includes a time trend as an additional exogenous regressor. The solid line is the median of 10,000 draws from the posterior distribution. Hyperparameters were optimized. The shaded area is the 68% credible region. Lags are those reported in Table 1 under the gray shaded rows.



**Figure 3.** IRFs: Cumulative Emissions and Global Temperature. **Notes:** We show impulse response functions from bivariate Vector Autoregressions of GMTA and  $\text{CO}_2$ , and Total Forcing, respectively. The right column includes a time trend as an additional exogenous regressor. The solid line is the median of 10,000 draws from the posterior distribution. Hyperparameters were optimized. The shaded area is the 68% credible region. Lags are those reported in Table 1 under the gray shaded rows.

Lastly, we conduct a robustness check. Our variables are clearly non-stationary. While this does not pose a problem for Bayesian inference, it is nevertheless natural to wonder if the results would be significantly altered by including a time trend. Accordingly, the second column reports the same IRFs, but for VAR specifications augmented with trends. Those show that adding such an exogenous explanatory variable does not change the dynamics of a GMTA response to an unexpected shock to CO<sub>2</sub>. The addition of a trend to the bivariate model of GMTA and total forcing allows GMTA to slowly revert to a lower impact—which is nevertheless highly persistent.

#### 4.3. Are VAR Estimates Quantitatively Reasonable?

The transient climate response (TCR) is a frequently used metric of the impact of rising atmospheric CO<sub>2</sub> concentration on temperature. It is not only an indication of the trajectory of ongoing climate change, but also serves as a benchmark to evaluate the results of climate model projections (Phillips et al. 2020). The TCR is defined as the increase in temperature, between  $h_0$  and  $h_T$ , under the assumption that CO<sub>2</sub> increases annually by 1%.  $h_T$  is defined as that point in time when—due to the steady annual increase of 1%—CO<sub>2</sub> concentration is twice as high as at date  $h_0$  (Montamat and Stock 2020; Pretis 2020). Such a doubling of CO<sub>2</sub> would occur approximately after 70 years (Otto et al. 2013). Following the transformation of ppm to W/m<sup>2</sup> as suggested by Myhre et al. (2013) and Montamat and Stock (2020), a doubling of CO<sub>2</sub> under an annual increase in concentration of 1% would generate a radiative forcing of  $5.35 \times \ln(2) \approx 3.7 \text{ W/m}^2$ .<sup>13</sup>

Typical estimates for TCR fall within a range of 1 °C–2.5 °C with a 66% probability, as summarized in the IPCC 5th Assessment Report (Bindoff et al. 2013). More recent estimates are well aligned with this range. Bruns et al. (2020) report a point estimate of TCR ranging from 1.17 °C to 1.85 °C, depending on the type of data and model specification. Pretis (2020) embeds a two-component energy balance model into a cointegrated vector autoregressive model. His estimates vary across model specification and range from 1.24 °C to 1.38 °C. Phillips et al. (2020) report a global transient climate sensitivity of 2.05 °C. The IV regression in Montamat and Stock (2020) allows for a differentiation of TCR measurements across different horizons, normalized to giving 70-year-horizon estimates. Their point estimates range in the neighborhood of 1.5 °C within a 95% confidence interval of roughly 0.9 °C to 2.1 °C.

Despite being much more simplistic than the models of the aforementioned works on TCR, bivariate VARs also allow for the estimation of the impact of a doubling of CO<sub>2</sub> on temperature. Here we make use of the concept of IRFs, as presented in Equation (9). In particular, we estimate the impact on temperature when  $RF^{CO_2}$  increases by one standard deviation of its reduced-form residuals,  $\sigma_{\varepsilon_j}$ , where  $j = RF^{CO_2}$ , from a bivariate VAR of  $RF^{CO_2}$  and GMTA. Recalling the definition of TCR, a doubling of CO<sub>2</sub> concentration, which is achieved by an annual increase of 1% in atmospheric CO<sub>2</sub> concentration, generates an additional radiative forcing of approximately  $5.35 \times \ln(2) \approx 3.7 \text{ W/m}^2$ . In our case, the shock  $\sigma_{\varepsilon_j}$  to CO<sub>2</sub> is a one-time event at horizon  $h = 0$ , but its effects are distributed over horizons  $h = 1, 2, 3, \dots, H$ . This allows us to measure the cumulative increase,  $\Xi_i$ , in  $i \in \{RF^{CO_2}, GMTA\}$  generated by  $\sigma_{\varepsilon_j}$ , where  $j = RF^{CO_2}$ , at any horizon  $h$ :

$$\Xi_{j,h} = \sum_{s=0}^h IRF(j \rightarrow i, s) ,$$

where  $IRF(j \rightarrow i, s)$  is defined as in Equation (9). Adapting the formula of Otto et al. (2013), we estimate  $TCR_h$  as the increase in GMTA at horizon  $h$  as follows:

$$TCR_h = \Xi_{GMTA,h} \times \frac{5.35 \times \ln(2)}{\Xi_{RF^{CO_2},h}} , \quad (10)$$

where  $\Xi_{GMTA,h}$  is the cumulative increase in global temperature at horizon  $h$ , resulting from the shock  $\sigma_{\varepsilon_j}$ , where  $j = RF^{CO_2}$ , at horizon  $h = 0$ . Likewise,  $\Xi_{RF^{CO_2},h}$  is the cumulative

increase in radiative forcing of CO<sub>2</sub> at horizon  $h$ , resulting from the shock  $\sigma_{\varepsilon_j}$ , where  $j = RF^{CO_2}$ , at horizon  $h = 0$ .

In Table 2, we present median point estimates of TCR <sub>$h$</sub>  for  $h = 20$  and  $h = 70$  (as in Montamat and Stock (2020)) from bivariate VARs of  $RF^{CO_2}$  and GMTA—with and without an exogenous time trend. Thus, Table 2 reports the TCR corresponding to the model specifications in Figure 3a,b. That is, we use Bayesian estimation techniques and deploy a Minnesota prior on our parameter estimates. Our VAR has four lags and the estimation is based on annual observations between 1850 and 2005. Results for 1850–2017 are available in Table A3 (Appendix C).

**Table 2.** Transient Climate Response.

Ordering	Without Trend		With Trend	
	TCR <sub>20</sub>	TCR <sub>70</sub>	TCR <sub>20</sub>	TCR <sub>70</sub>
CO <sub>2</sub> , GMTA	1.99 °C	2.06 °C	2.17 °C	2.58 °C
GMTA, CO <sub>2</sub>	0.57 °C	1.82 °C	0.85 °C	2.39 °C

The main message of Table 2 is two-fold: first, the ordering of the variables heavily influences the final results for TCR<sub>20</sub>, demonstrating the importance of respecting the possibility of cross-correlated residuals. Mechanically, the discordance brought up by the ordering choice vanishes at much longer horizons, and, as a result, TCR<sub>70</sub> estimates are largely similar. However, at that horizon, it is the choice of whether or not to include a trend that can alter results significantly. Second, even though the TCR <sub>$h$</sub>  point estimates of the trend models are rather located at the upper bound of the IPCC range of 1 °C–2.5 °C, a simplistic bivariate VAR model including a constant and a time trend as additional exogenous regressors is capable of providing a reasonable approximation of the rise in global mean temperature, triggered by a doubling of atmospheric CO<sub>2</sub> concentration.

TCR estimates can be helpful in choosing which ordering is most plausible. For instance, only by ordering CO<sub>2</sub> first do we achieve TCR<sub>20</sub> within the IPCC range. IRFs can also help sort things out. Ordering CO<sub>2</sub> second leads to a surprisingly lasting *negative* effect of CO<sub>2</sub> shocks on GMTA. An increasingly popular approach to VAR identification in macroeconomics is to use sign restrictions, where implausible IRF draws (based on economic theory) are tossed out (Uhlig 2005). This dispenses the researcher from formulating a likely contentious causal ordering of variables. In a climate application, one could identify the VAR by rejecting specifications generating implausible  $IRF(CO_2 \rightarrow GMTA, h)$  or TCRs. Applying this sort of reasoning leads us to favor—to nobody’s surprise—the specification where simultaneous causality runs from CO<sub>2</sub> to GMTA. However, it is important to stress that this choice is obtained from prior knowledge on what is deemed reasonable and what is not, rather than our two time series.

## 5. Conclusions

This note is a cautionary tale about how seemingly innocuous simplifying assumptions can go wrong—especially when they are formulated without consulting the data. IFs, as proposed by SMCGL, are a concept that hinges on the assumption of zero correlation between the residuals of a bivariate VAR(1) process. In discrete time, especially with observations at lower frequencies, such an assumption is most often not justified. Both stylized simulations and an empirical application in the form of the transient climate response demonstrate that being negligent about cross-correlated residuals can lead to markedly different outcomes. Our results show that, already in a bivariate system of CO<sub>2</sub> and GMTA, the resulting TCR depends on how one deals with correlated residuals.<sup>14</sup> Nevertheless, provided reasonable modeling choices, the practical implications of our VAR results are in line with previous works (Richardson et al. 2016; Schurer et al. 2018), though when including an exogenous time-trend, he estimates range at the upper bound of estimates found in Otto et al. (2013), Montamat and Stock (2020), or Forster et al. (2021).



That is, for a doubling of atmospheric CO<sub>2</sub> concentration our estimates suggest GMTA to increase by 1.99 °C (2.17 °C with trend) after 20 years and 2.06 °C (2.58 °C with trend) after 70 years.

FEVDs provide an alternative to IFs. Although a decisive improvement over IFs, FEVDs are as good as their underlying statistical model. To avoid departing too much from the SMCGL's framework, we only considered bivariate models. Climate systems obviously comprise of numerous additional variables. [Granville Tunnicliffe \(2015\)](#) considers four in a VAR setup, but there could be many more. Additionally, the dynamics were assumed to be linear and time invariant, an approximation that should be eventually tested. [Estrada et al. \(2013\)](#)'s trend "breaks" model is one way to do it and they report results pointing in the same direction as ours. However, with use of machine learning tools becoming increasingly common, more flexible alternatives could be used to yield further insights—like those developed for macroeconomic time series in [Goulet Coulombe \(2020\)](#). Finally, we considered simplistic identification schemes that implied a causal ordering of variables. There is a plethora of more sophisticated schemes available ([Kilian and Lütkepohl 2017](#)) and those could be used in future work, especially when moving beyond bivariate systems. Another (simpler) avenue is the use of data sampled at higher frequencies (daily for example) which, by construction, makes the simultaneity problem much less of a Damocles sword.

**Author Contributions:** All authors contributed equally. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data are either available from public sources, as indicated in Table A1, or from the authors' websites.

**Acknowledgments:** We are grateful to Boyuan Zhang for many helpful comments and suggestions.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A. Data Sources

For the empirical part of Section 4, we refer to the variables presented in Table 1 of SMCGL. Due to the sparse description of the data sources, our data set reduces to the annual global mean surface temperature anomalies (GMTA) and seven other representatives of radiative forcing. For five variables we could match the reported correlation coefficients as well as the IF values of SMCGL, Table 1. For *total forcing* and *solar* these measures do not coincide. For the analysis in Sections 4.2 and 4.3, we use the CO<sub>2</sub> series based on [Meinshausen et al. \(2017\)](#). We transform *ppm* into W/m<sup>2</sup> as described in Section 4.

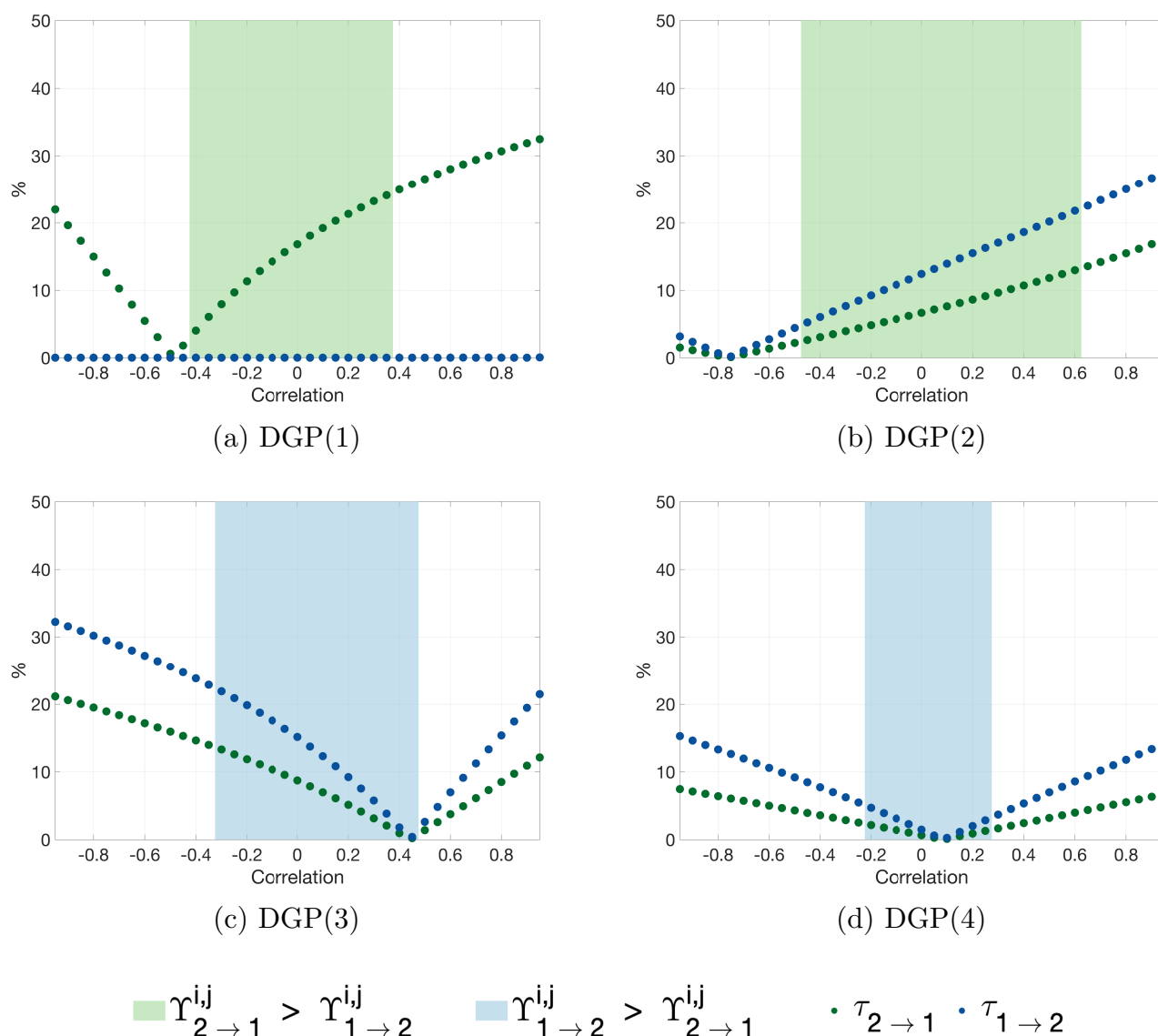
**Table A1.** List of Variables.

Abbreviation	Description	Data Source
Total Forcing	annual; 1850–2005	<a href="#">KNMI Climate Explorer</a>
Anthropogenic	annual; 1850–2005	<a href="#">KNMI Climate Explorer</a>
CO <sub>2</sub> -ERF (W/m <sup>2</sup> )	annual; 1850–2005	<a href="#">KNMI Climate Explorer</a>
Aerosol	annual; 1850–2005	<a href="#">KNMI Climate Explorer</a>
Solar	annual; 1850–2005	<a href="#">KNMI Climate Explorer</a>
Volcanic	annual; 1850–2005	<a href="#">KNMI Climate Explorer</a>
PDO	annual averages of monthly observations; 1900–2005	<a href="#">KNMI Climate Explorer</a>
GMTA	annual; global; 1900–2005	<a href="#">HadCRUT4</a>
CO <sub>2</sub> (Million Tonnes/Year)	annual; global production-based emissions; 1850–2005	<a href="#">Our World in Data</a> —not in use
CO <sub>2</sub> (ppm)	annual; global 1850–2014 ( <a href="#">Meinshausen et al. 2017</a> ): 2015–2017 (NOAA-ESRL):	<a href="#">IAC ETH Zürich</a> <a href="#">NOAA-ESRL</a>



## Appendix B. Additional Simulation Results

In Figure A1 below, we compare NIF and  $FEVD_{h=2}$ . One might be concerned that results in Figure 1 suffer from a horizon mismatch between NIF and  $FEVD_{h=10}$ . Results suggest that reducing  $h$  worsens NIFs' problems by shrinking the "safe" regions. This is intuitive; the effect of assumptions on simultaneous relationships becomes milder as we move further from  $h = 0$ .



**Figure A1.** Ranking of NIFs ( $\tau_{i \rightarrow j}$ ) vs. Ranking of FEVDs ( $Y_{i \rightarrow j}^{i,j}$ ); Different Levels of Correlation; Horizon  $h = 2$ . **Notes:** Color specification: the *light blue* area shows the region, in which FEVDs suggest  $Y_{1 \rightarrow 2}^{i,j} > Y_{2 \rightarrow 1}^{i,j}$ , regardless of the ordering. Similarly, the *light green* area shows the region, in which FEVDs suggest  $Y_{2 \rightarrow 1}^{i,j} > Y_{1 \rightarrow 2}^{i,j}$ —independent of the ordering. The *white/non-colored* areas shows those regions for which the ordering of  $X_1$  and  $X_2$  does matter. An unambiguous determination of  $Y_{1 \rightarrow 2}^{i,j} > Y_{2 \rightarrow 1}^{i,j}$ , respectively,  $Y_{2 \rightarrow 1}^{i,j} > Y_{1 \rightarrow 2}^{i,j}$ , is not possible.

## Appendix C. Additional Empirical Results

In this section, we show results for the sample period to 1850–2017.

**Table A2.** Empirical Results for the Bivariate Relationship Between Various Forcings and GMTA Sample Period: 1850–2017.

	Correlation	Normalized IF (IF × 100)		Lags (P)	Correlation of Residuals ( $\rho_u$ )	FEVD			
		$i \rightarrow$ GMTA	GMTA $\rightarrow i$			Ordering: $i$ , GMTA		Ordering: GMTA, $i$	
						$i \rightarrow$ GMTA	GMTA $\rightarrow i$	$i \rightarrow$ GMTA	GMTA $\rightarrow i$
Total Forcing	0.82	36.8 (17.2)	29.5 (15.3)	4 1	0.23 *** 0.29 ***	48.0 55.7	15.6 14.4	28.3 30.0	29.3 36.6
Anthropogenic	0.91	43.7 (39.9)	−18.9 (−0.6)	4 1	−0.16 ** −0.15 **	6.5 4.2	5.0 5.1	4.9 2.4	13.5 13.7
CO <sub>2</sub> —ERF (W/m <sup>2</sup> ) SMCGL	0.91	43.5 (39.0)	−13.0 (−0.3)	4 1	−0.10 −0.11	5.75 2.1	9.2 3.1	5.3 1.2	15.4 8.0
Aerosol	−0.84	37.9 (19.4)	−45.7 (−1.3)	4 1	−0.17 ** −0.00	3.9 2.0	2.9 33.3	6.1 2.0	0.0 33.1
Solar	31.4	7.0 (1.1)	2.1 (0.6)	8 1	0.05 0.06	4.6 6.7	0.9 1.0	3.1 4.4	1.4 2.0
Volcanic	0.11	1.3 (0.2)	−0.3 (−0.2)	4 1	0.18 ** 0.20 ***	10.1 7.1	0.5 0.3	2.7 0.6	2.4 3.7
PDO	0.15	−2.3 (−0.3)	−0.4 (−0.3)	4 1	0.4 *** 0.38 ***	31.0 9.5	0.8 0.3	3.9 0.7	13.7 13.5
CO <sub>2</sub> (Mt/yr)	0.89	42.0 (31.0)	1.0 (0.00)	2 1	−0.12 −0.10	7.9 5.4	1.9 0.0	8.8 5.4	0.3 0.7
CO <sub>2</sub> (W/m <sup>2</sup> )	0.91	43.4 (38.3)	−13.4 (−0.3)	2 1	0.18 ** 0.06	5.0 1.5	16.1 3.4	6.1 1.0	6.2 1.6

**Notes:**  $i$  corresponds to the type of radiative forcing, listed in the left most column. The second column ("Correlation") documents the correlation between GMTA and variable  $i$ . FEVD values are taken at horizon  $h = 15$ , which translates into the contribution of variable  $i$  in the variance of the forecast error of variable  $j$  a decade and a half after the in-sample end date. Numbers in bold underline the highest absolute causal flow among a ( $i$ , GMTA) pair for a given measure. "\*\*\*", "\*\*", and "\*" means that the null of the residuals cross-correlation of residuals is rejected at the 10%, 5%, and 1% level, respectively. Sample period: 1850–2017.

**Table A3.** Transient Climate Response Sample Period: 1850–2017

Ordering	Without Trend		With Trend	
	TCR <sub>20</sub>	TCR <sub>70</sub>	TCR <sub>20</sub>	TCR <sub>70</sub>
CO <sub>2</sub> , GMTA	1.46 °C	1.94 °C	1.76 °C	2.35 °C
GMTA, CO <sub>2</sub>	0.58 °C	1.79 °C	0.97 °C	2.22 °C

## Notes

- In [Tawia Hagan et al. \(2019\)](#), IFs are used on daily data, which can alleviate the problem if there are no intra-day relationships. This last condition is something that should be verified, not assumed.
- Of course, there are identification schemes outside of the family of "orderings" obtained by Choleski decomposition, but those are beyond the scope of this paper and unnecessary to make our main point.
- For a discussion on how to think about "shocks" in a physical system, see [Goulet Coulombe and Göbel \(2021\)](#).
- Variances of  $u_t^{X_1}$  and  $u_t^{X_2}$  are one.
- It is important to note that while we consider cases where either  $b_{12} = 0$  or  $b_{21} = 0$ , there is a continuum of possibilities between those. We do so for simplicity of exposition (it makes the problem dichotomous). Moreover, setting either  $b_{12} = 0$  or  $b_{21} = 0$  to zero corresponds to a causal ordering that is by far the most common identification scheme used in practice, which happens to be what we will be using in Section 4.
- Resolving this puzzle has led to re-evaluating the role of oceans in the interplay of radiative forcing and the climatic response ([Marotzke and Forster 2015](#); [Tollefson 2014](#)).
- Data on the Pacific Decadal Oscillation (PDO) range from 1900 to 2005.
- The sample was restricted to 1850–2005 to match that of SMCGL. Table A2 reports results extending the sample to 2017.
- There are other identification schemes that cannot be cast as "orderings". That is, there are rotations of  $u_t$  (even when  $\rho_u = 0$ ) giving different structural shocks. Hence, while the two orderings span a lot of possibilities (and those traditionally considered first in practice), they do represent the universe of rotations of  $u_t$  into  $e_t$ .
- [Hansen et al. \(2006\)](#) states that global warming did not start to accelerate prior to the 1970s. Only after 1975 did the global temperature increase by approximately 0.2°C per decade.

- 11 See: Available online: [https://drive.google.com/file/d/1yhaJi92dvY\\_Lax0H5BFlzJeNc517Gn44/view?usp=sharing](https://drive.google.com/file/d/1yhaJi92dvY_Lax0H5BFlzJeNc517Gn44/view?usp=sharing) (accessed on 31 August 2021).
- 12 Especially NO<sub>2</sub> is found to be well-correlated with CO<sub>2</sub> emissions (Forster et al. 2020).
- 13 An annual increase of 1% in atmospheric CO<sub>2</sub> concentration results in a doubling of CO<sub>2</sub> after approximately 70 years, which is described more formally as:  $h \times \ln\left(\frac{1.01}{1}\right) = \ln(2)$ , for  $h \approx 70$  (Montamat and Stock 2020).
- 14 With the proliferation of new methods to extract information flows between time series (e.g., fractal regressions (Kristoufek and Ferreira 2018) or multiscale transfer entropy (Zhao et al. 2018)), there is much research to be done on dealing with the simultaneity problem within those heterogeneous frameworks.

## References

- Andrews, T., P. M. Forster, O. Boucher, N. Bellouin, and A. Jones. 2010. Precipitation, Radiative Forcing and Global Temperature Change. *Geophysical Research Letters* 37. [CrossRef]
- Bindoff, N. L., P. A. Stott, P. A. AchutaRao, M. R. Allen, N. Gillett, D. Gutzler, K. Hansingo, G. Hegerl, Y. Hu, S. Jain, and et al. 2013. Chapter 10—Detection and Attribution of Climate Change: From Global to Regional. In *Climate Change 2013: The Physical Science Basis. IPCC Working Group I Contribution to AR5*. Cambridge: Cambridge University Press.
- Bruns, S. B., Z. Csereklyei, and D. I. Stern. 2020. A Multicointegration Model of Global Climate Change. *Journal of Econometrics* 214: 175–97. Annals Issue: Econometric Models of Climate Change. [CrossRef]
- Chen, Gang, Daniel R. Glen, Ziad S. Saad, J. Paul Hamilton, Moriah E. Thomason, Ian H. Gotlib, and Robert W. Cox. 2011. Vector autoregression, structural equation modeling, and their synthesis in neuroimaging data analysis. *Computers in Biology and Medicine* 41: 1142–55. [CrossRef]
- Estrada, Francisco, Pierre Perron, and Benjamín Martínez-López. 2013. Statistically derived contributions of diverse human influences to twentieth-century temperature changes. *Nature Geoscience* 6: 1050–5. [CrossRef]
- Etminan, M., G. Myhre, E. J. Highwood, and K. P. Shine. 2016. Radiative Forcing of Carbon Dioxide, Methane, and Nitrous Oxide: A significant Revision of the Methane Radiative Forcing. *Geophysical Research Letters* 43: 12614–23. [CrossRef]
- Forster, P. M., H. I. Forster, M. J. Evans, M. J. Gidden, C. D. Jones, C. A. Keller, R. D. Lamboll, C. L. Quéré, J. Rogelj, D. Rosen, and et al. 2020. Current and Future Global Climate Impacts Resulting from COVID-19. *Nature Climate Change* 10: 913–19. [CrossRef]
- Forster, P., T. Storelvmo, K. Armour, W. Collins, J. L. Dufresne, D. Frame, D. J. Lunt, T. Mauritsen, M. D. Palmer, M. Watanabe, and et al. 2021. The Earth's Energy Budget, Climate Feedbacks, and Climate Sensitivity. In *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Edited by V. Masson-Delmotte, P. Zhai, A. Pirani, S. Connors, C. Péan, S. Berger, Y. Caud, N. Chen, L. Goldfarb, M. Gomis and et al. Cambridge: Cambridge University Press.
- Goulet Coulombe, Philippe. 2020. The Macroeconomy as a Random Forest. Available online: <https://ssrn.com/abstract=3633110> (accessed on 31 August 2021).
- Goulet Coulombe, Philippe, and Maximilian Göbel. 2021. Arctic amplification of anthropogenic forcing: A vector autoregressive analysis. *Journal of Climate* Forthcoming.
- Granger, Clive W. J. 1969. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society* 37: 424–38. [CrossRef]
- Granville Tunnicliffe, Wilson. 2015. Atmospheric co2 and global temperatures: The strength and nature of their dependence. In *Working Paper*. Lancaster: Lancaster University.
- Hansen, J., M. Sato, R. Ruedy, K. Lo, D. W. Lea, and M. Medina-Elizade. 2006. Global Temperature Change. *Proceedings of the National Academy of Sciences* 103: 14288–93. [CrossRef]
- Kilian, Lutz, and Helmut Lütkepohl. 2017. *Structural Vector Autoregressive Analysis*. Cambridge: Cambridge University Press.
- Koutsoyiannis, Demetris, and Zbigniew W Kundzewicz. 2020. Atmospheric temperature and co2: Hen-or-egg causality? *Sci* 2: 83. [CrossRef]
- Kristoufek, Ladislav, and Paulo Ferreira. 2018. Capital asset pricing model in portugal: Evidence from fractal regressions. *Portuguese Economic Journal* 17: 173–83. [CrossRef]
- Li, C., D. Notz, S. Tietsche, and J. Marotzke. 2013. The Transient versus the Equilibrium Response of Sea Ice to Global Warming. *Journal of Climate* 26: 5624–36. [CrossRef]
- Liang, X. S. 2008. Information Flow within Stochastic Dynamical System. *Phys. Rev. E* 78: 031113. [CrossRef] [PubMed]
- Liang, X. S. 2014. Unraveling the Cause-Effect Relation between Time Series. *Physical Review E* 90: 052150. [CrossRef] [PubMed]
- Liang, X. S. 2015. Normalizing the Causality between Time Series. *Physical Review E* 92: 022126. [CrossRef]
- Liang, X. S. 2016. Information Flow and Causality as rigorous Notions ab initio. *Physical Review E* 94: 052201. [CrossRef]
- Marotzke, J., and P. M. Forster. 2015. Forcing, Feedback and Internal Variability in Global Temperature Trends. *Nature* 517: 565–70. [CrossRef]

- Masson-Delmotte, V., P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P. R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, and et al. 2018. 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*. Available online: <https://www.ipcc.ch/2018/10/08/summary-for-policymakers-of-ipcc-special-report-on-global-warming-of-1-5c-approved-by-governments/> (accessed on 31 August 2021).
- Meinshausen, M., E. Vogel, A. Nauels, K. Lorbacher, N. Meinshausen, D. M. Etheridge, P. J. Fraser, S. A. Montzka, P. J. Rayner, C. M. Trudinger, and et al. 2017. Historical greenhouse gas concentrations for climate modelling (cmip6). *Geoscientific Model Development* 10: 2057–116. [\[CrossRef\]](#)
- Montamat, G., and J. H. Stock. 2020. Quasi-experimental estimates of the transient climate response using observational data. *Climatic Change* 160: 361–71. [\[CrossRef\]](#)
- Myhre, G., D. Shindell, F.-M. Bréon, W. Collins, J. Fuglestad, J. Huang, D. Koch, J.-F. Lamarque, D. Lee, B. Mendoza, and et al. 2013. Anthropogenic and Natural Radiative Forcing. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Edited by T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Doschung, A. Nauels, Y. Xia, V. Bex and P. Midgley. Cambridge: Cambridge University Press, pp. 659–740. [\[CrossRef\]](#)
- Nordhaus, W. 2014. Estimates of the social cost of carbon: Concepts and results from the dice-2013r model and alternative approaches. *Journal of the Association of Environmental and Resource Economists* 1: 273–312. [\[CrossRef\]](#)
- Notz, D., and J. Stroeve. 2016. Observed arctic sea-ice loss directly follows anthropogenic co2 emission. *Science* 354: 747–50. [\[CrossRef\]](#) [\[PubMed\]](#)
- Otto, A., F. E. L. Otto, O. Boucher, J. Church, G. Hegerl, P. M. Forster, N. P. Gillett, J. Gregory, G. C. Johnson, R. Knutti, and et al. 2013. Energy Budget Constraints on Climate Response. *Nature Geoscience* 6: 415–6. [\[CrossRef\]](#)
- Phillips, P. C. B., T. Leirvik, and T. Storelvmo. 2020. Econometric Estimates of Earth's Transient Climate Sensitivity. *Journal of Econometrics* 214: 6–32. [\[CrossRef\]](#)
- Pretis, Felix. 2020. Econometric Modelling of Climate Systems: The Equivalence of Energy Balance Models and Cointegrated Vector Autoregressions. *Journal of Econometrics* 214: 256–73. [\[CrossRef\]](#)
- Richardson, M., K. Cowtan, E. Hawkins, and M. B. Stolpe. 2016. Reconciled climate response estimates from climate models and the energy budget of earth. *Nature Climate Change* 6: 931–5. [\[CrossRef\]](#)
- Schurer, A., G. Hegerl, A. Ribes, D. Polson, C. Morice, and S. Tett. 2018. Estimating the Transient Climate Response from Observed Warming. *Journal of Climate* 31: 8645–63. [\[CrossRef\]](#)
- Shukla, P. R., J. Skea, E. Calvo Buendia, V. Masson-Delmotte, H.-O. Pörtner, D. C. Roberts, P. Zhai, R. Slade, S. Connors, R. van Diemen, and et al. 2019. Climate Change and Land: An IPCC Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse Gas Fluxes in Terrestrial Ecosystems. Available online: <https://spiral.imperial.ac.uk/bitstream/10044/1/76618/2/SRCCL-Full-Report-Compiled-191128.pdf> (accessed on 6 September 2021).
- Sims, Christopher A. 1980. Macroeconomics and reality. *Econometrica: Journal of the Econometric Society* 48: 1–48. [\[CrossRef\]](#)
- Stips, A., D. Macias, C. Coughlan, E. Garcia-Gorriz, and X. S. Liang. 2016. On the Causal Structure between CO<sub>2</sub> and Global Temperature. *Scientific Reports* 6: 21691. [\[CrossRef\]](#) [\[PubMed\]](#)
- Tawia Hagan, Daniel Fiifi, Guojie Wang, X San Liang, and Han AJ Dolman. 2019. A time-varying causality formalism based on the liang–kleeman information flow for analyzing directed interactions in nonstationary climate systems. *Journal of Climate* 32: 7521–37. [\[CrossRef\]](#)
- Tollefson, J. 2014. Climate Change: The Case of the Missing Heat. *Nature* 505: 276–8. [\[CrossRef\]](#) [\[PubMed\]](#)
- Uhlig, H. 2005. What are the effects of monetary policy on output? results from an agnostic identification procedure. *Journal of Monetary Economics* 52: 381–419. [\[CrossRef\]](#)
- Zhao, Xiaojun, Yupeng Sun, Xuemei Li, and Pengjian Shang. 2018. Multiscale transfer entropy: Measuring information transfer on multiple time scales. *Communications in Nonlinear Science and Numerical Simulation* 62: 202–12. [\[CrossRef\]](#)