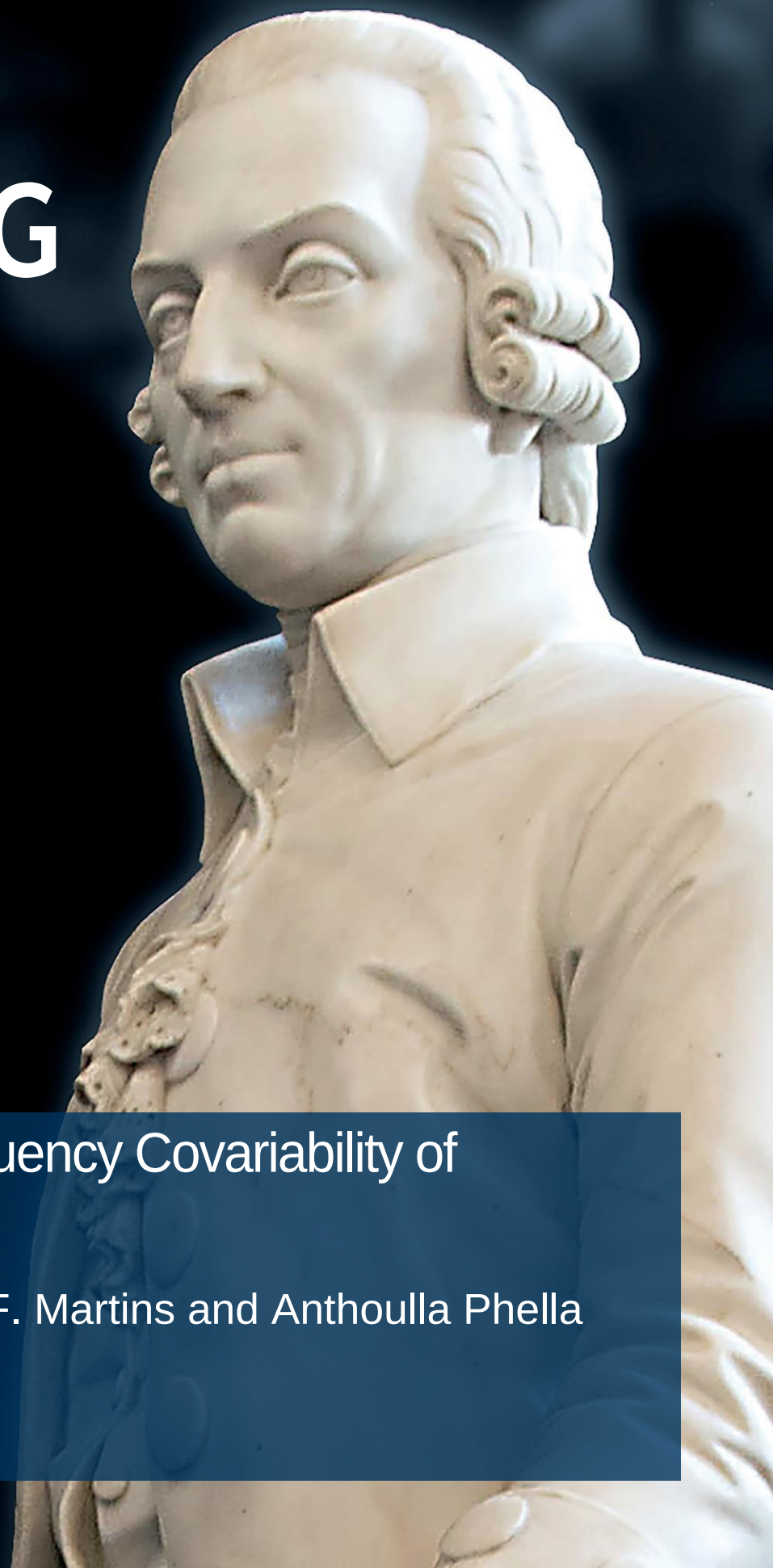
University of Glasgow

Adam Smith Business School

# WORKING PAPER SERIES



## Modelling Low-Frequency Covariability of Paleoclimatic Data

Vasco J. Gabriel, Luis F. Martins and Anthoulla Phella

Paper No. 2022-17
January 2022

# Modelling Low-Frequency Covariability of Paleoclimatic Data

Vasco J. Gabriel, University of Victoria, Canada

Luis F. Martins, ISCTE - Instituto Universitário de Lisboa, Portugal and CIMS, UK

Anthoulla Phella, University of Glasgow, UK

This version: December 2021

**Abstract**

This paper explores formal statistical procedures that allow us to quantify low-frequency comovement amongst a range of paleoclimate times series. Our first contribution is methodological: we extend the long-run covariability approach of Müller and Watson (2018) to higher dimensional settings by means of a first-pass partialling out of exogenous sources of variation. Our second contribution is empirical: we provide new estimates for the long-run relationship between temperatures and $CO_2$, concluding that in the long-run a 100 ppm increase in $CO_2$ levels would raise temperatures around $1°C$. Finally, we illustrate how joint modelling of this set of paleoclimate time series can be carried out by factor analysis and how long-term projections about temperature increases and ice-sheet retreat can be constructed.

Keywords: Paleoclimate data; Glaciar cycles; Equilibrium climate sensitivity; Low frequency analysis.

JEL Classification: C22, C53, Q54.

# 1 Introduction

The analysis of paleoclimate data has played an important role in informing the debate around the causes of climate change. While Milankovitch cycles provide a useful baseline to understand the long term drivers of Earth's climate, the complex dynamics of glacial/interglacial cycles, with persistent and non-negligible deviations from 'Milankovitch-induced' equilibria, present challenges for the study of the long-run features of climate data. Indeed, understanding the long-run relationship between temperatures and $CO_2$ levels is crucial for climate sensitivity simulations and to better understand the effects of anthropogenically emitted $CO_2$ (see Knutti, Rugenstein and Hegerl, 2017, for example).

A strand of the literature has focused on the application of cointegrated vector autoregressions, making use of inherent error-correction mechanisms to model long-run relationships between climate and orbital geometry - for example, see Kaufmann and Juselius (2013) and Kaufmann and Pretis (2020, 2021), which find support for a long-run equilibrium driven by solar insolation, but disturbed by interactions among components of the climate system. However, evidence for the presence of stochastic trends in the relevant time series is mixed at best. Many studies generate results that are consistent with the presence of a stochastic trend (Gordon, 1991, Woodward and Gray, 1993, Woodward and Gray, 1995, Gordon et al., 1996, Kärner, 1996). Conversely, many other studies generate results that are consistent with the presence of a deterministic trend with possibly highly persistent noise (Bloomfield, 1992, Bloomfield and Nychka, 1992, Baillie and Chung, 2002, Fomby and Vogelsang, 2002). The results in Davidson, Stephenson and Turasie (2016), for example, do not support the hypothesis of integrated behaviour.

While the role of orbital variation in driving ice ages is well established, considerable attention has been devoted to the interplay of temperatures, ice volume and atmospheric concentrations of $CO_2$. The treatment of the latter and, in particular, whether it should be treated as a forcing variable or an endogenous response, has been subject of some debate. See Lea (2004) and Jaccard et al. (2016), for example, who suggest that orbital forcing drive variations in temperatures, which in turn affect ice volume and how much trapped $CO_2$ is released into the atmosphere. In turn, Davidson et al. (2016) find that $CO_2$ and other greenhouse gases are Granger-caused by temperatures.

Consequently, Castle and Hendry (2020) model these variables as a jointly endogenous system, with orbital forcing variables deemed to be strongly exogenous. They find evidence of an endogenous response of $CO_2$ to orbital forcing, as well as support for a 'weak' form of the Milankovitch hypothesis, in that to account for all aspects of glacial cycles one needs to also consider nonlinear interactions between the different orbital components.

In an interesting by-product of their estimations, and given that the path of orbital variables can be obtained well into the future, Castle and Hendry (2020) look at the long-run implications of anthropogenically determined levels of $CO_2$. Indeed, using the solved long-run coefficient of

$CO_2$, they simulate the impact on temperatures of increasing atmospheric levels of $CO_2$: for instance, an increase of 105 ppm (as seen since 1958) from 280 ppm raises temperatures by 6.9 $^{\circ}C$, *ceteris paribus*.

Thus, and unlike the works of Kaufmann and co-authors, as well as Davidson et al. (2016), Castle and Hendry (2020) or Proietti and Maddanu (2021), whose focus is on the cyclical properties of the paleoclimate data, in our study we reconsider the long-run comovement between temperatures and $CO_2$ by employing recently developed methods (Müller and Watson, 2018, 2021, MW henceforth) that highlight low-frequency covariability of time series. Indeed, we suggest that variation in the usual paleoclimate time series is dominated by glaciation cycles and therefore contain only limited information about the very long-run relation between temperatures and $CO_2$. Using the procedures of MW allows us to isolate a small number of low-frequency trigonometric weighted averages, which are then used to conduct inference about the long-run (co)variability of temperatures and $CO_2$. Moreover, in order to attenuate 'curse-of-dimensionality' issues, we extend the MW techniques by allowing for a group of effects to be 'partialled out'. This allow us to hone in on the long-run relationship between temperatures and $CO_2$ while controlling for orbital forcing.

This approach is quite novel and distinctive in this literature and there are several advantages in using these low-frequency techniques. First, they allow us to focus on time spans that go beyond cyclical dynamics and thus obtain better estimates of long-run coefficients that are uncontaminated by short or medium run variations. Second, the methodology is flexible in that we can conduct inference on bivariate or multivariate low-frequency features of several climate variables. Third, these methods are fairly robust to the persistence patterns of paleoclimate data, permitting combinations of non-stationary, near-nonstationary or stationary series, thus circumventing the need to (pre) test for unit roots and cointegration. Indeed, the data transformations are approximately Gaussian and therefore standard inference tools and confidence intervals can be employed. Finally, the construction of long horizon forecasts is relatively straightforward, thus allowing us to complement the scenario analysis in Castle and Hendry (2020).

Put simply, the method consists in obtaining low-frequency weighted averages of the data by using trigonometric projections that ensure approximate Gaussianity of the transformed series. Inference is then based on a (relatively) small number $q$ of low-frequency averages, whereby correlation or regression coefficients (and respective standard errors) can be computed in the usual way. The choice of $q$ determines the cyclicality that the researcher wishes to study. In our case, we are agnostic about the choice of $q$ and suggest selecting $q$ such that covariability is maximized. As it turns out, in our case $q$ ranges from 16 to 20, which indicates spans of data between 80,000 to 100,000 years.

Thus, we start by obtaining measures of long-run covariability between temperatures and levels of $CO_2$, with our results suggesting that correlation is indeed strong and quite significant,

with a 90% confidence interval of $[0.90, 0.98]$, while the long-run regression coefficient is estimated at around 0.09. Given that the exclusively bivariate focus is likely to lead to overestimation, we then develop an extension of the MW approach that allows us to control for measures capturing orbital forcing, i.e., a selection of control variables is 'partialled out' prior to the application of low-frequency projections. The resulting confidence intervals for the regression coefficient of temperatures on $CO_2$ range between 0.048 and 0.106, and therefore not too dissimilar to the results of Castle and Hendry (2020).

We also pursue the alternative approach of summarizing relevant exogenous information by extracting a common factor driving orbital forcing and $CO_2$ levels, with these results confirming our previous findings. Finally, emulating Castle and Hendry (2020), we present long-term forecasts for temperatures and ice-volume that can be constructed from low-frequency (multivariate) factor models, conditioning on (anthropogenically determined) current levels of $CO_2$ concentrations, which are far higher than those typical during previous glaciar cycles. Our results suggest a steady increase in temperatures, coupled with a substantial decline in ice volumes well below historical minima, pointing towards an ice-free planet under current $CO_2$ levels.

The paper is organised as follows. Section 2 provides a brief description of the data and the low-frequency inference procedures used throughout the paper, presenting baseline results for the simple bivariate relationship between temperatures and $CO_2$ levels. Section 3 develops the extension to the MW setup by considering additional controls, which can either be 'partialled out' or encapsulated into a common factor extracted via principal component analysis. Section 4 discusses multivariate low-frequency analysis and forecasting of paleoclimate time series, while section 5 concludes.

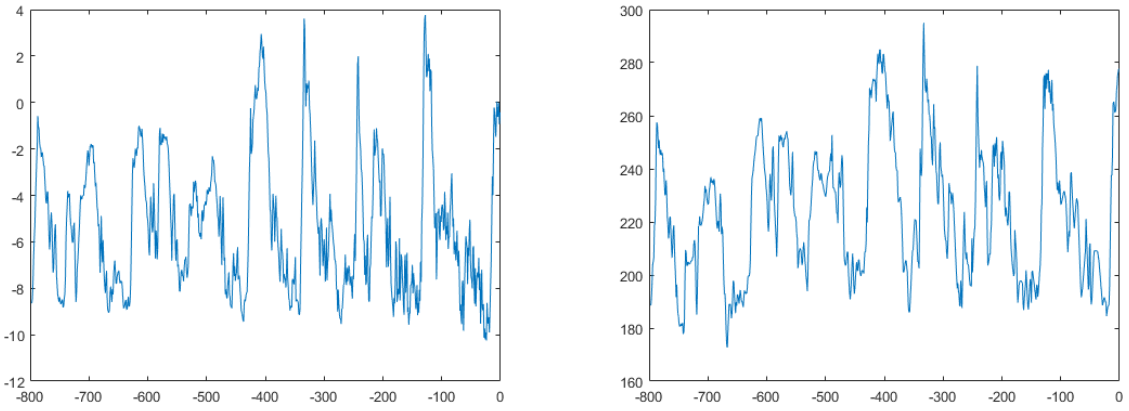
## 2 Low-frequency inference: the bivariate case

### 2.1 Data and basic method

We focus on three paleoclimate time series: ice cores reconstructions of temperatures, $CO_2$ atmospheric levels and ice volume. The first two series are sourced from the European Project for Ice Core in Antarctica (EPICA) (Jouzel et al., 2007; Loulergue et al., 2008; Lüthi et al., 2008), with $CO_2$ measured in ppm (parts per million, where 1 ppm= 7.8 gigatonnes of $CO_2$, while the ice volume series is from Lisiecki and Raymo (2005). In addition, and following Kaufmann and Juselius (2013), we consider standard ice-age orbital drivers, namely Eccentricity ($Ec$), Obliquity ($Ob$) and Precession ($Pr$).

The period comprises 800,000 years (800 kyr), with all observations adjusted to the common EDC3 time scale and linearly interpolated for missing observations. The total sample size in 1000 year intervals is thus $T = 801$ with the last 100 observations (i.e., 100,000 years, ending 1000 years before the present) used to evaluate the predictive ability of our models (see Kaufmann

and Juselius 2013, Davidson et al. 2016 or Castle and Hendry 2020 for further details on the construction of the data, as well as Miller, 2019 for the consequences of interpolation on subsequent statistical analysis).




(a) Temperature

(b) $CO_2$

Figure 1: Temperatures and $CO_2$ concentration levels

The most striking feature of the paleoclimate series is their cyclicality, i.e. significant recurring comovements, such that temperature and $CO_2$ levels stay below their mean for long periods during glaciations, with the opposite pattern for ice volume (see Figure 1). As extensively discussed in Davidson et al. (2016) this behaviour makes it difficult to reconcile with the cointegration approach used to study long-run relationships and followed in much of the literature. When we strip away the cycles, how can we quantify with precision the relationship between temperatures and $CO_2$? Given its importance for accurate climate sensitivity projections, our focus is on inference about the long-run relationship between temperature and gases concentrations, so in order to abstract from 'cyclical noise', we propose using the low-frequency 'filtering' procedures of MW.

Indeed, low-frequency variation can be extracted by using relatively small number $q$ of weighted averages, where the weights are deterministic (and known) low-frequency trigonometric series. Consider the simplest case of a single time series $x_t$ observed over $t = 1, ..., T$, letting $\Psi_j(s) = \sqrt{2}cos(js\pi)$, so that $\Psi_j(t/T)$ has period $2T/j$, with $\Psi(s) = (\Psi_1(s), \Psi_2(s), ..., \Psi_q(s))'$ a $\Re^q$ valued function, with $\Psi(T) = (\Psi((1-1/2)/T), \Psi((2-1/2)/T), ..., \Psi((T-1/2)/T))'$ a $T \times q$ matrix obtained by evaluating the function at $s = (t-1/2)/T$ for $t = 1, ..., T$. Then, we can obtain low-frequency projections by obtaining the fitted values from running OLS regressions of $x_t$ on $\Psi(T)$, such that $\hat{x}_t = \bar{x} + \Psi((t-1/2)/T)'X_T$, where $\bar{x}$ and $X_T$ are OLS coefficients, the latter having the simple form $X_T = T^{-1}\sum_{t=1}^{T}\Psi((t-1/2)/T)'x_t$, such that the $j^{th}$ regression coefficient $X_{jt}$ is the $j^{th}$ cosine transform of $(x_1, x_2, ..., x_T)'$.

Müller and Watson (2018) and we in section 3 show that this can be extended to a multivariate

Table 1: Long-Run Covariation Measures for $CO_2$ and Temperature

| $q$ | **LR correlation ($\rho$)** | **LR regression coefficient ($\beta$)** | **LR regression st. error ($\sigma$)** |
|---|---|---|---|
| 10 | 0.865 [0.450, 0.960] | 0.068 [0.025, 0.090] | 0.395 [0.279, 0.809] |
| 11 | 0.828 [0.500, 0.947] | 0.068 [0.026, 0.093] | 0.447 [0.313, 0.811] |
| 12 | 0.841 [0.450, 0.952] | 0.069 [0.028, 0.094] | 0.447 [0.317, 0.864] |
| 13 | 0.841 [0.450, 0.947] | 0.072 [0.028, 0.096] | 0.464 [0.332, 0.866] |
| 14 | 0.865 [0.550, 0.953] | 0.077 [0.040, 0.100] | 0.498 [0.364, 0.825] |
| 15 | 0.877 [0.637, 0.959] | 0.082 [0.048, 0.102] | 0.516 [0.382, 0.874] |
| **16** | **0.945 [0.900, 0.980]** | **0.090 [0.069, 0.100]** | **0.533 [0.386, 0.921]** |
| 17 | 0.937 [0.850, 0.974] | 0.092 [0.073, 0.104] | 0.621 [0.459, 1.013] |
| 18 | 0.940 [0.850, 0.974] | 0.093 [0.072, 0.105] | 0.627 [0.455, 1.005] |
| 19 | 0.943 [0.850, 0.974] | 0.094 [0.072, 0.105] | 0.628 [0.463, 1.029] |
| 20 | 0.943 [0.850, 0.974] | 0.095 [0.075, 0.106] | 0.643 [0.475, 1.033] |
| 21 | 0.937 [0.850, 0.974] | 0.095 [0.074, 0.107] | 0.663 [0.496, 1.123] |

Notes: $x$ is $CO_2$ and $y$ is temperature; results based on the Posterior Median and a Coverage Probability of 0.90 in square brackets

setting, in a way that long-run correlation or regression coefficients can be obtained in a standard way based on the $q$ cosine projections. Considering the scarcity of low-frequency information in the data, it is natural that only a small number of projection coefficients are employed to capture low-frequency variability, in turn leading to a typical "small-sample" problem. An advantage of these procedures is that statistical inference is straightforward and is applicable to both weakly and highly persistent time series, with little requirements regarding error and model assumptions.

## 2.2 Bivariate Long-Run Covariability

We start by examining the long-run covariability of temperatures and $CO_2$ emissions by computing low-frequency correlation and bivariate regression coefficients. From visual inspection of both series, we will assume that our time span of interest is smaller than 150 (150,000 years) and greater than 75 (75,000 years) datapoints. This then corresponds to a range of about $q = 10$ to $q = 21$ cosine transforms, respectively. The maximum span is by default 1,596,000 years. Thus, we take an agnostic view as to what the 'long run' is by not imposing a specific value for $q$, not least because, beyond Milankovitch cyclicality, there is no consensus in climate science on what this should be. Hence, we present results obtained for $q = 10, ..., 21$ in Table 1.

From Table 1, we observe that there is strong evidence that $CO_2$ and temperature are highly correlated over the long run. Indeed, the in-sample long-run correlation of the two series is positive and very large - always above 0.83, achieving a maximum of 0.95 with $q = 16$ and remaining high

thereafter, with the confidence intervals noticeably narrowing for $q > 15$.[1] The estimated long-run regression coefficient is, as expected, positive, thus confirming that an increase in $CO_2$ emissions is associated with an estimated $^{\circ}C$ increase in temperatures in the long-run. Note, however, that the slope coefficient is very sensitive to the choice of $q$, monotonically increasing from 0.068 for $q = 10$ to 0.095 with $q = 21$.

Given this, we define a criterion that allows us to select the most reasonable long-run measures: pick $q$ such that the estimated long-run correlation is the largest and the range for the 90% confidence set is the narrowest. For the pair ($CO_2$, Temperature) it corresponds to $q = 16$ (in bold, Table 1), implying a periodicity smaller than 99.75 (i.e. of about 100,000 years). Also, note that for $q < 16$, the 90% confidence sets for the correlation and the regression coefficient are substantially wider, therefore suggesting that this strategy is a sensible one.

Thus, when $q = 16$, the correlation equals 0.945 with a corresponding 90% confidence set of $[0.900, 0.980]$, while the regression coefficient is 0.090, with a standard error 0.533. This value is somewhat large, especially when compared to the ones found in the literature as, for example, by Castle and Hendry (2020), who estimated an impact of only $0.060^{\circ}C$. The lower bound of the 90% confidence set [0.069, 0.100] is close to the by Castle and Hendry (2020) estimate, but does not include it.

For the sake of completeness, we also computed the long-run measures for the bivariate relationships of ice volume with temperatures or $CO_2$ emissions (see Table A1 in the Appendix). As in the previous case, $q < 16$ delivers wide confidence sets for the pair {Ice, Temperature}. The in-sample long-term correlation for $q = 16$ equals $-0.871$ with a confidence set of $[-0.946, -0.638]$ (similar results are obtained for $q > 16$) and an estimated long-run regression coefficient of $-4.991$. For {Ice, $CO_2$} the correlation is also large ($-0.918$ with an interval of $[-0.945, -0.800]$) and the regression coefficient is $-55.619$, highlighting the long run association between higher $CO_2$ emissions and reduced ice volumes.

## 3 Long-Run Covariability under Multivariate Partialling Out

The value of 0.090 and corresponding 90% confidence set of [0.069, 0.100] for the long-term in-sample regression coefficient of temperature on $CO_2$ discussed above are probably overestimated. The most likely explanation is that the standard bivariate Müller and Watson (2018) approach is not taking into account the effect of exogenous variables on temperatures, namely orbital forcing and non-linear functions thereof. Therefore, in order to be able to assess the long-run covariability between temperature and $CO_2$ emissions without omitting orbital forcing, we extend the Müller

[1] To put it in perspective, it is worth mentioning that in their empirical applications with macro variables, Müller and Watson (2018) found for fixed $q$ only a few number of pairs of series with such a large degree of covariability. Despite the differences in terms of the observed data, it seems that climate data has a more significant long-term covariability compared to macro economic data.

and Watson (2018) approach so that these effects are partialled out. We do so because of the 'small sample' problem induced by using only a small number of trigonometric projections. To avoid compounding this problem by adding additional regressors, and given that in our application we can safely treat orbital forcing as exogenous, we suggest this approach as a straightforward way of addressing this issue. Moreover, as a complementary alternative, we also consider long-run covariability measures between temperature and a common factor capturing the main dynamics of all (or part of) the other variables in the model, i.e. $CO_2$, ice volume and orbital variables. .

### 3.1 Long-Run Projections After Partialling Out

In this Section, we derive the main results for the long-run projections and covariability measures after partialling out the effects of other control variables, thus providing an extension of the original work by Müller and Watson (2017, 2018, 2021) (the Appendix contains further details). The data generating process is given by Assumption 1.

**Assumption 1**: Let $z_t = (y_t', x_t')'$ denote a $(2+k)$-dimensional vector of time series with $y_t = (y_{1t}, y_{2t})'$ and assume that $z_t$ follows a multivariate I(0) model $z_t = \mu + u_t$, such that

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{\lfloor rT \rfloor} u_t \Longrightarrow \Omega^{1/2} W(r), r \in [0,1], \tag{1}$$

where $\Omega = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \Omega_{1x} \\ \sigma_{21} & \sigma_2^2 & \Omega_{2x} \\ \Omega_{x1} & \Omega_{x2} & \Omega_{xx} \end{pmatrix}$ is the $(2+k)\times(2+k)$ long-run covariance matrix of $u_t$ and $W(r)$ is a $(2+k)\times 1$ multivariate standard Wiener process.[2] Here, $\mu = \left(\mu_y', \mu_x'\right)'$, $u_t = (u_{1t}, u_{2t}, u_{xt}')'$, and $W(r) = \left(W_1(r), W_2(r), W_x(r)'\right)'$.

We link $y_{2t}$ with $x_t$ through the model $y_{2t} = \theta' x_t + v_t$ and define

$$\widehat{v}_t = y_{2t} - \widehat{\theta}' x_t = y_{2t} - \left( \left( \sum_{t=1}^{T} y_{2t} x_t' \right) \left( \sum_{t=1}^{T} x_t x_t' \right)^{-1} \right) x_t \tag{2}$$

which is also an I(0) process. For the cosine-weighted averages, define $\Psi_j(s) = \sqrt{2}\cos(js\pi)$, $s \in [0,1]$, denote the function with period $2/j$, let $\Psi(s) = (\Psi_1(s), ..., \Psi_q(s))'$ be a vector $q \times 1$ of the functions $\Psi_j(s)$ from periods 2 through $2/q$, and denote $\Psi_T$ as the $T \times q$ matrix with $t^{th}$ row given by $\Psi((t-1/2)/T)'$, so the $j^{th}$ column of $\Psi_T$ has period $2T/j$. Consequently, $\widehat{y}_{1t} = Y_{1T}' \Psi((t-1/2)/T)$ and $\widetilde{v}_t = \widehat{V}_T' \Psi((t-1/2)/T)$ where $Y_{1T}'$ and $\widehat{V}_T$ are the projection (linear regression) coefficients, namely,

$$\widehat{V}_T = T^{-1} \sum_{t=1}^{T} \Psi((t-1/2)/T) \widehat{v}_t \tag{3}$$

[2] As in Müller and Watson (2017, 2018, 2021), we could consider other different data generating processes with similar results.

(cosine-weighted averages of the residuals).

**Lemma 1**: Under Assumption 1, the average covariance matrix of the long-run projections $(\widehat{y}_{1t}, \widetilde{v}_t)$, which also measures the covariability of the corresponding cosine transforms, is given by:

$$\Omega_T \to \sum_{j=1}^{q} \begin{pmatrix} \sigma_1^2 & \sigma_{1V} \\ \sigma_{1V} & \sigma_V^2 \end{pmatrix}, \tag{4}$$

as $T \to \infty$, where

$$\sigma_{1V} = \sigma_{12} - Q'_{2x} Q_{xx}^{-1} \Omega'_{1x} \Omega_{1x} Q_{xx}^{-1} Q_{2x}. \tag{5}$$

$$\sigma_V^2 = \sigma_2^2 + Q'_{2x} Q_{xx}^{-1} \Omega_{xx} Q_{xx}^{-1} Q_{2x} - 2 Q'_{2x} Q_{xx}^{-1} \Omega'_{2x} \Omega_{2x} Q_{xx}^{-1} Q_{2x} \tag{6}$$

such that

$$\frac{1}{T} \sum_{t=1}^{T} y_{2t} x'_t \to Q'_{2x} = E\left(y_{2t} x'_t\right); \frac{1}{T} \sum_{t=1}^{T} x_t x'_t \to Q_{xx} = E\left(x_t x'_t\right), \tag{7}$$

as $T \to \infty$.

From Lemma 1 we obtain the low-frequency measures between $y_{1t}$ and $\widehat{v}_t$. The low-frequency covariance is defined as the population covariance between the low-frequency trend values $(\widehat{y}_{1t}, \widetilde{v}_t)$ averaged over the length of the sample

$$T^{-1} \sum_{t=1}^{T} E\left(\widehat{y}_{1t} \widetilde{v}_t\right) \to q \sigma_{1V}. \tag{8}$$

On the other hand, the low-frequency correlation between $y_{1t}$ and $\widehat{v}_t$ is

$$\rho_T = \frac{tr\left[E\left(Y_{1jT} \widehat{V}_{jT}\right)\right]}{\sqrt{tr\left[E\left(Y_{1T}^2\right)\right] tr\left[E\left(\widehat{V}_{jT}^2\right)\right]}} \to \frac{\sigma_{1V}}{\sigma_1 \sigma_V}. \tag{9}$$

The long-run linear regression parameter in the regression of $\widehat{y}_{1t}$ on $\widetilde{v}_t$ (best linear prediction of the long-run projection $\widehat{y}_{1t}$ by the long-run projection $\widetilde{v}_t$) is

$$\beta_T = \arg\min_b E\left[T^{-1} \sum_{t=1}^{T} \left(\widehat{y}_{1t} - b \widetilde{v}_t\right)^2\right] \tag{10}$$

$$= \frac{tr\left[E\left(Y_{1jT} \widehat{V}_{jT}\right)\right]}{tr\left[E\left(\widehat{V}_{jT}^2\right)\right]} \to \frac{\sigma_{1V}}{\sigma_V^2} \tag{11}$$

and the average variance of the prediction error is

$$\sigma^2_{y_1|\widehat{v},T} = tr\left[E\left(Y_{1T}^2\right)\right] - \frac{tr\left[E\left(Y_{1jT} \widehat{V}_{jT}\right)\right]^2}{tr\left[E\left(\widehat{V}_{jT}^2\right)\right]} \tag{12}$$

$$\to q\left(\sigma_1^2 - \frac{\sigma_{1V}^2}{\sigma_V^2}\right) \tag{13}$$

These are the parameters we estimate in our empirical analysis, which measure the population linear dependence of the long-run variation of $(y_{1t}, \widehat{v}_t)$.

### 3.1.1 Empirical Results

In this subsection, we apply the partialling out approach in long-run covariability methods to the paleoclimate data analysed above. The variable of interest is temperature (our $y_{1t}$ term) and how it depends on $CO_2$ emissions (the $y_{2t}$ variable) in the long-run, after extracting the effects of other determinants ($x_t$ vector). The choice for $x_t$ is not clear-cut, therefore we follow the existing literature and try different options to check the robustness and sensitivity of the results to the choice made.

As in Castle and Hendry (2020), we assume that $CO_2$ depends, at most, on lagged temperatures and ice volume and the current and lagged non-linear impacts of the exogenous orbital variables, i.e., the largest set for $x_t$ is

$$x_t = (Temp_{t-1}, Ice_{t-1}, O_t, O_{t-1}), \text{ where} \tag{14}$$

$$O = (Ec, Ob, Pr, EcOb, EcPr, PrOb, Ec^2, Ob^2, Pr^2). \tag{15}$$

For comparison, we also consider smaller sets by dropping those variables that were not found statistically significant in explaining $CO_2$ levels in previous studies. Interestingly enough, we did not find important differences in the OLS residuals $\widehat{v}_t$ compared to the observed $CO_2$ emissions (see Figure 1) under the regression model $CO_{2t}=\theta' x_t + v_t$. In Figure 2, we have $\widehat{v}_{1t}$ for $x_t = O_t$ (left panel) and $\widehat{v}_{2t}$ for $x_t = \left(Ec, Ob, EcOb, EcPr, Ob^2\right)_t$ (right panel).

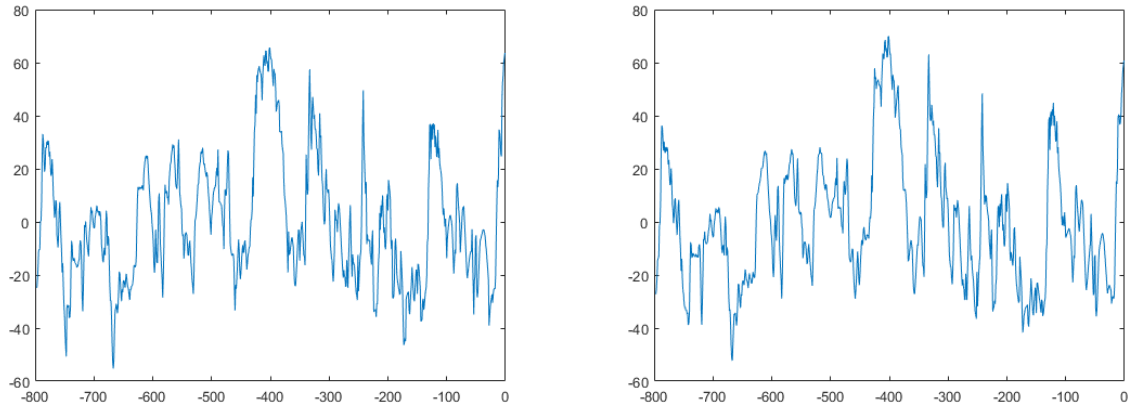
In Table 2, we present the long-run covariability measures for temperatures and (distinct) $CO_2$ residuals $\widehat{v}_t$ for the 'optimal' $q$, in the sense that the correlation is the largest and the range for the confidence set is the narrowest.[3] The results differ depending on whether lagged temperatures and ice volume as determinants for $CO_2$ emissions are included or not. If included, we only find a significant long-run relationship between temperatures and partialled out $CO_2$ for a confidence set of 67%. Even in that case, the long-run correlation is not large (0.317) and the long-run regression coefficient equals 0.108. Once we include only the exogenous orbital variables, the results become very consistent. The long-run correlation is large (around 0.8) and the long-run regression coefficient is about 0.076 with a 90% confidence set of [0.048, 0.103].

[3]Results for other values of $q$ produce qualitatively and quantitatively similar results, and are available upon request.

Table 2: Long-Run Covariation Measures for $v_t$ and $Temperature$

| $q$ | **LR correlation ($\rho$)** | **LR regression coefficient ($\beta$)** | **LR regression st. error ($\sigma$)** |
|---|---|---|---|
| $x_t = (Temp_{t-1}, Ice_{t-1}, O_t, O_{t-1})$ | | | |
| 16 | 0.317 [-0.129, 0.539] | 0.108 [-0.014, 0.254] | 1.680 [1.265, 2.440] |
| | 0.317 (0.131, 0.443) | 0.108 (0.036, 0.188) | 1.680 (1.413, 2.065) |
| $x_t = (O_t, O_{t-1})$ | | | |
| 16 | 0.744 [0.502, 0.917] | 0.077 [0.048, 0.105] | 1.009 [0.762, 1.479] |
| $x_t = O_t$ | | | |
| 16 | 0.744 [0.511, 0.917] | 0.078 [0.048, 0.106 ] | 1.012 [0.847, 1.245] |
| $x_t = \left(Ec, Ob, EcOb, EcPr, Ob^2\right)_t$ | | | |
| 16 | 0.804 [0.564, 0.918] | 0.076 [0.048, 0.103] | 0.980 [0.733, 1.425] |
| $x_t = (Ec_t, Ec_{t-1}, Ob_{t-1}, EcOb_{t-1}, Ob_t^2)$ | | | |
| 16 | 0.798 [0.512, 0.917] | 0.075 [0.046, 0.102] | 1.002 [0.752, 1.463] |

Notes: Results based on the Posterior Median and a Coverage Probability of 0.90 in square brackets or 0.67 in parentheses




(a) $\widehat{v}_{1t}$ (b) $\widehat{v}_{2t}$

Figure 2: Residuals $\widehat{v}_{1t}$ and Residuals $\widehat{v}_{2t}$

The main takeaways are that 1) regardless of whether or not orbital effects are partialled out, the long-run correlation between temperatures and $CO_2$ emissions is very large (above 0.7); 2) with respect to the long-run impact of $CO_2$ emissions on temperature levels, we obtain larger point estimates than those reported in Castle and Hendry (2020), which can be rather interpreted as a lower bound. Castle and Hendry (2020) obtain a solved long-run coefficient of $0.066°C$ increase in temperatures for a 1 ppm increase in the $CO_2$ emissions, whereas our long-run estimated impact ranges from about $0.070°C$ to $0.100°C$, depending on the exact specification.

### 3.2 Using Common Factors as Controls

An alternative to partialling out using the OLS residuals is to control for exogenous forcing by extracting common factors using principal components analysis. For comparability, we only study common factors extracted from the data for which there is evidence that it determines the temperature levels. Thus, following Castle and Hendry (2020), we assume that temperatures in period $t$ are explained by the set of variables $x_t = (Temp_{t-1}, CO_{2t}, Ec_t, EcOb_t, EcOb_{t-1}, EcPr_t)$ and formulate a factor model in which low-frequency measures of $Temp_t$ are related with the main common factor of $x_t$,, denoted as $\widehat{f}_t$. In Table 3, we present the results for $\widehat{f}_t$ extracted from $x_t$, also considering different subsets of its elements, using $q = 16$ as before (results are similar for any $q > 15$).

Table 3: Long-Run Covariation Measures for $\widehat{f}_t$ and $Temperature$

| $q$ | **LR correlation ($\rho$)** | **LR regression coefficient ($\beta$)** | **LR regression st. error ($\sigma$)** |
|---|---|---|---|
| $x_t = (Temp_{t-1}, CO_{2t}, Ec_t, EcOb_t, EcOb_{t-1}, EcPr_t)$ | | | |
| $PTV = 97.688$; $\widehat{\lambda} = (0.102, 0.994, 0.000, 0.003, 0.003, 0.027)$ | | | |
| 16 | 0.945 [0.900, 0.980] | 0.090 [0.069, 0.099] | 0.528 [0.383, 0.912] |
| $x_t = (CO_{2t}, Ec_t, EcOb_t, EcPr_t)$ | | | |
| $PTV = 97.914$; $\widehat{\lambda} = (0.999, 0.000, 0.003, 0.027)$ | | | |
| 16 | 0.945 [0.900, 0.980] | 0.090 [0.069, 0.100] | 0.532 [0.386, 0.920] |
| $x_t = (CO_{2t-1}, CO_{2t}, Ice_{t-1}, Ice_t)$ | | | |
| $PTV = 98.712$; $\widehat{\lambda} = (0.706, 0.708, -0.010, -0.010)$ | | | |
| 16 | 0.943 [0.900, 0.974] | 0.063 [0.048, 0.071] | 0.558 [0.404, 0.962] |
| $x_t = (CO_{2t}, Ice_t)$ | | | |
| $PTV = 99.999$; $\widehat{\lambda} = (0.999, -0.015)$ | | | |
| 16 | 0.945 [0.900, 0.980] | 0.090 [0.069, 0.100] | 0.533 [0.387, 0.922] |

Notes: Results based on the Posterior Median and a Coverage Probability of 0.90 in square brackets. $PTV$ is the percentage of the total variance explained by the principal component. $\widehat{\lambda}$ is the principal component coefficients (loadings)

First, we note that the percentage of the variance (PTV) explained by the principal component is almost 100%, even for the largest of sets $x_t$. On the other hand, analysing the corresponding factor loadings (displayed in $\widehat{\lambda}$), the loading associated with $CO_{2t}$ is essentially one, while the other loadings are quite small and close to zero. These two results together mean that the principal component $\widehat{f}_t$ captures the dynamic features of $CO_{2t}$ levels after extracting the information from the remaining variables. Thus, in essence, we accomplish the goal of obtaining low-frequency estimates of the relationship between temperatures and $CO_2$ levels in a much similar way to the partialling out procedure expounded above.

Thus, considering the results so far, our estimates point to (Antarctic) temperature long

run increases between $6°C$ and $10°C$. These are in the neighbourhood of, but larger than, the estimates of Castle and Hendry (2020), as well as those of Kaufmann and Juselius (2013) (see also Knutti, Rugenstein and Hegerl, 2017). In section 4, we return to this type of scenario analysis by considering long run forecasts obtained from low-frequency multivariate factor models.

# 4 Low-Frequency Factor Model and Long-Term Forecasts

In the previous sections, we focused on the long run relationship between temperatures and $CO_2$ concentration levels - first, a stripped down low-frequency bivariate analysis, followed by approaches that allow us to control for exogenous forcing. In this section we turn our attention to modelling the joint dynamics of the whole climate system by employing the low-frequency factor models developed by Müller and Watson (2021), which, as a useful by-product, allow us to compute long-range forecasts.[4] Indeed, given the very low-frequency nature of our data, as well as the mixture of endogenous and exogenous variables we study, we deem this forecasting model to be more appropriate than a standard dynamic factor model. Next we present the main features of the model and analyse the estimated factor loadings, after which we focus on long-horizon predictive distributions for the variables of interest, under relevant scenarios for $CO_2$ levels.

## 4.1 Factor Model Estimation

Following Müller and Watson (2021) and Müller, Stock and Watson (2021), the model for the observed $x_t \in \Re^n$ is

$$x_t = \mu + \lambda f_t + e_t, \tag{16}$$

where $f_t$ denotes the (scalar) unobserved common factor(s), $\lambda = (\lambda_1, ..., \lambda_n)$ contains the factor loadings, $e_t$ represents a vector of mutually independent errors that captures the residual variability in $x_t$, and $\mu = (\mu_1, ..., \mu_n)$ is the intercept. Moreover, $f_t$ follows a local-level model, $e_{j,t}$ , $j = 1, ..., n$ stationary $I(d_j)$ models and $\{f_t, e_{1,t}, ..., e_{n,t}\}$ are independent. The local-level model is the sum of uncorrelated I(0) and I(1) processes with common long-run variance $\sigma^2$. The scale of $f_t$ and $\lambda$ are not separately identified and the factor loading $\lambda_1$ is normalized to unity. The estimation of the model is carried out by Bayesian methods, assuming the same priors for the parameters as in Müller and Watson (2021) (see paper for details). Once the posterior distribution is recovered, an additional advantage of this approach is that we can analyse how the relationships differ across quantiles - we consider the median, the tails at 5% and 95% and two intermediate quantiles, 17% and 83%.

We consider three models: 1) a simple bivariate one with the main variables $x_t = (Temp_t, CO_{2t})'$ with $n = 2$; 2) a model that adds ice volume, i.e. $x_t = (Temp_t, CO_{2t}, Ice_t)'$, $n = 3$; and 3) a

[4] See also Müller and Watson (2016) for a discussion on long-horizon prediction intervals and Müller, Stock and Watson (2021) for an interesting application about the long-run path of GDP for a list of 113 countries.

model that also includes exogenous orbital variables, such that $x_t = (Temp_t, CO_{2t}, Ice_t, Ec_t, Ob_t, EcPr_t, Ob_t^2)'$, $n = 7$.

The results are displayed in Table 4. We can see that $CO_2$ is a strong driver of low-frequency comovements with temperatures, particularly at the highest quantiles. Interestingly, the salience of $CO_2$ remains strong even when orbital variables are included. Indeed, as expected and in accordance with the Milankovich hypothesis, orbital variables capture a significant proportion of the joint variation, but $CO_2$ levels still have a non-negligible impact for quantiles 83 and 95. As anticipated, variation in ice volume goes in an opposite direction, while the joint effects of nonlinear interactions amongst the orbital variables appear to be the most significant contributors to low-frequency movements in the respective set of variables in $x_t$.

Table 4: Factor Loadings for the Low-Frequency Factor Model

| **Variables** | **PostMean** | **PostQ0.05** | **PostQ0.17** | **PostQ0.50** | **PostQ0.83** | **PostQ0.95** |
|---|---|---|---|---|---|---|
| $x_t = (Temp_t, CO_{2t})'$; $n = 2$ | | | | | | |
| $Temp$ | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| $CO_2$ | 9.29 | 7.98 | 8.66 | 9.37 | 9.96 | 10.37 |
| $x_t = (Temp_t, CO_{2t}, Ice_t)'$; $n = 3$ | | | | | | |
| $Temp$ | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| $CO_2$ | 9.30 | 8.00 | 8.66 | 9.37 | 9.97 | 10.39 |
| $Ice$ | -0.15 | -0.19 | -0.17 | -0.15 | -0.12 | -0.11 |
| $x_t = (Temp_t, CO_{2t}, Ice_t, Ec_t, Ob_t, EcPr_t, Ob_t^2)'$; $n = 7$ | | | | | | |
| $Temp$ | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| $CO_2$ | 1.50 | -1.78 | -0.40 | 1.50 | 3.42 | 4.79 |
| $Ice$ | 1.36 | -1.92 | -0.56 | 1.33 | 3.35 | 4.71 |
| $Ec$ | 1.35 | -1.99 | -0.66 | 1.37 | 3.33 | 4.65 |
| $Ob$ | 0.78 | -1.64 | -0.62 | 0.56 | 2.32 | 3.76 |
| $EcPr$ | 1.68 | -1.70 | -0.34 | 1.59 | 3.63 | 5.43 |
| $Ob^2$ | 1.44 | -1.85 | -0.49 | 1.43 | 3.36 | 4.74 |

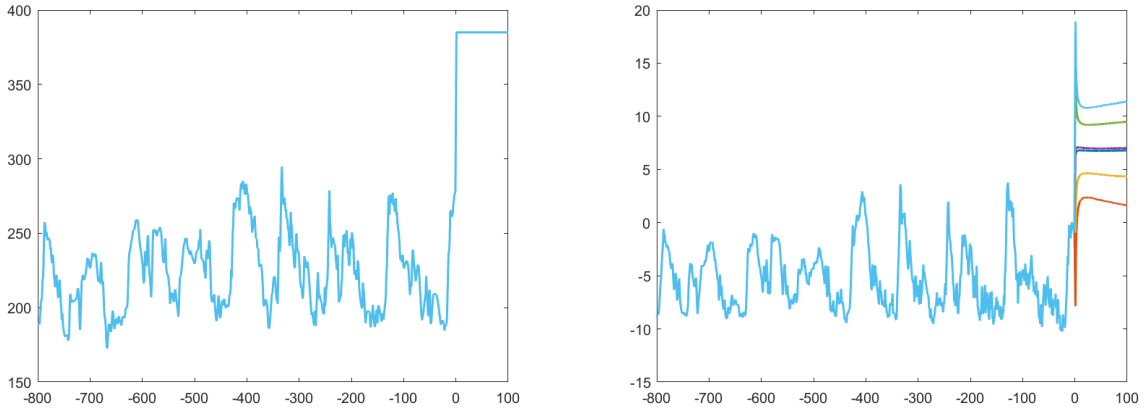
Notes: PostQ stands for "posterior quantile"

## 4.2 Long-Horizon Projections

Having documented and quantified the low-frequency relationship between temperatures and $CO_2$ levels, and given the recent developments in the climate debate, it is important to understand how temperatures can evolve in the long-term for specific scenarios of $CO_2$ emissions. Here, we follow Castle and Hendry (2020) by conditioning on the relatively stable path of orbital variables and on a stable level of 385 ppm for $CO_2$ (consistent with current anthropogenically induced levels) to obtain long-term conditional forecasts for temperatures and ice volume using the low-frequency

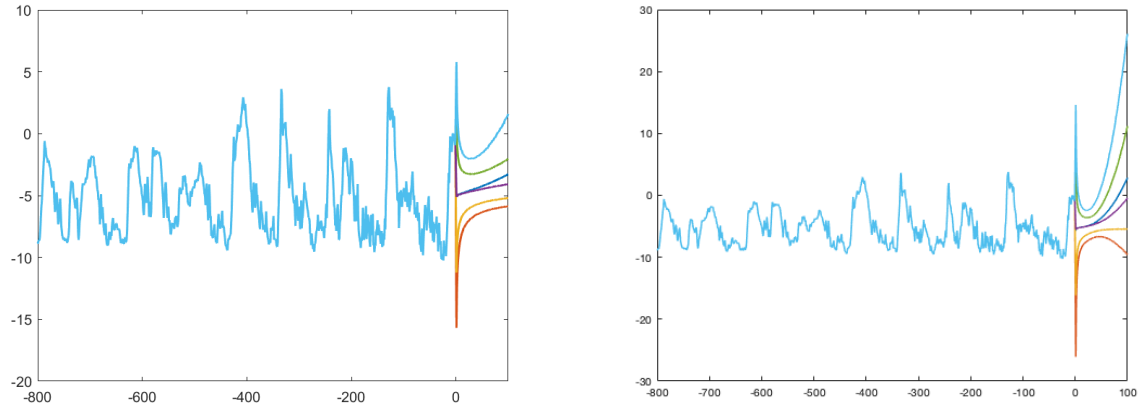
factor models estimated in the previous section. We consider the horizon of 100 observations (100,000 years, the lowest period associated with the optimal $q$ as explained in Section 2).

In Figures 3 and 4 we have the conditional forecasts for temperatures, while Figure 5 displays projections for ice volumes. In the case of $n = 2$, i.e. with $x_t = (Temp_t, CO_{2t})'$, temperatures are predicted to remain well above the peak in the sample. This can be expected, as we can see from Figure 3a that the value of $CO_2$, on which we are conditioning, is also well above historical levels. Once other variables are included in $x_t$ ($n = 3$ and $n = 7$), the projections for temperatures are less extreme, closer to the sample mean, but, crucially, predicted to increase.




(a) $CO_2$ (b) Temperatures

Figure 3: CO2 Levels and Conditional Forecasts of Temperatures ($n = 2$)

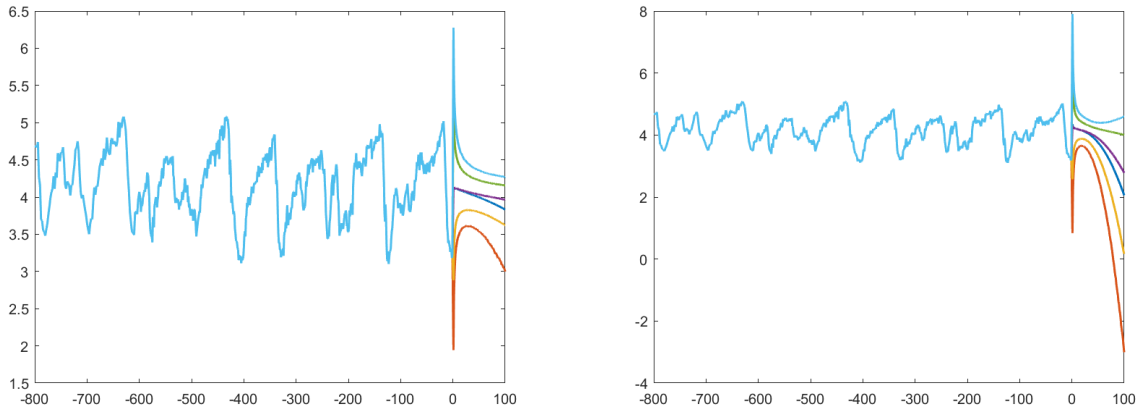



(a) $n = 3$ (b) $n = 7$

Figure 4: Conditional Forecasts of Temperatures

Turning to the long term of implications for ice volumes, we confirm the prediction that ice sheets will tend to reduce for high levels of $CO_2$ (naturally in tandem with the higher temperatures predicted above). Interestingly, the speed at which ice volumes vanish increases when other variables are included in the dynamic system ($n = 3$ and $n = 7$), with (median) ice volume levels

attaining values similar (or even below) to sample minima (see Diebold and Rudebusch, 2022 for a discussion, albeit at shorter horizons).[5]




(a) $n = 3$ (b) $n = 7$

Figure 5: Conditional Forecasts of Ice Volumes

## 5 Conclusion

This paper explores formal inference procedures that focus on low-frequency comovement amongst paleoclimate time series. By employing the recently developed methods of Müller and Watson (2017, 2018, 2021), we attempt to abstract from the strong cyclical features of the data, spurred mainly by orbital forcings, and focus on relationships over longer periods. This approach involves obtaining weighted averages (constructed by means of trigonometric projections) and then employing these to conduct inference in a small-sample, yet standard, setting. Given that complete consideration of all drivers of temperatures aggravates the small-sample issue, we extend the above mentioned framework by allowing for a first round of 'partialling out' of exogenous forcings.

Our results confirm the strong relationship between temperatures and $CO_2$ for a periodicity of around 100,000 years. The long-run coefficient thus obtained validates, to a large extent, previous equilibrium climate sensitivity exercises conducted by several authors. If anything, our findings suggest that the recent simulations of Castle and Hendry (2020) are a lower bound, indicating that warming due to the increase of $CO_2$ concentration levels can be even more substantial. This holds even when orbital forcings are taken into account, either by partialling out or by means of an extracted common factor. Joint modelling of all time series allows us to consider a long-term scenario exercise, in which we show the extent of acceleration in temperature increases, as well as ice-sheet recession.

[5]Note that we do not constrain the quantiles, hence the negative values for Q95.

In this setup, the choice of $q$ is usually left at the researchers' discretion. We proposed a 'data-driven' approach, whereby we picked $q$ such that the target long-run correlation is maximized. An interesting issue to consider is the choice of $q$ when two or more time series display different cyclicality - it is not clear whether different $q$'s should be selected for each series, or whether a common $q$ would be preferable. We leave this for future research.

## References


[1] Baillie, R. T. and Chung, S.-K. (2002), Modeling and forecasting from trend-stationary long-memory models with applications to climatology, *International Journal of Forecasting*, 18, 215–226.

[2] Bierens, H.J. (1994), *Topics in Advanced Econometrics: Estimation, Testing and Specification of Cross-Section and Time Series Models*, Cambridge University Press.

[3] Bloomfield, P. (1992), Trends in Global Temperature, *Climatic Change*, 21, 1–16.

[4] Bloomfield, P. and Nychka, D. (1992), Climate spectra and detecting climate change *Climatic Change*, 21, 275–287.

[5] Castle, J. L. and Hendry, D. F. (2020), *Climate Econometrics: An Overview*, Foundations and Trends in Econometrics: Vol. 10, No. 3–4, 145–322.

[6] Davidson, J. E. H., Stephenson, D. B. and Turasie, A. A. (2016), Time series modeling of paleoclimate data, *Environmetrics*, 27, 55-65.

[7] Diebold, F.X. and Rudebusch, G.D. (2022), Probability Assessments of an Ice-Free Arctic: Comparing Statistical and Climate Model Projections, *Journal of Econometrics*, forthcoming.

[8] Fomby, T. B. and Vogelsang, T. J. (2002), The Application of Size-Robust Trend Statistics to Global-Warming Temperature Series, *Journal of Climate*, 15, 117-123.

[9] Gordon, A. H. (1991), Global warming as a manifestation of a random walk, *Journal of Climate*, 4 , 589–597.

[10] Gordon, J.D., Boudreau, P.R., Mann, K.H., Ong, J.-E., Silvert, W.L., Smith, S.V., Wattayakorn, G., Wulff, F. and Yanagi, T. (1996), LOICZ Biogeochemical modelling Guidelines. LOICZ Reports and Studies No. 5.

[11] Jaccard, S. L., Galbraith, E. D., Martínez-García, A. and Anderson, R. F. (2016), Covariation of deep Southern Ocean oxygenation and atmospheric CO2 through the last ice age, *Nature* 530, 207–210.

[12] Jouzel, J., Masson-Delmotte, V., Cattani, O., Dreyfus, G., Falourd, S. and Hoffmann, G. (2007), Orbital and millennial Antarctic climate variability over the past 800,000 years, *Science* 317, 793–796.

[13] Kaufmann, R.K. and Juselius, K. (2013), Testing hypotheses about glacial cycles against the observational record, *Paleoceanography*, 28, 1–10

[14] Kaufmann, R. K. and Pretis, F. (2020), Testing Hypotheses About Glacial Dynamics and the Stage 11 Paradox Using a Statistical Model of Paleo-Climate, *Climate of the Past: Discussions*, forthcoming.

[15] Kaufmann, R. K. and Pretis, F. (2021), Understanding glacial cycles: A multivariate disequilibrium approach, *Quaternary Science Reviews*, 251, 106694

[16] Kärner, O. (1996), Global Temperature Deviations as a Random Walk, *Journal of Climate*, 9, 656-658.

[17] Knutti, R., Rugenstein, M. A. A. and Hegerl, G. C. (2017), Beyond equilibrium climate sensitivity, *Nature Geoscience*, 10, 727-734.

[18] Lea, D. W. (2004), The 100,000-yr cycle in tropical SST, greenhouse forcing and climate sensitivity, *Journal of Climate*, 17, 2170–2179.

[19] Lisiecki, L. E. and Raymo, M. E. (2005), A Pliocene-Pleistocene stack of 57 globally distributed benthic $\delta^{18}$O records, *Paleoceanography*, 20, PA1003.

[20] Loulergue, L., Schilt, A. and Spahni, R. (2008), Orbital and millennial-scale features of atmospheric CH4 over the past 800,000 years, *Nature*, 453, 383–386.

[21] Miller, J. I. (2019), Testing Cointegrating Relationships Using Irregular and Non-Contemporaneous Series with an Application to Paleoclimate Data, *Journal of Time Series Analysis*, 40, 936-950.

[22] Müller, U. K. and Watson, M. W. (2016), Measuring Uncertainty about Long-Run Predictions, *Review of Economic Studies*, 83, 1711 – 1740.

[23] Müller, U. K. and Watson, M. W. (2017), *Low-Frequency Econometrics*, in Advances in Economics and Econometrics: Eleventh World Congress of the Econometric Society, Volume II, ed. by B. Honoré, and L. Samuelson, Cambridge University Press, 53 – 94.

[24] Müller, U. K. and Watson, M. W. (2018), Long-Run Covariability, *Econometrica*, 86, 775–804.

[25] Müller, U. K. and Watson, M. W. (2021), *Low-Frequency Analysis of Economic Time Series*, Handbook of Econometrics, Vol. 7, ed. by S. Durlauf, L.P. Hansen, J.J. Heckman, and R. Matzkin, forthcoming.

[26] Müller, U. K., Stock, J.H. and Watson, M. W. (2021), An Econometric Model of International Growth Dynamics for Long-Horizon Forecasting, *Review of Economics and Statistics*, forthcoming.

[27] Proietti, T. and Maddanu, F. (2021), Modelling Cycles in Climate Series: The Fractional Sinusoidal Waveform Process, CEIS Working Paper No. 518.

[28] Woodward, W. A. and Gray H. L. (1993), Global warming and the problem of testing for trend in time series data, *Journal of Climate*, 6 , 953–962.

[29] Woodward, W. A. and Gray, H. L. (1995), Selecting a model for detecting the presence of a trend, *Journal of Climate*, 8, 1929–1937.

# 6 Appendix

## 6.1 Bivariate Long-run Covariability

Table A1 - Long-Run Covariation Measures for Ice-Temperature and Ice-$CO_2$

| $q$ | **LR correlation ($\rho$)** | **LR regression coefficient ($\beta$)** | **LR regression st. error ($\sigma$)** |
|---|---|---|---|
| Ice and Temperature | | | |
| 16 | -0.871 [-0.946, -0.638] | -4.991 [-6.190, -3.745] | 0.814 [0.608, 1.196] |
| Ice and $CO_2$ | | | |
| 22 | -0.918 [-0.945, -0.680] | -55.343 [-66.609, -44.553] | 9.891 [7.080, 14.917] |

Notes: In the first row $x$ is Ice and $y$ is temperature, in the second $x$ is Ice and $y$ is $CO_2$; results based on the Posterior Median and a Coverage Probability of 0.90 in square brackets

## 6.2 Proof of Lemma 1

First, we study the properties of the cosine-weighted average of $\widehat{v}_t$. In the main text, we defined $\Psi_j(s) = \sqrt{2}\cos(js\pi)$, $s \in [0,1]$, from which we have the $q \times 1$ vector $\Psi(s) = (\Psi_1(s), ..., \Psi_q(s))'$ and the $T \times q$ matrix $\Psi_T$ with $t^{th}$ row given by $\Psi((t-1/2)/T)'$. The projection of $\widehat{v}_t$ onto $\Psi((t-1/2)/T)$ for $t = 1, ..., T$ yields the fitted values

$$\widetilde{v}_t = \widehat{V}_T' \Psi((t-1/2)/T) \tag{17}$$

where $\widehat{V}_T$ are the projection coefficients

$$\widehat{V}_T = \left(\Psi_T'\Psi_T\right)^{-1} \Psi_T' \widehat{v}_{1:T} = T^{-1} \sum_{t=1}^{T} \Psi((t-1/2)/T)\, \widehat{v}_t, \tag{18}$$

with $\widehat{v}_{1:T} = (\widehat{v}_1, \widehat{v}_2, ..., \widehat{v}_T)'$, because $T^{-1}\Psi_T'\Psi_T = I_q$. Given that $\Psi_T'\iota_T = 0$, the process $\widetilde{v}_t$ corresponds to the projection of $\widehat{v}_t - \overline{\widehat{v}}_{1:T}$ onto $\Psi((t-1/2)/T)$, where $\overline{\widehat{v}}_{1:T}$ is the sample mean of $\widehat{v}$. That is, by including an intercept, $\Psi_T^0 = (\iota_T, \Psi_T)$, we have

$$\widehat{V}_T^0 = \left(\Psi_T^{0\prime}\Psi_T^0\right)^{-1} \Psi_T^{0\prime} \widehat{v}_{1:T} = \begin{pmatrix} T^{-1}\iota_T'\widehat{v}_{1:T} \\ T^{-1}\Psi_T\widehat{v}_{1:T} \end{pmatrix} = \begin{pmatrix} \overline{\widehat{v}}_{1:T} \\ \widehat{V}_T \end{pmatrix}. \tag{19}$$

Morever,

$$\widetilde{v}_t^0 = \widehat{V}_T^{0\prime}\Psi^0((t-1/2)/T) = \overline{\widehat{v}}_{1:T} + \widetilde{v}_t \text{ and} \tag{20}$$

$$\widetilde{v}_{1:T}^0 = \Psi_T^0\widehat{V}_T^0 = I_T\overline{\widehat{v}}_{1:T} + \Psi_T\widehat{V}_T. \tag{21}$$

In the case of $\widehat{v}_t$ being the OLS residuals of $y_{2t}$ on $x_t$, with a zero mean $\widehat{v}_t$ and therefore $\overline{\widehat{v}}_{1:T} = 0$,

$$\widetilde{v}_t^0 = \widetilde{v}_t \text{ and } \widehat{V}_T^0 = \widehat{V}_T. \tag{22}$$

In fact, even if we have $\widehat{v}_t^0 = \eta + \widehat{v}_t$, the $\eta$ has no effect on $\widehat{V}_T^0$ because, by $\sum_{t=1}^T \Psi\left(\left(t-1/2\right)/T\right) = 0$,

$$T^{-1}\sum_{t=1}^{T}\Psi\left(\left(t-1/2\right)/T\right)\widehat{v}_t^0 = T^{-1}\sum_{t=1}^{T}\Psi\left(\left(t-1/2\right)/T\right)\widehat{v}_t. \tag{23}$$

Next, we derive the large-sample properties of $\widehat{V}_T$ (the $q$ cosine transforms of $\widehat{v}_{1:T}$). For the $j$th cosine transform, the scaled $\widehat{V}_{jT}$ is

$$\begin{aligned} & T^{1/2}\widehat{V}_{jT} \\ = \;& T^{-1/2}\sum_{t=1}^{T}\Psi_j\left(\left(t-1/2\right)/T\right)\widehat{v}_t \\ = \;& T^{-1/2}\sum_{t=1}^{T}\sqrt{2}\cos\left(j\pi\left(t-1/2\right)/T\right)y_{2t} \\ & -\left(\sum_{t=1}^{T}y_{2t}x_t'\right)\left(\sum_{t=1}^{T}x_tx_t'\right)^{-1}T^{-1/2}\sum_{t=1}^{T}\sqrt{2}\cos\left(j\pi\left(t-1/2\right)/T\right)x_t. \end{aligned} \tag{24}$$

Following Müller and Watson (2017, 2018, 2021), cf Bierens (1994, Lem. 9.6.3, p. 200),

$$\begin{aligned} & T^{-1/2}\sum_{t=1}^{T}\sqrt{2}\cos\left(j\pi\left(t-1/2\right)/T\right)y_{2t} \\ = \;& T^{-1/2}\sum_{t=1}^{T}\sqrt{2}\cos\left(j\pi\left(t-1/2\right)/T\right)u_{2t} \\ = \;& \frac{\sin\left(j\pi/(2T)\right)}{j\pi/(2T)}\left[\Psi_j\left(1\right)G_{Tu_2}\left(1\right)-\int_0^1\psi_j\left(s\right)G_{Tu_2}\left(s\right)ds\right] \\ \Longrightarrow \;& \Psi_j\left(1\right)G_{u_2}\left(1\right)-\int_0^1\psi_j\left(s\right)G_{u_2}\left(s\right)ds = \int_0^1\Psi_j\left(s\right)dG_{u_2}\left(s\right), \end{aligned} \tag{25}$$

as $T\to\infty$, where $\psi_j\left(s\right) = \frac{\partial\Psi_j(s)}{\partial s}$ and

$$\begin{aligned} G_{Tu_2}\left(s\right) \;&=\; T^{-1/2}\sum_{t=1}^{\lfloor sT\rfloor}u_{2t} + T^{-1/2}\left(sT-\lfloor sT\rfloor\right)u_{2,\lfloor sT\rfloor+1} \\ &\Longrightarrow\; G_{u_2}\left(s\right) = \sigma_2 W_2\left(s\right), \end{aligned} \tag{26}$$

as $T\to\infty$. Similarly,

$$\begin{aligned} & T^{-1/2}\sum_{t=1}^{T}\sqrt{2}\cos\left(j\pi\left(t-1/2\right)/T\right)x_t \\ \Longrightarrow \;& \Psi_j\left(1\right)G_{u_x}\left(1\right)-\int_0^1\psi_j\left(s\right)G_{u_x}\left(s\right)ds = \int_0^1\Psi_j\left(s\right)dG_{u_x}\left(s\right), \end{aligned} \tag{27}$$

as $T\to\infty$, where $G_{u_x}\left(s\right) = \Omega_{xx}^{1/2}W_x\left(s\right)$. Hence,

$$\begin{aligned} T^{1/2}\widehat{V}_{jT} \;&\Longrightarrow\; \int_0^1\Psi_j\left(s\right)dG_{u_2}\left(s\right) - Q_{2x}Q_{xx}^{-1}\int_0^1\Psi_j\left(s\right)dG_{u_x}\left(s\right) \\ &=\; \sigma_2\int_0^1\Psi_j\left(s\right)dW_2\left(s\right) - Q_{2x}Q_{xx}^{-1}\Omega_{xx}^{1/2}\int_0^1\Psi_j\left(s\right)W_x\left(s\right), \end{aligned} \tag{28}$$

as $T \to \infty$, where

$$\frac{1}{T}\sum_{t=1}^{T} y_{2t}x_t' \to Q_{2x}' = E\left(y_{2t}x_t'\right); \quad \frac{1}{T}\sum_{t=1}^{T} x_t x_t' \to Q_{xx} = E\left(x_t x_t'\right), \tag{29}$$

as $T \to \infty$, a $1 \times k$ vector and a $k \times k$ matrix, respectively. This representation holds jointly for the $q$ elements $\widehat{V}_T$, so

$$\begin{aligned} T^{1/2}\widehat{V}_T \implies V &= \int_0^1 \Psi(s)\, dG_{u_2}(s) - \int_0^1 \Psi(s)\, Q_{2x}' Q_{xx}^{-1} dG_{u_x}(s) \\ &= \sigma_2 \int_0^1 \Psi(s)\, dW_2(s) - \int_0^1 \Psi(s)\, Q_{2x}' Q_{xx}^{-1} \Omega_{xx}^{1/2} dW_x(s). \end{aligned} \tag{30}$$

The limiting distribution of $T^{1/2}\widehat{V}_T$ is known. Let $V = V_1 - V_2$ with

$$V_1 = \sigma_2 \int_0^1 \Psi(s)\, dW_2(s), \tag{31}$$

$$V_2 = \int_0^1 \Psi(s)\, Q_{2x}' Q_{xx}^{-1} \Omega_{xx}^{1/2} dW_x(s). \tag{32}$$

As in Müller and Watson (2017, 2018, 2021), under I(0),

$$V_1 \sim N(0, \Sigma_1), \Sigma_1 = \sigma_2^2 I_q. \tag{33}$$

For the case of $V_2$, we have

$$\begin{aligned} & Var\left[\left(\sum_{t=1}^{T} y_{2t}x_t'\right)\left(\sum_{t=1}^{T} x_t x_t'\right)^{-1} T^{-1/2}\sum_{t=1}^{T}\Psi_j\left(\left(t-1/2\right)/T\right)x_t\right] \\ = & Var\left[T^{-1/2}\sum_{t=1}^{T}\Psi_j\left(\left(t-1/2\right)/T\right)\left(\sum_{t=1}^{T} y_{2t}x_t'\right)\left(\sum_{t=1}^{T} x_t x_t'\right)^{-1} u_{xt}\right] \\ = & \left(\sum_{t=1}^{T} y_{2t}x_t'\right)\left(\sum_{t=1}^{T} x_t x_t'\right)^{-1}\widehat{\Omega}_{xx}\left(\sum_{t=1}^{T} x_t x_t'\right)^{-1}\left(\sum_{t=1}^{T} y_{2t}x_t'\right)' T^{-1}\sum_{t=1}^{T}\Psi_j\left(\left(t-1/2\right)/T\right)^2 \end{aligned} \tag{34}$$

plus an $o_p(1)$ term, so that jointly

$$\begin{aligned} & Var\left[\left(\sum_{t=1}^{T} y_{2t}x_t'\right)\left(\sum_{t=1}^{T} x_t x_t'\right)^{-1} T^{-1/2}\sum_{t=1}^{T} x_t\Psi\left(\left(t-1/2\right)/T\right)\right] \\ = & \left(\sum_{t=1}^{T} y_{2t}x_t'\right)\left(\sum_{t=1}^{T} x_t x_t'\right)^{-1}\widehat{\Omega}_{xx}\left(\sum_{t=1}^{T} x_t x_t'\right)^{-1}\left(\sum_{t=1}^{T} y_{2t}x_t'\right)' T^{-1}\Psi_T'\Psi_T + o_p(1) \\ = & \left(\frac{1}{T}\sum_{t=1}^{T} y_{2t}x_t'\right)\left(\frac{1}{T}\sum_{t=1}^{T} x_t x_t'\right)^{-1}\widehat{\Omega}_{xx}\left(\frac{1}{T}\sum_{t=1}^{T} x_t x_t'\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^{T} y_{2t}x_t'\right)' I_q + o_p(1) \end{aligned} \tag{35}$$

and therefore

$$V_2 \sim N(0, \Sigma_2), \Sigma_2 = Q_{2x}' Q_{xx}^{-1}\Omega_{xx}Q_{xx}^{-1}Q_{2x}I_q. \tag{36}$$

If $u_{xt}$ and $u_{2t}$ are not independent ($\Omega_{2x} \neq 0$), the nonzero $Cov(V_1, V_2)$ follows from

$$\begin{aligned}
& Cov\left(\begin{array}{c} T^{-1/2}\sum_{t=1}^{T}\Psi_j\left(\left(t-1/2\right)/T\right)u_{2t}, \\ T^{-1/2}\sum_{t=1}^{T}\Psi_j\left(\left(t-1/2\right)/T\right)\left(\sum_{t=1}^{T}y_{2t}x_t'\right)\left(\sum_{t=1}^{T}x_tx_t'\right)^{-1}u_{xt}\end{array}\right) \\
= & T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T}\Psi_j\left(\left(t-1/2\right)/T\right)\Psi_j\left(\left(s-1/2\right)/T\right)Cov\left(u_{2t},\left(\sum_{t=1}^{T}y_{2t}x_t'\right)\left(\sum_{t=1}^{T}x_tx_t'\right)^{-1}u_{xs}\right) \\
= & \left(\sum_{t=1}^{T}y_{2t}x_t'\right)\left(\sum_{t=1}^{T}x_tx_t'\right)^{-1}\widehat{\Omega}_{2x}'\widehat{\Omega}_{2x}\left(\sum_{t=1}^{T}x_tx_t'\right)^{-1}\left(\sum_{t=1}^{T}y_{2t}x_t'\right)'T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T}\Psi_j\left(\left(t-1/2\right)/T\right)\Psi_j\left(\left(s-1/2\right)/\right.
\end{aligned}$$

so that jointly

$$\begin{aligned}
& Cov\left(\begin{array}{c} T^{-1/2}\sum_{t=1}^{T}\Psi\left(\left(t-1/2\right)/T\right)u_{2t}, \\ T^{-1/2}\sum_{t=1}^{T}\Psi\left(\left(t-1/2\right)/T\right)\left(\sum_{t=1}^{T}y_{2t}x_t'\right)\left(\sum_{t=1}^{T}x_tx_t'\right)^{-1}u_{xt}\end{array}\right) \\
= & \left(\sum_{t=1}^{T}y_{2t}x_t'\right)\left(\sum_{t=1}^{T}x_tx_t'\right)^{-1}\widehat{\Omega}_{2x}'\widehat{\Omega}_{2x}\left(\sum_{t=1}^{T}x_tx_t'\right)^{-1}\left(\sum_{t=1}^{T}y_{2t}x_t'\right)'I_q. \qquad (38)
\end{aligned}$$

which converges to

$$Q_{2x}'Q_{xx}^{-1}\Omega_{2x}'\Omega_{2x}Q_{xx}^{-1}Q_{2x}I_q. \qquad (39)$$

To sum up, we have multivariate normality for the limiting distribution of $T^{1/2}\widehat{V}_T$ with $\widehat{V}_T \overset{a}{\sim} N\left(0, T^{-1}\Sigma\right)$,

$$\begin{aligned}
V &\sim N\left(0,\Sigma\right), \text{ where} && (40)\\
\Sigma &= \left(\sigma_2^2 + Q_{2x}'Q_{xx}^{-1}\Omega_{xx}Q_{xx}^{-1}Q_{2x} - 2Q_{2x}'Q_{xx}^{-1}\Omega_{2x}'\Omega_{2x}Q_{xx}^{-1}Q_{2x}\right)I_q && (41)
\end{aligned}$$

and

$$\begin{aligned}
V_j &\sim iidN\left(0,\sigma_V^2\right), \text{ where} && (42)\\
\sigma_V^2 &= \sigma_2^2 + Q_{2x}'Q_{xx}^{-1}\Omega_{xx}Q_{xx}^{-1}Q_{2x} - 2Q_{2x}'Q_{xx}^{-1}\Omega_{2x}'\Omega_{2x}Q_{xx}^{-1}Q_{2x} && (43)
\end{aligned}$$

After deriving the limiting properties of the residuals, we now consider the multivariate (bivariate) process $w_t = (y_{1t}, \widehat{v}_t)'$ under the I(0) model $(y_{1t}, \widehat{v}_t)' = (\mu_1, 0)' + (u_{1t}, \widehat{v}_t)'$. From Müller and Watson (2017, 2018, 2021) we have

$$T^{1/2}Y_{1jT} \Longrightarrow Y_{1j} = \sigma_1\int_0^1\Psi_j\left(s\right)dW_1\left(s\right) \sim N\left(0,\sigma_1^2\right). \qquad (44)$$

Hence, $T^{1/2}\begin{pmatrix} Y_{1jT} \\ \widehat{V}_{jT}\end{pmatrix} \Longrightarrow \begin{pmatrix} Y_{1j} \\ V_j\end{pmatrix}$, where

$$\begin{pmatrix} Y_{1j} \\ V_j\end{pmatrix} \sim N\left(0, \begin{pmatrix} \sigma_1^2 & \sigma_{1V} \\ \sigma_{1V} & \sigma_V^2\end{pmatrix}\right), \qquad (45)$$

jointly bivariate normal, with

$$\sigma_{1V} = \sigma_{12} - Q_{2x}'Q_{xx}^{-1}\Omega_{1x}'\Omega_{1x}Q_{xx}^{-1}Q_{2x}. \qquad (46)$$

To prove the formula for $\sigma_{1V}$, notice that the covariance between $T^{-1/2}\sum_{t=1}^{T}\Psi_j\left(\left(t-1/2\right)/T\right)u_{1t}$ and

$$\begin{aligned}&T^{-1/2}\sum_{t=1}^{T}\Psi_j\left(\left(t-1/2\right)/T\right)u_{2t}\\&-T^{-1/2}\sum_{t=1}^{T}\Psi_j\left(\left(t-1/2\right)/T\right)\left(\sum_{t=1}^{T}y_{2t}x_t'\right)\left(\sum_{t=1}^{T}x_tx_t'\right)^{-1}u_{xt}\end{aligned}\tag{47}$$

equals

$$\begin{aligned}&T^{-1}Cov\left(\begin{array}{c}\sum_{t=1}^{T}\Psi_j\left(\left(t-1/2\right)/T\right)u_{1t},\\\sum_{t=1}^{T}\Psi_j\left(\left(t-1/2\right)/T\right)\left[u_{2t}-\left(\sum_{t=1}^{T}y_{2t}x_t'\right)\left(\sum_{t=1}^{T}x_tx_t'\right)^{-1}u_{xt}\right]\end{array}\right)\\=\;&T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T}\Psi_j\left(\left(t-1/2\right)/T\right)\Psi_j\left(\left(s-1/2\right)/T\right)Cov\left(u_{1t},\left[u_{2s}-\left(\sum_{t=1}^{T}y_{2t}x_t'\right)\left(\sum_{t=1}^{T}x_tx_t'\right)^{-1}u_{xs}\right]\right)\end{aligned}\tag{48}$$

where the covariance term is

$$\begin{aligned}&Cov\left(u_{1t},u_{2s}\right)-Cov\left(u_{1t},\left(\sum_{t=1}^{T}y_{2t}x_t'\right)\left(\sum_{t=1}^{T}x_tx_t'\right)^{-1}u_{xs}\right)\\=\;&\sigma_{12}-\left(\sum_{t=1}^{T}y_{2t}x_t'\right)\left(\sum_{t=1}^{T}x_tx_t'\right)^{-1}\widehat{\Omega}_{1x}'\widehat{\Omega}_{1x}\left(\sum_{t=1}^{T}x_tx_t'\right)^{-1}\left(\sum_{t=1}^{T}y_{2t}x_t'\right)'.\end{aligned}\tag{49}$$

Finally, as in Müller and Watson (2017, 2018, 2021), we derive the LR parameters of interest in terms of $\left(Y_{1T}',\widehat{V}_T'\right)'$, the $2q\times 1$ cosine transforms of $w_t=(y_{1t},\widehat{v}_t)'$. As was shown these authors them, because of the orthogonality of the cosine regressors $\Psi_T$, there is a tight connection between the variability and covariability in the long-run projections $\widehat{w}_t=(\widehat{y}_{1t},\widetilde{v}_t)'$, where $\widehat{y}_{1t}=Y_{1T}'\Psi\left(\left(t-1/2\right)/T\right)$ and $\widetilde{v}_t=\widehat{V}_T'\Psi\left(\left(t-1/2\right)/T\right)$, and the cosine transforms $\left(Y_{1jT},\widehat{V}_{jT}\right)$, i.e.,

$$T^{-1}\sum_{t=1}^{T}\widehat{w}_t\widehat{w}_t'=\left(\begin{array}{cc}Y_{1T}'Y_{1T}&Y_{1T}'\widehat{V}_T\\\widehat{V}_T'Y_{1T}&\widehat{V}_T'\widehat{V}_T\end{array}\right).\tag{50}$$

Thus, the average covariance matrix of the long-run projections $(\widehat{y}_{1t},\widetilde{v}_t)$ also measures the covariability of the cosine transforms $\left(Y_{1T},\widehat{V}_T\right)$:

$$\Omega_T=T^{-1}\sum_{t=1}^{T}E\left[\left(\begin{array}{c}\widehat{y}_{1t}\\\widetilde{v}_t\end{array}\right)(\widehat{y}_{1t},\widetilde{v}_t)\right]=\sum_{j=1}^{q}E\left[\left(\begin{array}{c}Y_{1jT}\\\widehat{V}_{jT}\end{array}\right)\left(Y_{1jT},\widehat{V}_{jT}\right)\right]\tag{51}$$

$$\begin{aligned}&=\left(\begin{array}{cc}tr\left[E\left(Y_{1T}^2\right)\right]&tr\left[E\left(Y_{1jT}\widehat{V}_{jT}\right)\right]\\tr\left[E\left(Y_{1jT}\widehat{V}_{jT}\right)\right]&tr\left[E\left(\widehat{V}_{jT}^2\right)\right]\end{array}\right)\\&\to\sum_{j=1}^{q}\left(\begin{array}{cc}\sigma_1^2&\sigma_{1V}\\\sigma_{1V}&\sigma_V^2\end{array}\right),\end{aligned}\tag{52}$$

as $T\to\infty$, where

$$\sigma_V^2=\sigma_2^2+Q_{2x}'Q_{xx}^{-1}\Omega_{xx}Q_{xx}^{-1}Q_{2x}-2Q_{2x}'Q_{xx}^{-1}\Omega_{2x}'\Omega_{2x}Q_{xx}^{-1}Q_{2x}.\tag{53}$$