Department of Economics Discussion Paper Series

Identifying the Causal Role of CO2 during the Ice Ages

Jennifer L. Castle, David F. Hendry

Number 898
January, 2020

Department of Economics
Manor Road Building, Manor Road
Oxford, OX1 3UQ

ISSN 1471-0498

Identifying the Causal Role of CO_2 during the Ice Ages

Jennifer L. Castle and David F. Hendry*
Institute for New Economic Thinking at the Oxford Martin School
and
Climate Econometrics at Nuffield College, University of Oxford

January 13, 2020

Abstract

We investigate past climate variability over the Ice Ages, where a simultaneous-equations system is developed to characterize land ice volume, temperature and atmospheric CO_2 levels as non-linear functions of measures of the Earth's orbital path round the Sun. Although the orbital variables were first theorised as the fundamental causes of glacial variation by Croll in 1875 following Agassiz's conception of a 'Great Ice Age' in 1840, their minor variations were thought insufficient to drive such major changes, especially the relative rapidity of shifts between glacial and warmer periods. The changes over the ice ages in atmospheric CO_2 closely matched changes in land ice volumes, and since temperature changes are in turn affected by CO_2 and also closely tracked ice volumes, a key identification issue is the causal role of CO_2 in the process. As any links between CO_2 and temperature above the forces from the orbital drivers (which of course are still operating) must have been natural ones hundreds of thousands of years ago, understanding their interactions at that time is important now that additional CO_2 emissions are anthropogenic. We develop a simultaneous equation system over the last 800,000 years that allows a test of the role of CO_2 as endogenously driven by the orbital variations, or an 'exogenous' influence as it now is.

JEL classifications: C01, C51, C87, Q54.
Keywords: Climate Econometrics; Model Selection; Outliers; Identification; Saturation Estimation; *Autometrics*; Ice Ages.

1 Introduction

While many contributions led to the discovery of massive past glaciation on land, that by Louis Agassiz (1840), based on the contemporaneous movements of glaciers in his native Switzerland and using those to explain a number of previously puzzling features of the landscape in Scotland, was a major step forward in understanding the variability of past climate. Agassiz conceived of a 'Great Ice Age', an intense, global winter lasting ages, rather than multiple Ice Ages as now, but Archibald Geikie (1863) discovered plant fragments between different layers of glacial deposits, implying that sustained warm periods separated

*Financial support from the Robertson Foundation (award 9907422), and the Institute for New Economic Thinking (grant 20029822) is gratefully acknowledged. All calculations and graphs use *PcGive* (Doornik and Hendry, 2018) and *OxMetrics* (Doornik, 2018). email: jennifer.castle@magd.ox.ac.uk and david.hendry@nuffield.ox.ac.uk

cold glacial periods in prehistory. The calculations by James Croll (1875) using just the variations in the Earth's orbit then gave a theoretical mechanism for how ice ages could occur and a time line, where the changing albedo of ice coverage helped explain the relative rapidity with which glacial periods switched, although he predicted that the last ice age was older than observed. Croll's research was later amplified by Milutin Milankovitch (1969) (originally 1941) who calculated solar radiation at different latitudes from changes in obliquity and precession of the Earth as well as eccentricity. Milankovitch also corrected Croll's assumption that minimum winter temperatures mattered, to show that cooler summer maxima were more important in leading to glaciation.

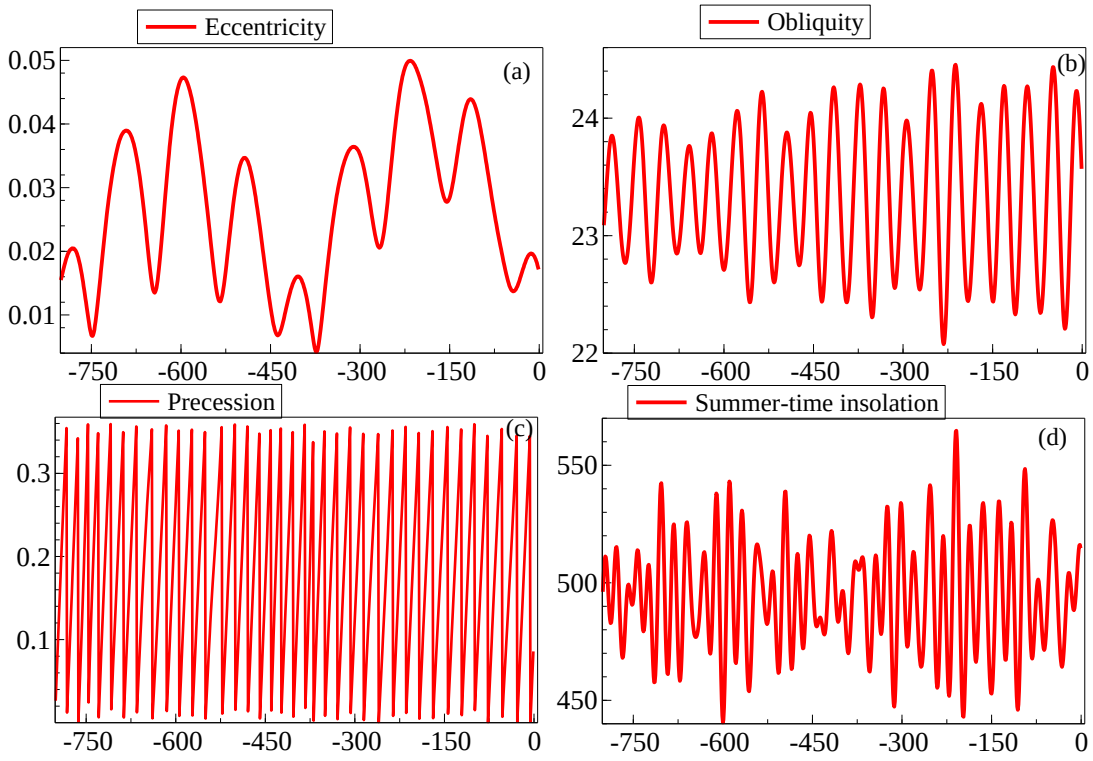

Figure 1: Ice-age orbital drivers: (a) eccentricity (*Ec*); (b) obliquity (*Ob*); (c) precession (*Pr*); (d) Summer-time insolation at 65° south (*St*).

Even a century after Agassiz, there was limited evidence to support such ideas and the timings of glacial episodes. However, these general explanations have since been corroborated by many empirical observations of past oceanic and atmospheric climate changes: see e.g., John Imbrie *et al.* (1992). As we show below, an important reason for analyzing what may seem like the distant past is its relevance today. The climate then was little affected by the activities of the various human species on the planet, partly as they were too sparse and partly did not have the technology. Consequently, any links between, say, CO_2 and temperature above the forces from the orbital drivers (which of course are still operating) must have been natural ones, so can help us understand their present interactions when CO_2 emissions are anthropogenic.

There are three main interacting orbital changes over time affecting incoming solar radiation (insolation) that could drive ice ages and inter-glacial periods. These are: (a) 100,000 year periodicity deriving

from the non-circularity of the Earth's orbit round the Sun induced by the gravitational influences of other planets in the solar system (eccentricity: *Ec* below); (b) a 41,000 year periodicity coming from changes in the tilt of the Earth's rotational axis relative to the ecliptic (obliquity: *Ob*); (c) about 23,000 and 19,000 year periodicities due to the precession of the equinox, which changes the season at which the Earth's orbit is nearest to the Sun, resulting in part from the Earth not being an exact sphere (*Pr*).[1] These three are shown measured at 1000-year intervals in Figure 1(a), (b) and (c), together with summer-time insolation at 65° south (*St*) in Panel (d) (see Didier Paillard, Laurent Labeyrie, and Pascal Yiou, 1996). The X-axes in such graphs are labelled by the time before the present in 1000-year intervals, starting 800,000 years ago. *Ec* and *St* show two major long-swings pre and post about −325 and within each, a number of shorter 'cycles' of varying amplitudes, levels and durations. *Ob* appears to have increased in amplitude since the start of the sample, whereas it is difficult to discern changes in the patterns of *Pr*. The orbital series are clearly strongly exogenous, and most seem non-stationary, but due to shifting distributions, not unit roots.

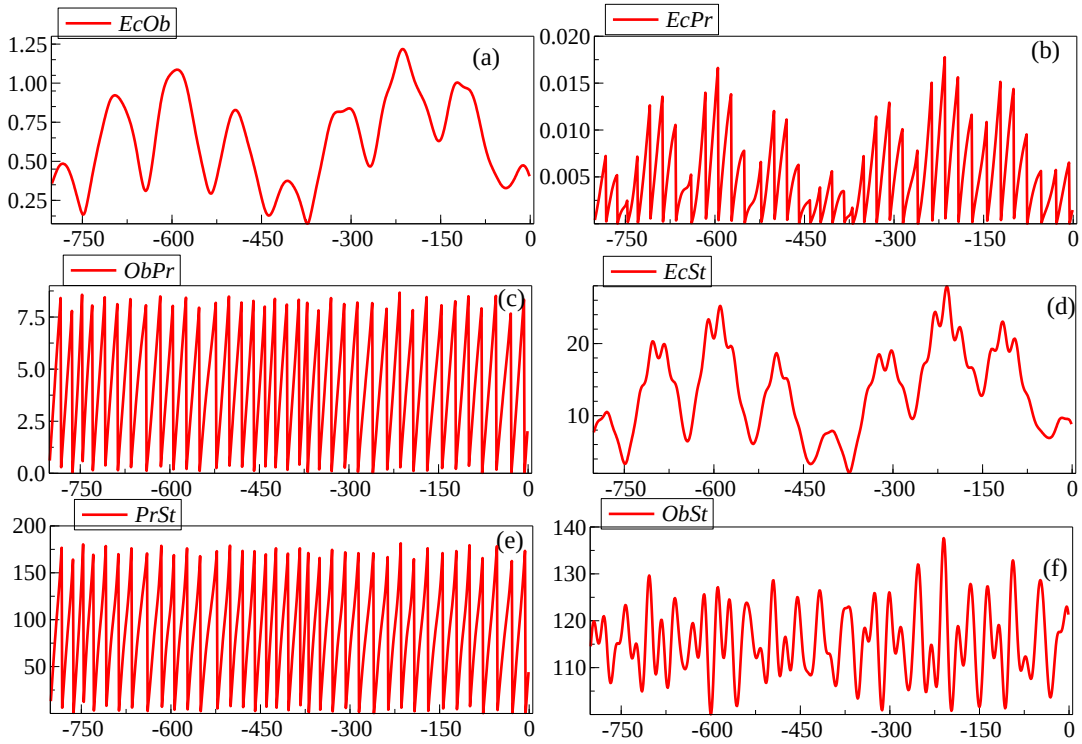

Figure 2: Ice-age orbital driver interactions: (a) *EcOb*; (b) *EcPr*; (c) *ObPr*; (d) *EcSt*; (e) *PrSt*; (f) *ObSt*.

Orbital variations are not the only forces that affect glaciation. The Earth's energy balance is determined by incoming and outgoing radiation: for a cointegrated econometric model thereof, see Pretis (2019). The role of *St* is to summarize changes in the former, but an exogenous summary measure of outgoing radiation is not clear as changes that also affect climate include:
(i) variations in the Sun's radiation output (radiative forcing);

[1]The measure was divided by 1000 to avoid awkwardly small estimated coefficients.

(ii) particulates in the atmosphere from volcanic eruptions, and
(iii) the atmosphere's water vapor and greenhouse gas content (primarily CO_2, N_2O, CH_4);
(iv) reflectivity (albedo) from alterations in ice cover, especially in polar regions, and from
(v) wind blown and volcanic dust covering existing ice;
(vi) ocean temperatures (which lag behind land);
(vii) sea levels and induced ocean circulation patterns;
(viii) cloud cover and its distribution in location and season;
(ix) changes in the magnetic poles.
Of these, (i), (ii) and (ix) seem strongly exogenous, as do the volcanic contributions to (iii) and (v), whereas the rest of (iii) and (iv)–(viii) must be endogenously determined within the global climate system by the strongly exogenous drivers, although anthropogenic greenhouse gas emissions are now 'exogenously' changing the composition of the atmosphere: see Jean-François Richard (1980) for an analysis of modeling changes in a variable's status as endogenous or exogenous, although here that effect would at most affect the last few (1000 year) observations.

That the distance from the Sun matters seems rather natural, as such variations change radiative forcing and hence global temperatures. However, the variations due purely to the eccentricity of the orbit are small. Obliquity also must matter: if the Northern Hemisphere directly faced the Sun, ice would usually be absent there; and if it never faced the Sun, would generally be frozen. Precession seems the smallest driving force of these, but interactions may be important: when the Earth is furthest from the Sun and tilts away in the Northern Hemisphere summer, that may cool faster: see Paillard (2010) for an excellent discussion of these interactions. Even so, a problem with the theory that 'purely orbital' variations drove ice ages over the last 800,000 years is that the known orbital variations should not result in sufficiently large changes in radiative forcing on the Earth to cause the rapid arrival and especially the rapid ending, of glacial periods: see Paillard (2001). Although *St* could provide some additional explanation, and in particular seems to help capture changes at peaks and troughs, we decided to only use the strongly exogenous orbital drivers. An equation regressing *St* just on these and its first lag produced $\mathsf{R}^2 = 0.988$, so we leave to the reader the exercise of building a model with *St* included in the list of variables.[2]

A probable reason for 'rapid' changes in the system is the presence of non-linear feedbacks or interactions between the drivers. Thus, Figure 2 shows their interactions in Panels (a) [$Ec \times Ob$], (b) [$Ec \times Pr$], (c) [$Ob \times Pr$], (d) [$Ec \times St$], (e) [$Pr \times St$] and (f) [$Ob \times St$] although the model developed here includes only the first three interactions together with the squares to capture non-linear influences. Recently, Kristina Pistone, Ian Eisenman, and Veerabhadran Ramanathan (2019) have shown that the complete disappearance of Arctic sea ice would be (in temperature terms) 'equivalent to the effect of one trillion tons of CO_2 emissions' (roughly 140 ppm) because an open ocean surface typically absorbs approximately six times more solar radiation than a high albedo surface covered with sea ice. Such an effect reducing ocean ice as the climate gradually warmed after the peak of glacial extent would accelerate melting, and conversely for cooling. Based on ice-age evidence, Anton Vaks, Andrew Mason and Sebastian Breitenbach *et al.* (2019) show that a sea-ice-free Arctic makes permafrost vulnerable to thawing, which would further accelerate warming, although they find that thawing permafrost last occurred about 400,000 years ago, about mid-way through our sample. However, the non-linear model in Francis Diebold and Glenn

[2]For more comprehensive systems that endogenously model measures for all the variables in (iii)–(vii), see Kaufmann and Katarina Juselius (2013), and Pretis and Kaufmann (2018). We are also grateful to those authors for providing the data series analyzed here.

Rudebusch (2019) suggests there is a 60 percent chance of an ice-free Arctic Ocean in the 2030s, which would exacerbate current warming.

Explaining glaciation over the ice ages has garnered a huge literature, only a small fraction of which is cited here. As the following quote illustrates, the possibility of the Northern Hemisphere facing another ice age was still considered in the 1950s, but by 1982, Hubert Lamb emphasized global warming as the more serious threat to climate stability.

> We do not yet know whether the latest turn in our climatic fortunes, since the optimum years of the 1930s, marks the beginning of a serious downward trend or whether it is merely another wobble...
> Lamb (1959)

2 Data series over the past 800,000 years

A vast international effort over many decades has been devoted to measuring the behavior of a number of variables over the ice ages. Naturally, proxies or indirect but closely associated observables that remain in the ground, ice, oceans and ocean floors are used based on well-established physical and chemical knowledge. Econometricians are essentially mere end users of this impressive research base.

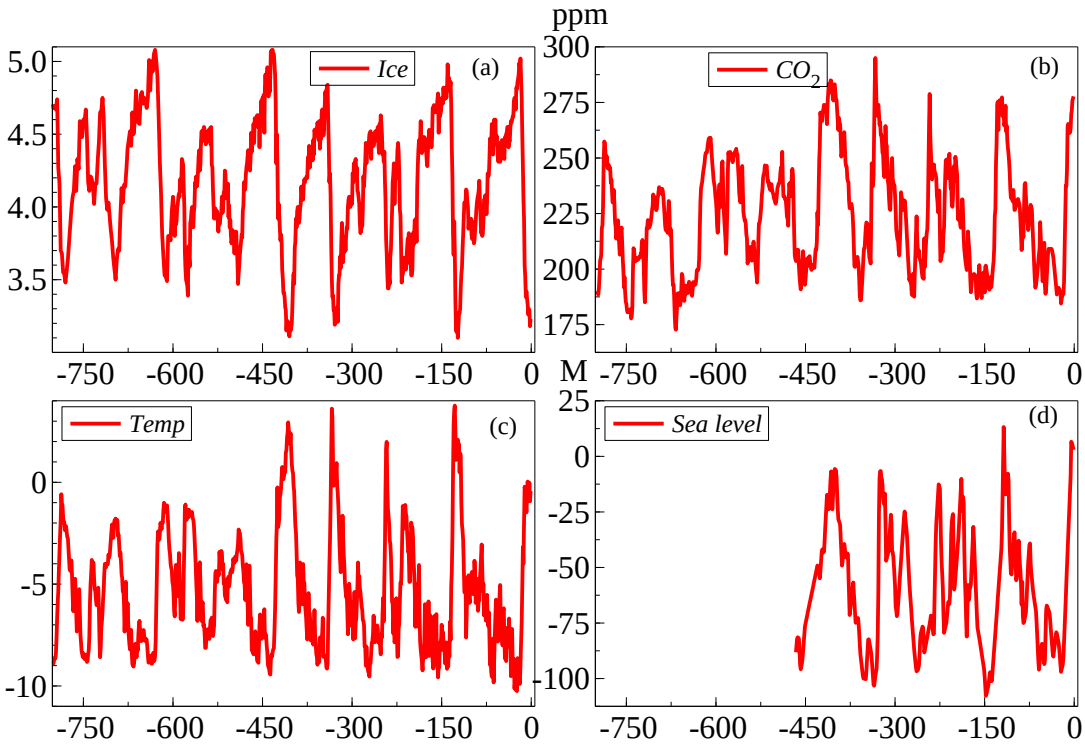

Figure 3: Ice-age time series: (a) Ice volume (*Ice*); (b) atmospheric CO_2 in parts per million (CO_2); (c) temperature (*Temp*); (d) shorter-sample sea level changes in meters.

Antarctic-based land surface temperature proxies (denoted *Temp* below) were taken from Jean Jouzel, Valérie Masson-Delmotte, O. Cattani, and Gabrielle Dreyfus *et al.* (2007). The paleo record from deep ice cores show that atmospheric CO_2 varied between 170ppm and 300ppm over the ice ages, where 1ppm $= 7.8$ gigatonnes of CO_2 (see Lüthil *et al.*, 2008). Ice volume estimates (denoted *Ice* below) were from Lisiecki and Raymo (2005) (based on δ^{18}O as a proxy measure). To capture orbital variations, *Ec*, *Ob* and *Pr* and their interactions are conditioned on. All observations had been adjusted to the common EDC3 time scale and linearly interpolated for missing observations to bring all observations on a 1000 year time interval (EDC3 denotes the European Project for Ice Coring in Antarctica–EPICA–Dome C, where drilling in East Antarctica has been completed to a depth of 3260 meters, just a few meters above bedrock (see Frédéric Parrenin *et al.*, 2007). Synchronization between the EPICA Dome C and Vostok ice core measures over the period $-145,000$ to the present was based on matching residues from volcanic eruptions (see Parrenin *et al.*, 2012). The total sample size in 1000 year intervals is $T = 801$ with the last 100 observations (i.e., 100,00 years, ending 1000 years before the present) used to evaluate the predictive ability of the estimated system. Figure 3 records the time series together with a shorter sample of sea level data.[3]

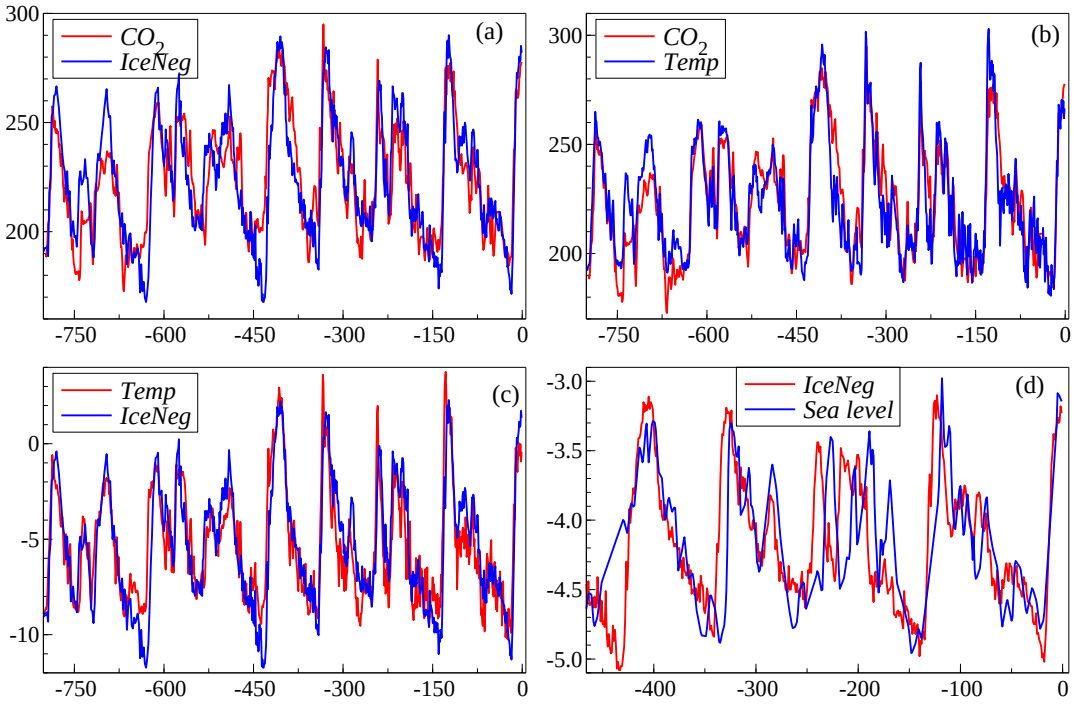

Figure 4: (a) CO_2 and the negative of ice volume (*IceNeg*); (b) CO_2 and temperature; (c) temperature and *IceNeg*; (d) *IceNeg* and sea level (only for the last 465,000 years).

[3]Sea surface temperature data are available from Alfredo Martínez-Garcia, Antoni Rosell-Melé, Walter Geibert, Rainer Gersonde, Pere Masqué, Vania Gaspari, and Carlo Barbante (2009) which could help explain oceanic CO_2 uptake and interactions with land surface temperature. Sea level data, based on sediments, can be obtained from M. Siddall *et al.* (2003), over a shorter sample, but are not analyzed here.

We focus on modeling *Ice*, CO_2 and *Temp* as jointly endogenous functions of the orbital variables which we take to be strongly exogenous, so feedbacks onto their values from Earth's climate are negligible. The patterns of these three time series are remarkably similar, all rising (or falling) at roughly the same times. Figure 4 emphazises how close these movements are by plotting pairs of time series: (a) CO_2 and the negative of ice volume (denoted *IceNeg*); (b) CO_2 and *Temp*; (c) *Temp* and *IceNeg*; (d) *IceNeg* and sea level (only for the last 465,000 years). The pattern in the atmospheric CO_2 levels is closely similar to that of the negative ice volume, the temperature record and sea level, as are other pairs.

If ice ages are due to orbital variations, why should atmospheric CO_2 levels also correlate so closely with ice volume? David Lea (2004) relates changes in tropical sea surface temperature to atmospheric CO_2 levels over the last 360,000 years to suggest that CO_2 was the main determinant of tropical climate. Conversely, in https://climateaudit.org/2005/12/18/gcms-and-ice-ages/, Stephen McIntyre argues that CO_2 should not be treated as a forcing variable in statistical models of ice-age climate, as it is an endogenous response. So is the mechanism not orbital variations, but instead that changes in atmospheric CO_2 levels alter global temperatures which in turn drive changes in ice volume? The answer lies in the deep oceans, in particular, the Southern Ocean, which acts as a carbon sink during cold periods, and releases some of that CO_2 as the planet warms, in turn enhancing cooling and warming: see e.g., Samuel Jaccard, Eric Galbraith, Alfredo Martínez-García, and Robert Anderson (2016). Thus, the exogenous orbital variations drive temperature, which drives changes in ice volume and in turn CO_2 levels. By modeling the 3-variable simultaneous-equations system estimated using full information maximum likelihood (FIML: see e.g., Hendry, 1976), treating all three as endogenous, the roles of *Temp* and CO_2 as endogenous determinants of *Ice* can be investigated. The next section describes the formulation of simultaneous equations models, Section 3.1 discusses identification, then Section 3.2 considers the problem of 'weak instruments'.

3 Selecting simultaneous equations models

A simultaneous equations representation is a model of a system of n variables, $\mathbf{y}_t$, that are to be modeled as jointly endogenous by m other variables, $\mathbf{z}_t$, that are non-modeled. The properties of such systems were first analyzed in Haavelmo (1943). Their representations derive from the theory of reduction as representing the local data generation process (LDGP: see Hendry, 2009, 2018). To validly condition on $\mathbf{z}_t$ requires that those variables are known to be at least weakly exogenous (see Koopmans, 1950b, and Engle, Hendry, and Richard, 1983). In essence, the weak exogeneity of $\mathbf{z}_t$ requires that the DGP of $\mathbf{z}_t$ does not depend on the parameters of the DGP of $\mathbf{y}_t$ conditional on $\mathbf{z}_t$. When the status of $\mathbf{z}_t$ is not certain, as with CO_2 in Section 4.2, $\mathbf{z}_t$ should be treated, at least initially, as a component of $\mathbf{y}_t$. The strong exogeneity of $\mathbf{z}_t$, as later applies to the orbital drivers, requires that their DGP does not depend on lagged values of $\mathbf{y}_t$, in which case non-linear functions of the $\mathbf{z}_t$ can also be included as 'conventional' conditioning determinants of $\mathbf{y}_t$ even when they are non-stationary.

A dynamic representation of the system $(\mathbf{y}_t, \mathbf{z}_t)$ can be formulated as a vector autoregression (VAR) conditional on the $\mathbf{z}_t$ (often denoted by VARX), lags of all the variables, and deterministic terms such as intercepts and any indicator variables, denoted $\mathbf{d}_t$:

$$\mathbf{y}_t = \mathbf{\Psi}_0 \mathbf{z}_t + \sum_{j=1}^{s} \mathbf{\Psi}_j \mathbf{z}_{t-j} + \sum_{i=1}^{s} \mathbf{\Gamma}_i \mathbf{y}_{t-i} + \mathbf{\Upsilon} \mathbf{d}_t + \mathbf{u}_t \tag{1}$$

where $\mathbf{u}_t \sim \mathsf{IN}_n[\mathbf{0}, \mathbf{\Omega}_u]$. The assumptions on the error process require that s is sufficiently large to create a martingale difference process, and the variables including $\mathbf{d}_t$ remove any outliers, location shifts and parameter changes not captured by the other regressors so that homoskedasticity and constant parameters are viable. Then given that $\mathbf{z}_t$ is weakly exogenous, the error process will also be uncorrelated with the regressors.

To check the specification of (1), the system should be tested for congruence: once the initial system is congruent, all later reductions of it should be congruent as well to avoid relevant information being lost. Next, a parsimonious version of the system in (1) can then be selected, while ensuring that congruence is maintained (denoted PVARX): Hendry and Mizon (1993) propose evaluating such dynamic models by their ability to encompass the VAR. Since the initial system is identified, all later non-simultaneous reductions from eliminating insignificant variables must be as well. At this selection stage, the system should also have been reduced to a non-integrated (I(0)) representation so that conventional critical values can be used: if the data are I(1), cointegration and differencing can do so.

A system like (1), which is also called the 'reduced form' in the econometrics literature, is always identified, so that multivariate least-squares estimators of its parameters are unique, and under these assumptions will deliver consistent estimates. A simultaneous equations representation is a model of the system derived by reduction from (1): see Hendry, Adrian Neale and Frank Srba (1988). However, in econometrics there is often a pre-specified theory of that representation from which the 'reduced form' is derived (hence the terminology), inverting the correct order of the relationship between the system and a model thereof.

Written in a concise notation, with the $N \times 1$ vector $\mathbf{w}_t$ denoting all the right-hand side variables, the system in (1) is:

$$\mathbf{y}_t = \mathbf{\Pi}\mathbf{w}_t + \mathbf{u}_t \text{ where } \mathbf{u}_t \sim \mathsf{IN}_n[\mathbf{0}, \mathbf{\Omega}_n] \tag{2}$$

Then a simultaneous-equations model of (2) is a reduction to:

$$\mathbf{B}\mathbf{y}_t = \mathbf{C}\mathbf{w}_t + \boldsymbol{\epsilon}_t \text{ where } \boldsymbol{\epsilon}_t \sim \mathsf{IN}_n[\mathbf{0}, \mathbf{\Sigma}_{\boldsymbol{\epsilon}}] \tag{3}$$

with:

$$\mathbf{B}\mathbf{\Pi} = \mathbf{C} \tag{4}$$

A necessary condition for (4) to be solvable is that there are no more non-zero parameters in $\mathbf{B}$ and $\mathbf{C}$ than the $n \times N$ in $\mathbf{\Pi}$, which is called the order condition. In addition, when the rank condition discussed in §3.1 is satisfied, $\mathbf{B}$ and $\mathbf{C}$ have a unique relation to $\mathbf{\Pi}$ and (3) is fully identified. We use 'structure' (in quotation marks) to denote an equation with more than one endogenous variable as in (3), without any claim that it is a structural equation in the sense of being invariant to extensions of the information set for new variables, over time, and across regimes. A simultaneous-equations model like (3) then needs to be estimated appropriately, because including the i^{th} endogenous variable in the equation for the j^{th} will induce a correlation with its equation error. An infinite number of possible estimation methods exists, characterized by the estimator generating equation in Hendry (1976). Here we use full information maximum likelihood (FIML) first proposed in Koopmans (1950a). The general formulation and estimation procedures underlying FIML are described in Hendry, Neale, and Srba (1988). Since a simultaneous-equations model is a reduction from the system, automatic model selection is applicable as discussed in Hendry and Krolzig (2005): Doornik and Hendry (2017) propose an algorithm for doing so based on the multi-path search procedure of *Autometrics*, a variant of which is applied in Section 4.2.

3.1 Identification in the DGP

Identification, in the sense of uniqueness of $\mathbf{B}, \mathbf{C}$, in systems like (3) given $\mathbf{\Pi}$ has been extensively explored in the econometrics literature: see e.g. Koopmans (1949), Koopmans and Olaf Reiersøl (1950), Frank Fisher (1966) and Thomas Rothenberg (1973) *inter alia*. The rank condition for identification determines the extent to which each equation is or is not identified. In that literature, identification is usually a synonym for uniqueness, although usage also entails connotations of 'interpretable in the light of subject matter theory' and 'corresponding to reality' (as in 'identify a demand curve', as opposed to a supply relation, or a mix). Whether or not $\mathbf{B}, \mathbf{C}$ in (3) can be recovered uniquely from $\mathbf{\Pi}$ in (2) requires the exclusion of some different variables in every equation and the inclusion of some others, otherwise linear combinations of equations cannot be distinguished.[4] Given the appropriate exclusions and inclusions corresponding to particular elements of $\mathbf{B}$ and $\mathbf{C}$ being zero, the rank condition is then satisfied so (3) is fully identified. Consequently, $\mathbf{B}$ and $\mathbf{C}$ are unique related to $\mathbf{\Pi}$, which entails restrictions on the $\mathbf{\Pi}$ matrix in (2). The system for the three ice-age variables in Section 4.2 is highly overidentified, so all the $y_{i,t}, i \neq j$ can be included in the other equations for $y_{j,t}, j \neq i$.

To avoid later 'spurious identification' of a simultaneous representation, all the right-hand side variables need to be significant at a reasonable level both in the system and in their associated equations. Otherwise, claiming identification by excluding insignificant regressors from any equation based on their apparent presence in other equations, when in fact they are also insignificant there, will be misleading when such variables are actually irrelevant to the system as a whole. Throughout selection of a simultaneous representation, the rank condition for identification can be imposed as a constraint, both to ensure that essentially the 'same equation', but with different normalizations, is not included twice, and that at every stage, the current form is identified (see Hendry, Neale, and Srba, 1988).

There are three possibilities of lack of identification, exact identification, and over identification (subsets of parameters could be identified or not when others are the converse, in which case the following comments apply to the appropriate set). When $\mathbf{B}$, $\mathbf{C}$ are not identified, then (2) is the least restricted but still fully identified, representation. Any just-identified simultaneous representation with a form like (3) will also be minimally identified, so there is an equivalence class of such specifications with equal likelihood (see e.g., Rothenberg, 1971), although in such a setting, reductions may be possible by eliminating irrelevant regressor variables from the entire system.

When $\mathbf{B}, \mathbf{C}$ are over identified by the rank condition, then (3) is a unique representation for the given restrictions. However, Hendry, Maozu Lu, and Mizon (2009) show there may exist different sets of restrictions embodied in matrices $\mathbf{B}^*$, $\mathbf{C}^*$ which are not linear transforms of $\mathbf{B}$, $\mathbf{C}$ (precluded by their identifiability), but under which (3) is equally over identified. Thus, again an equivalence class of such specifications with equal likelihood can result: a given degree of over identification by itself does not ensure a unique model even when there is a unique DGP. The validity of any set of over-identified restrictions can be checked through parsimonious encompassing of the system by the 'structure'. When $\mathcal{L}$ is the log-likelihood of the system (2), and $\mathcal{L}_0$ that of the 'structural' form (3), in stationary DGPs, the test is $2(\mathcal{L} - \mathcal{L}_0) \sim \chi^2_{\mathsf{OR}}(p)$ for p over-identifying restrictions (see Koopmans, 1950a, and Hendry and Mizon, 1993).[5]

[4] Other forms of restriction than exclusions could identify 'structural' parameters that do not directly satisfy the rank condition, such as a diagonal error covariance matrix, or cross-equation links, but these are not considered here.

[5] Hence the earlier advice to obtain an $\mathsf{I}(0)$ representation of the system.

3.2 Weak instruments

Weak instruments show up as a poorly determined initial system, or requiring a loose significance level for regressors in (1) to be retained. That states, but does not resolve, the problem which lies in available information, not the performance of any selection algorithm or modeling approach. The choice of instruments can be made as proposed by Hendry and Krolzig (2001) (also see Hall, Rudebusch, and Wilcox, 1996) both to determine their relevance for each endogenous variable, and to test for instrument mis-specification as part of the congruence check.

A 'structure' could be identified in principle, yet the available instruments may be so weak that for practical purposes, the uncertainty is close to unbounded. A large literature exists on this problem: see e.g., Douglas Staiger and James Stock (1997), Eric Zivot, Richard Startz, and Charles Nelson (1998), Jiahui Wang and Zivot (1998), Peter Phillips (1989), Stock and Jonathan Wright (2000), and Sophocles Mavroeidis (2003). Again, the issue is intrinsic to the available information, and is not a problem created by model selection procedures.

4 System equation modeling of the Ice Age variables

In addition to the many dozens of climatology-based studies, there are several econometric analyses of ice-age data, examining issues of cointegration and the adequacy of using orbital variables as the exogenous explanatory regressors. Kaufmann and Juselius (2010, 2013) analyze the late Quaternary 'Vostok' period of four 'glacial cycles' and Pretis and Kaufmann (2018) build and simulate a statistical climate model over the paleo-climate record of the 800,000 years of data investigated here.

Our focus is on modeling *Ice* allowing for the endogeneity of *Temp* and CO_2, with dynamic feedbacks, non-linear impacts of the orbital variables and handling outliers. Consequently, the initial GUM is a VARX(1) for *Ice*, CO_2 and *Temp* conditional on the nine orbital measures and non-linear functions thereof discussed above, with a one-period lag (i.e., 1000 years earlier) on all variables. System impulse-indicator saturation (IIS: see Hendry Søren Johansen and Carlos Santos, 2008, and Johansen and Bent Nielsen, 2008) was implemented selecting indicators at 0.1%, with all the continuous variables retained. After locating outliers, the regressor variables were then selected at 1% to create a parsimonious VARX(1). Next, that system was transformed to a simultaneous-equations model of the PVARX(1), where only variables and outliers that were relevant in each equation were included, and finally contemporaneous links were investigated. Only retaining variables that are significant in a PVARX(1) avoids 'spurious identification' from using completely irrelevant variables that are then excluded differently in each equation to apparently achieve the order condition (see e.g., Hendry, Neale, and Srba, 1988, and the many references therein).

4.1 The general unrestricted model (GUM)

The GUM in this setting is a dynamic system with strongly exogenous regressors which can be written as:

$$\mathbf{y}_t = \boldsymbol{\gamma}_0 + \boldsymbol{\Gamma}_1 \mathbf{y}_{t-1} + \boldsymbol{\Psi}_2 \mathbf{z}_t + \boldsymbol{\Psi}_3 \mathbf{z}_{t-1} + \boldsymbol{\Upsilon} \mathbf{d}_t + \boldsymbol{\epsilon}_t \tag{5}$$

where:

$$\mathbf{y}_t = (Ice \;\; CO_2 \;\; Temp)_t \;\text{ and }\; \mathbf{z}_t' = \left(Ec \;\; Ob \;\; Pr \;\; EcOb \;\; EcPr \;\; PrOb \;\; Ec^2 \;\; Ob^2 \;\; Pr^2\right)_t \tag{6}$$

and $\mathbf{d}_t$ includes the vector of impulse indicators selected by system IIS. The difference from single-equation IIS described in Hendry, Johansen, and Santos (2008) is that indicators have to be significant at the target nominal significance level in the system, not just in any one equation therein. The lagged values are to capture dynamic inertia: when the ice covers a vast area, that will influence the ice sheet in the next period, even when periods are 1000 years apart. Moreover, that observation length is just 1% of the eccentricity periodicity, so the Earth will still be close to its previous position.[6]

First, all the $\mathbf{y}_{t-1}$, $\mathbf{z}_t$ and $\mathbf{z}_{t-1}$ in (5) are retained without selection when IIS is applied at $\alpha = 0.001$ for $T = 697$ keeping the last hundred observations for out-of-sample forecast evaluation. This led to 35 impulse indicators being selected, the earliest of which was $1_{\{-339\}}$. However, many of these were retained to avoid a failure of encompassing the first feasible GUM, and were not significant at $\alpha = 0.001$.

Table 1 records the correlations between the actual observations and the fitted values taking impulse indicators into account, so each variable can be explained in large measure by a model of the form (5). Table 2 shows the correlations between the residuals of the three equations, with residual standard deviations on the diagonal. There remains a high correlation between *CO*$_2$ and *Temp* residuals even conditional on all the orbital variables, but not between those of *Ice* and either *CO*$_2$ or *Temp*, although those correlations remain negative.

Ice	*CO*$_2$	*Temp*
0.981	0.981	0.972

Table 1: Correlations between actual and fitted values in the VARX(1).

	Ice	*CO*$_2$	*Temp*
Ice	0.090	—	—
CO$_2$	−0.179	5.13	—
Temp	−0.180	0.574	0.711

Table 2: Correlations between VARX(1) residuals, with standard deviations on the diagonal.

Next, the other regressors were selected at 1% resulting in a PVARX(1). Again, note that selection decisions are at the level of the system rather than individual equations. Finally, to avoid the spurious identification issue from indicators that were insignificant in the system, any that were also insignificant in every equation were manually deleted from the system, still leaving 32.

4.2 The simultaneous system estimates

Because many of the exogenous and lagged variables and impulse indicators were only significant in one equation, we reformulated the system as a simultaneous-equation model. This treats all three modeled variables as endogenous and was estimated by FIML. We then manually eliminated insignificant regressors in each equation in turn. The current dated values of *Temp* and *CO*$_2$ in the *Ice* equation and of *Temp* in the *CO*$_2$ equation were insignificant, but that of *CO*$_2$ was significant in the equation for *Temp*. This

[6]Residual autocorrelation suggests that a second lag may also matter, despite such variables being 2,000 years earlier.

delivered the system model in (7)–(9): the Appendix reports the retained impulse indicators.

$$\begin{aligned}\widehat{Ice}_t = \underset{(0.34)}{1.43} &+ \underset{(0.015)}{0.860}\ \mathit{Ice}_{t-1} - \underset{(0.002)}{0.020}\ \mathit{Temp}_{t-1} + \underset{(31)}{102}\ \mathit{Ec}_t - \underset{(32)}{101}\ \mathit{Ec}_{t-1} \\ &- \underset{(0.014)}{0.040}\ \mathit{Ob}_{t-1} - \underset{(1.30)}{5.07}\ \mathit{EcOb}_t + \underset{(1.36)}{5.05}\ \mathit{EcOb}_{t-1} - \underset{(1.03)}{4.97}\ \mathit{EcPr}_t\end{aligned} \tag{7}$$

$$\begin{aligned}\widehat{CO}_{2,t} = \underset{(32)}{218} &+ \underset{(0.018)}{0.853}\ \mathit{CO}_{2,t-1} + \underset{(0.18)}{1.34}\ \mathit{Temp}_{t-1} + \underset{(342)}{1400}\ \mathit{Ec}_t - \underset{(647)}{3070}\ \mathit{Ec}_{t-1} \\ &- \underset{(2.31)}{13.0}\ \mathit{Ob}_{t-1} + \underset{(23)}{70.7}\ \mathit{EcOb}_{t-1} + \underset{(0.047)}{0.232}\ \mathit{Ob}_t^2\end{aligned} \tag{8}$$

$$\begin{aligned}\widehat{Temp}_t = - \underset{(0.69)}{2.49} &+ \underset{(0.023)}{0.879}\ \mathit{Temp}_{t-1} + \underset{(0.0026)}{0.0080}\ \mathit{CO}_{2,t} - \underset{(37)}{301}\ \mathit{Ec}_t + \underset{(2.45)}{22.6}\ \mathit{EcOb}_t \\ &- \underset{(1.94)}{9.80}\ \mathit{EcOb}_{t-1} + \underset{(7.1)}{25.5}\ \mathit{EcPr}_t\end{aligned} \tag{9}$$

The correlations between the actual and fitted values for the three variables in the SEM are virtually identical to those in Table 1, consistent with the likelihood-ratio test of the over-identifying restrictions against the PVARX(1) being $\chi^2_{\mathsf{OR}}(64) = 69.7$, which is insignificant at even the 5% level. Although the inertial dynamics play a key role in the three equations, all the eigenvalues of the system dynamics are less than unity in absolute value at (0.97, 0.86, 0.77). The test for excluding all the non-linear functions yields $\chi^2(8) = 155^{**}$, and that for dropping all the impulse indicators $\chi^2(44) = 439^{**}$, both of which reject at any viable significance level.

Figure 5 records the actual and fitted values and residuals scaled by their standard deviations for the three equations. The tracking is very close, including over the final 100 'out-of-sample' observations, although the residuals show the occasional outlier: remember that IIS selection was at 0.1% to avoid overfitting. Figure 6 reports residual densities with a Normal matched by mean and variance, and correlograms. The densities are relatively close to the Normal for *Ice* and *Temp* after IIS, but less so for CO_2, probably because the restriction to one lag has left some residual autocorrelation. Most of the formal mis-specification tests rejected, possibly also reflecting the many omitted influences noted above, although most of those seem to be endogenous responses as the climate changed, such as dust from wind storms and sea level changes both varying with temperature. In the present context, outliers as represented by impulse indicators could derive from measurement errors in the variables, super-volcanoes either dramatically lowering temperature by erupted particulates, or raising by emitting large volumes of CO_2, or like wind-blown dust changing the albedo of ice sheets. Most of those indicators retained for *Ice* were negative around -0.2 (see appendix §6), whereas for CO_2 they were primarily positive and around $+15$, and for *Temp* around 2 but mixed in sign. However, outliers are relative to the model being estimated, so those found here by IIS could also represent features of variables omitted from the system in (7)–(9).

Table 3 records the correlations between the residuals of the simultaneous equations model, with the residual standard deviations on the diagonal: these are close to those in Table 2.

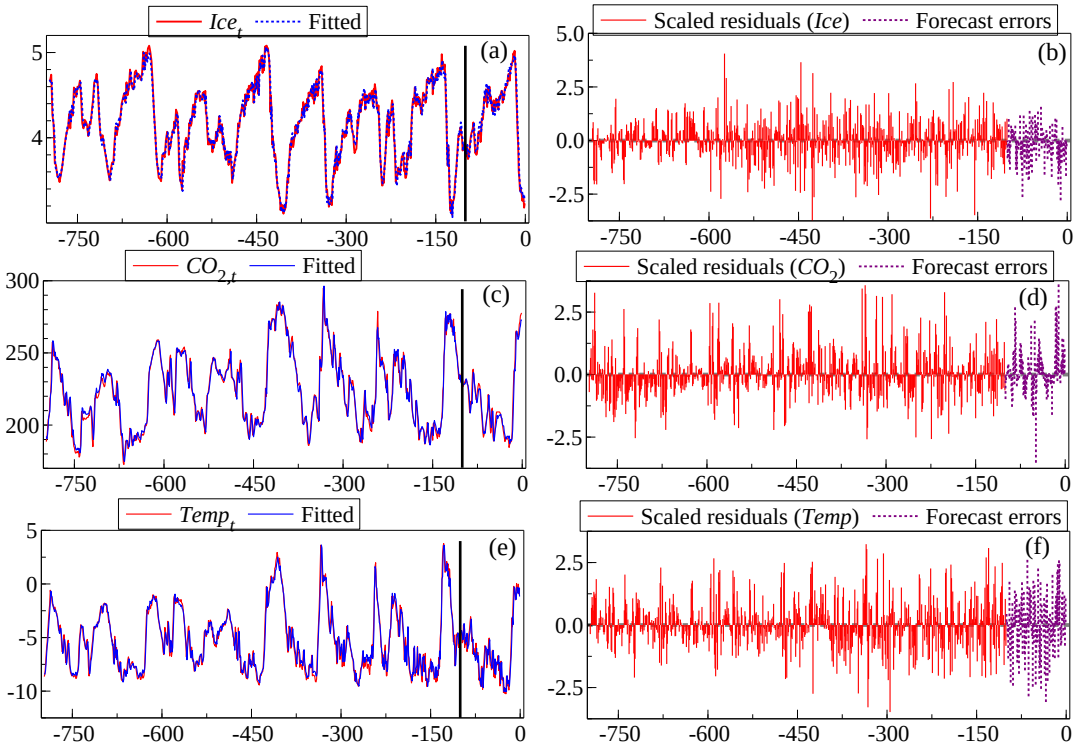

Figure 5: Actual, fitted and forecast values, with scaled residuals and forecast errors: (a) & (b) for *Ice* from (7); (c) & (d) for CO_2 from (8); (e) & (f) for *Temp* from (9). The vertical bar at $T = -100$ marks the start of the forecast period.

	Ice	CO_2	*Temp*
Ice	0.086	—	—
CO_2	−0.173	4.88	—
Temp	−0.184	0.509	0.667

Table 3: Correlations of the simultaneous model residuals, with standard deviations on the diagonal.

Considering the equations in more detail, the volume of ice in (7) depends on its previous level and on previous temperatures, as well as on eccentricity, on past obliquity and also on the current and lagged interactions of eccentricity with obliquity, and with current precession. Although *Ec* and *EcOb* appear to enter primarily as changes, more than their levels, the solved long-run outcome in Table 4 confirms they both enter significantly as levels. CO_2 has a similar coefficient on its lag, a feedback from the previous temperature, current and past levels of eccentricity, on past obliquity and their interaction, and on squared obliquity. Third, *Temp* responds to its previous value and positively to current CO_2: the coefficient in (9) entails that the 100 ppm increase seen since 1958 would raise temperatures by 0.80°C *ceteris paribus*. However, neither current CO_2 nor current *Temp* are significant if added to the equation for *Ice*; and current *Temp* is insignificant if added to the equation for CO_2.

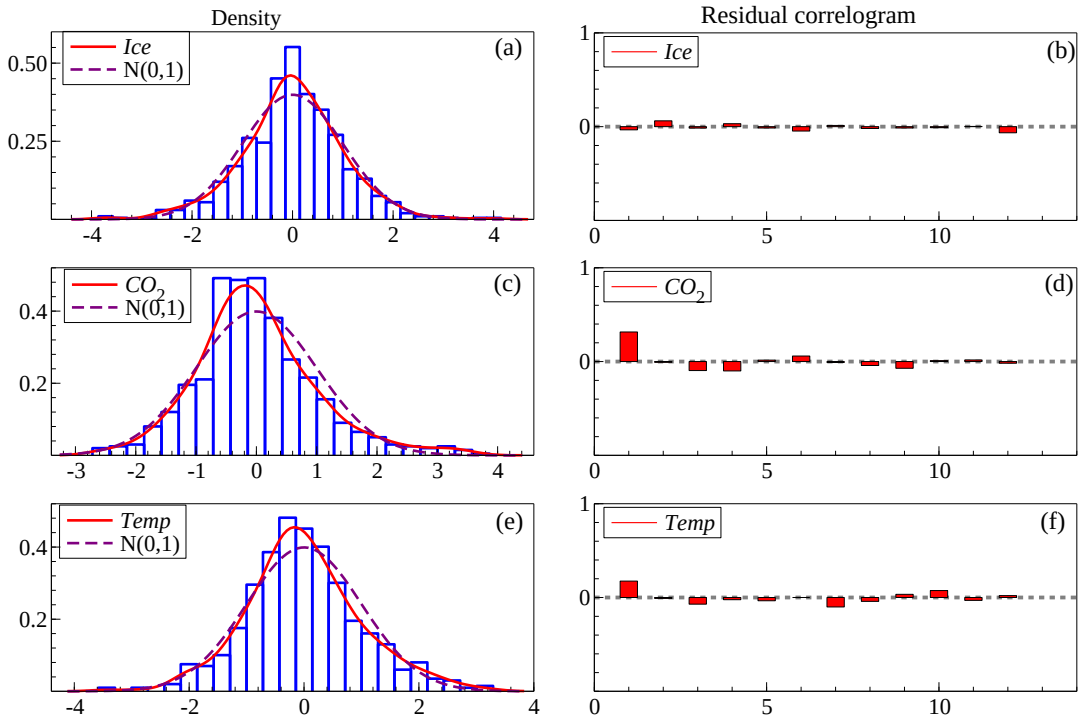

Figure 6: Residual densities and correlograms: (a) & (b) for *Ice* from (7); (c) & (d) for CO_2 from (8); (e) & (f) for *Temp* from (9).

5 Implications

We first modeled the dynamic three equation system of *Ice*, CO_2 and *Temp*, as a function of the strongly exogenous non-linear combinations of the orbital drivers, eccentricity *Ec*, obliquity *Ob*, and precession *Pr* over 700,000 years up to 100,000 years before the present. We robustified those estimates by removing large outliers by system-based impulse-indicator saturation (IIS), then simplified the VARX(1) to a PVARX(1) by first eliminating variables that were insignificant in the system, then in each of the equations in turn. The resulting system has many implications as discussed in this section. First, §5.1 derives the long-run implications and considers the computed time series based only on the strongly exogenous variables. Then §5.2 evaluates the 1-step and long-run (dynamic) forecasts over the last 100,000 years. From those findings, §5.3 asks when did humanity first influence the climate system? Finally, §5.4 examines the evidence on the role of CO_2 during past ice ages.

5.1 Long-run implications

The long-run solutions in Table 4 solve out the dynamics and lags to express each endogenous variable just as a function of the relevant strongly exogenous orbital variables. The original coefficients are not easy to interpret as they depend on the units of measurement of the orbital variables, so CO_2 has been

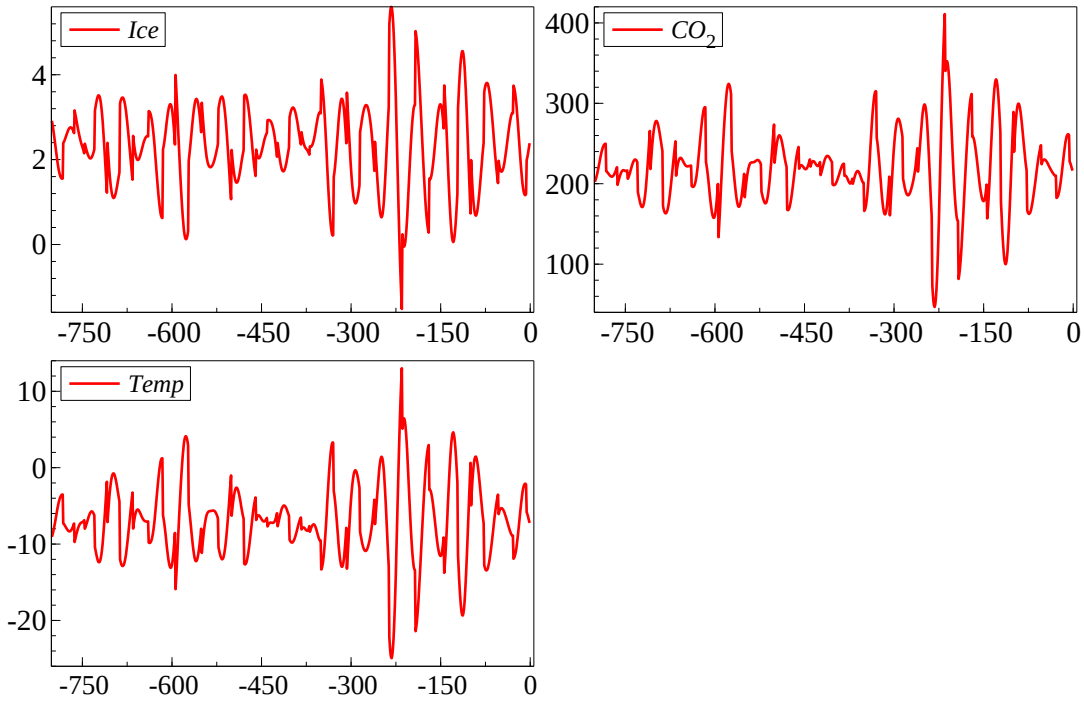

Figure 7: Computed time series of *Ice*, CO_2 and *Temp* from the long-run relationships in Table 4.

divided by 100 and *Temp* by 10 to align numerical coefficient values. Figure 7 graphs the computed time series of *Ice*, CO_2 and *Temp* from the long-run relationships in Table 4. These graphs include the last 100,000 years which are outside the estimation sample.

	1	*Ec*	*Ob*	*EcOb*	*EcPr*	*ObSq*
Ice	−17.3	1162	1.80	−49.3	−111	−0.037
SE	(16.3)	(402)	(1.2)	(17)	(31)	(0.020)
CO_2	32.0	−837	−2.19	35.6	47.2	0.039
SE	(11.9)	(276)	(0.87)	(11.7)	(19.1)	(0.016)
Temp	19.0	−798	−1.44	34.0	52.0	0.026
SE	(10.8)	(257)	(0.80)	(10.9)	(19.5)	(0.014)

Table 4: Long-run solutions as a function of the relevant strongly exogenous orbital variables where CO_2 has been divided by 100 and *Temp* by 10 to align numerical coefficient values.

Despite the somewhat different coefficients in the three long-run equations, the outcome time series are relatively similar, and the correlations between them all exceed $|0.977|$. These graphs are of course just recombinations of the orbital drivers weighted by the coefficients in Table 4, so reflect the relatively volatile and quiescent periods seen in Figure 1. Nevertheless, the increase in volatility starting around

250, 000 years ago is marked, so the series are not stationary. The inertial dynamics as modeled by the lagged dependent variables and the other lagged regressors smooths that over time.

5.2 1-step and long-run forecasts

Figure 8 records the hundred 1-step ahead forecasts with forecast intervals for ±2SE shown by error bands based on coefficient estimation variances as well as the residual variances. The resulting forecast errors (unscaled) are shown in the second column.

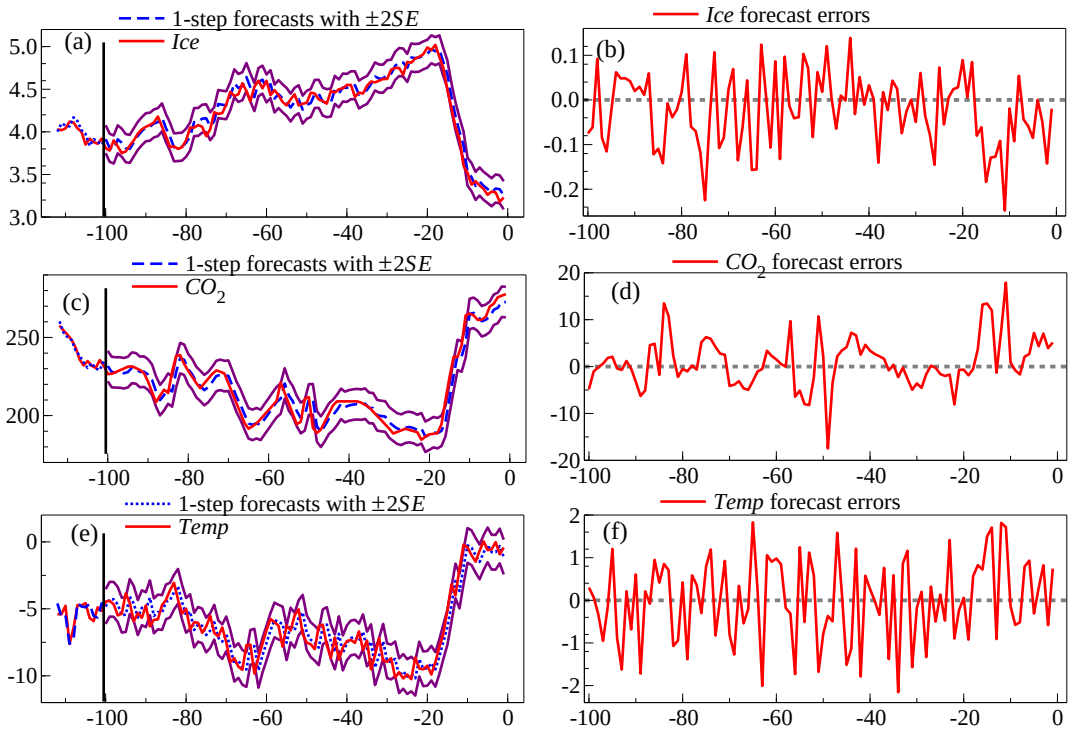

Figure 8: A hundred 1-step ahead forecasts at 1000-year measures with forecast intervals at ±2SE shown by error bands: (a) for *Ice* from (7); (c) for CO_2 from (8); (e) for *Temp* from (9). (b), (d) and (f) report the associated forecast errors.

Table 5 reports their RMSFEs which are close to the in-sample standard deviations shown in the following row as $\widehat{\sigma}$s (IIS) from Table 3 for comparison, or no IIS. The model for *Ice* provides a better description of the last 100 observations than the earlier sample, even though the in-sample residual standard deviations were calculated after outliers were removed by IIS. The forecast intervals in Figure 8 could be adjusted for the likely presence of outliers in the future at roughly their rate of occurrence in the past by calculating the in-sample residual standard deviations excluding impulse indicators, as those after IIS understate the future uncertainty to some extent. The last row in Table 5 reports those 'no-IIS' $\widetilde{\sigma}$ values.

The removal of outliers has not greatly improved the in-sample fit, and omitting impulse indicators

	Ice	CO_2	*Temp*
RMSFE	0.084	5.35	0.958
$\widehat{\sigma}$s (IIS)	0.086	4.88	0.667
$\widetilde{\sigma}$s (no IIS)	0.091	5.38	0.748

Table 5: 1-step ahead root mean square forecast errors; in-sample model residual standard deviations after IIS; and model residual standard deviations without IIS.

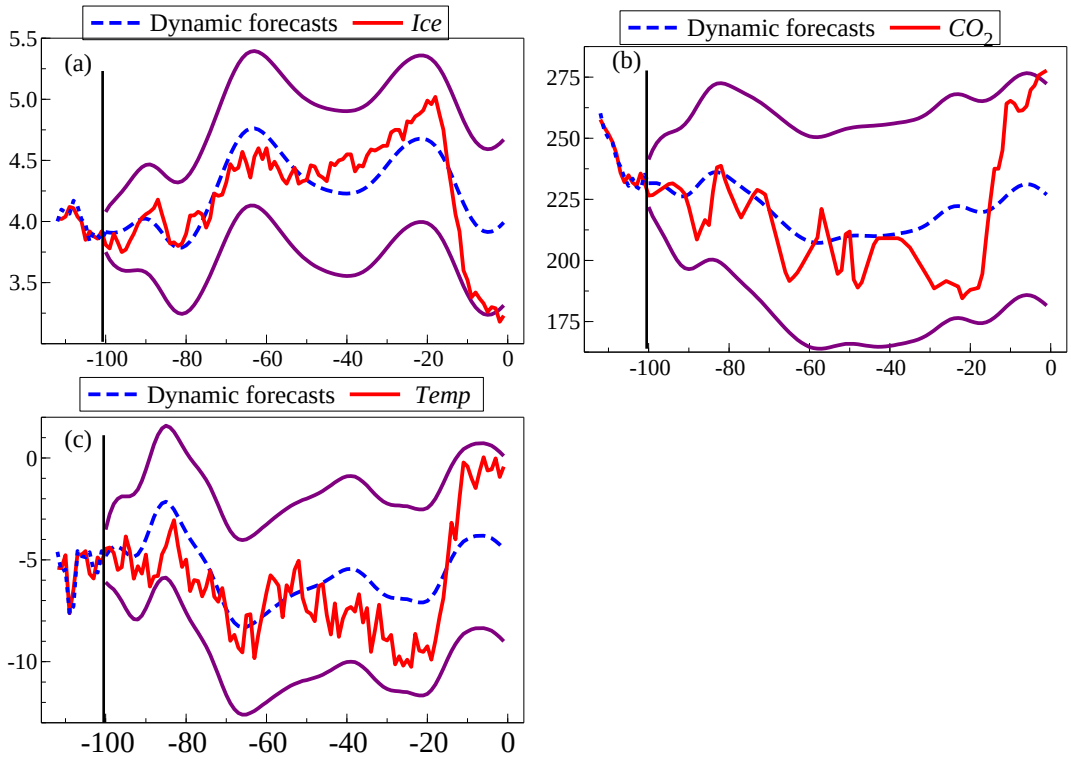

Figure 9: A hundred dynamic forecasts with ± 2.2SE error bands: (a) for *Ice* from (7); (b) for CO_2 from (8); (c) for *Temp* from (9).

would only increase the reported forecast intervals by about 10%. The table reveals that surprisingly, *Ice* provides a better description forecasting over the last 100 periods than in-sample with IIS, even though IIS could not be applied to such a 'future' period, whereas CO_2 and *Temp* forecasts are worse. Also, the RMSFE for *Ice* is smaller than the in-sample fitted $\widetilde{\sigma}$ without IIS, that for CO_2 is similar, whereas again *Temp* forecasts are worse. Looking back at Figure 5, the forecast errors for *Ice* seem less variable than the in-sample residuals on 'ocular' econometrics, less so those for CO_2, whereas those for *Temp* look somewhat more volatile.

Figure 9 shows the hundred dynamic, or multi-period ahead, forecasts with ± 2.2SE error bands to reflect the absence of indicators. These error bands depend crucially on two key assumptions, namely that the coefficients in the model remain constant, and that no new forces intervene. With the Industrial

Revolution, an additional driver of CO_2 from human fossil fuel emissions, and hence of temperature, was created, so any extension to forecast that era by an unchanged model, as considered below, is likely to reveal failure. Indeed, all three sets of forecasts either cross or are close to an error band at the end of the series. The first 60 periods are tracked quite well, but miss the changes around 20,000 years ago of ice declining rapidly, and CO_2 and temperature troughing then rising quickly. Compared to earlier changes seen in Figure 5, the last glacial period does not look unusual, a perception enhanced by Figure 10 showing the similar profiles of Ice and CO_2 over the last two cyclical periods, although the last glacial cycle persisted for longer.

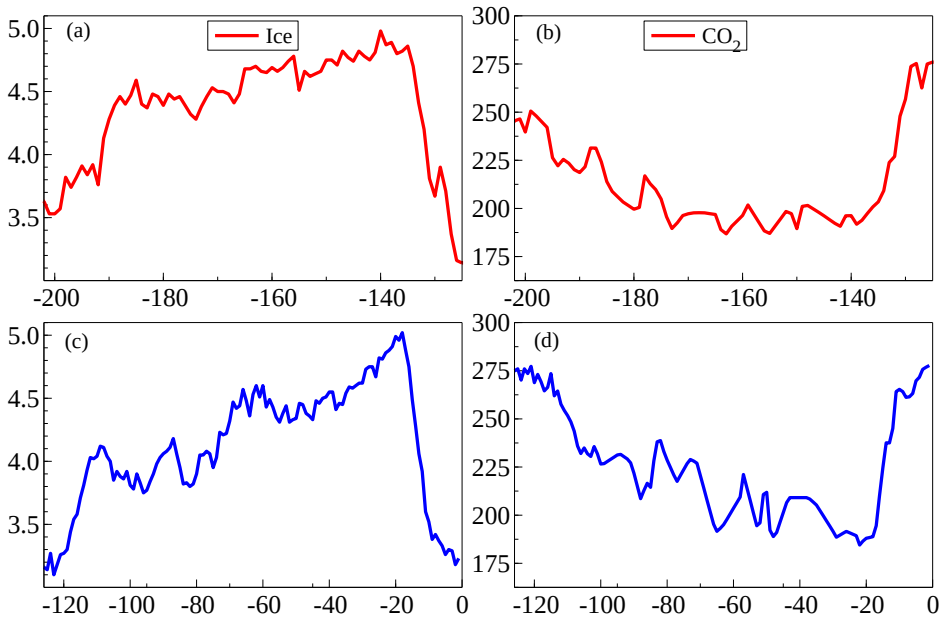

Figure 10: (a) Ice over -200 to -125; (b) CO_2 over -200 to -125; (c) Ice over -120 to -1; (d) CO_2 over -120 to -1.

5.3 When did humanity first influence the climate system?

Finally, it has been suggested that humanity began to influence the climate around the time of domesticating animals and starting farming (see e.g., William Ruddiman, 2005), so we 'zoom in' on the last 10,000 years, and re-estimate the system over the longer period to -10. The estimates are not much changed with $\chi^2_{\mathsf{OR}}(64) = 67.0$, although the contemporaneous coefficient of CO_2 on *Temp* has increased slightly to unity. Figure 11 records the multi-step forecasts over -10 to -1 for *Ice*, CO_2 and *Temp*. All the forecasts lie within their $\pm 2\mathsf{SE}$ error bands, but nevertheless they, and the fitted values for most of the previous 10,000 years, are systematically over for *Ice* and under for CO_2 and *Temp*. This matches the longer period dynamic forecasts in Figure 9, and could reflect model mis-specification, or the slowly growing divergence that might derive from the increasing influence of humanity envisaged by Ruddiman (2005). In fact, using the presence of proto-weeds that needed ground disturbance to grow in new ar-

eas, Ainit Snir *et al.* (2015) provide evidence of the origins of cultivation long before Neolithic farming occurred, dating such events to around 23,000 years ago.

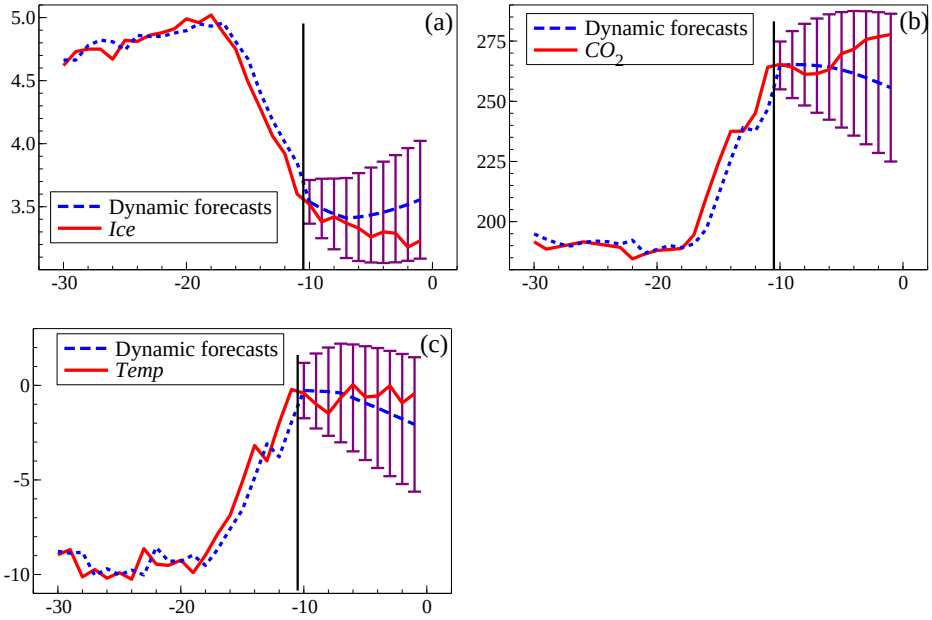

Figure 11: Multi-step forecasts over -10 to -1 of: (a) *Ice*; (b) CO_2; (c) *Temp*.

5.4 The role of CO_2

Having estimated the system up to 10,000 years ago, we changed the status of CO_2 to unmodeled and re-estimated the two-equation model for *Ice* and *Temp* conditional on CO_2. Neither fit was much improved, with $\widehat{\sigma}_{ice} = 0.085$ and $\widehat{\sigma}_{temp} = 0.688$, but now contemporaneous CO_2 is highly significant in the equation for *Ice*, with $\mathtt{t} = -3.37^{**}$, and its coefficient in the equation for *Temp* has more than doubled to 0.024, which seems implausibly large with $\mathtt{t} = 10.4^{**}$. Also, $\chi^2_{\mathsf{OR}}(43) = 198^{**}$ strongly rejects. Thus, the evidence here favours CO_2 over the Ice Ages being an endogenous response to the orbital drivers jointly with *Ice* and *Temp*.

References

Agassiz, L. (1840). *Études sur les glaciers*. Neuchâtel: Imprimerie de OL Petitpierre. Digital book on Wikisource accessed on July 22, 2019: https://fr.wikisource.org/w/index.php?title=%C3%89tudes_sur_les_glaciers&oldid=297457.

Castle, J. L. and N. Shephard (Eds.) (2009). *The Methodology and Practice of Econometrics*. Oxford: Oxford University Press.

Croll, J. (1875). *Climate and Time in Their Geological Relations, A Theory of Secular Changes of the Earth's Climate*. New York: D. Appleton.

Diebold, F. X. and G. D. Rudebusch (2019). Probability assessments of an ice-free Arctic: Comparing statistical and climate model projections. https://arxiv.org/abs/1912.10774v1 [stat.AP], 23 Dec 2019.

Doornik, J. A. and D. F. Hendry (2017). Automatic selection of multivariate dynamic econometric models. Unpublished typescript, Nuffield College, University of Oxford.

Engle, R. F., D. F. Hendry, and J.-F. Richard (1983). Exogeneity. *Econometrica 51*, 277–304.

Fisher, F. M. (1966). *The Identification Problem in Econometrics*. New York: McGraw Hill.

Geikie, A. (1863). On the Phenomena of the Glacial Drift of Scotland. *Transactions of the Geological Society of Glasgow 1*, 1–190.

Haavelmo, T. (1943). The statistical implications of a system of simultaneous equations. *Econometrica 11*, 1–12.

Hall, A. R., G. D. Rudebusch, and D. W. Wilcox (1996). Judging instrument relevance in instrumental variables estimation. *International Economic Review 37*, 283–298.

Hendry, D. F. (1976). The structure of simultaneous equations estimators. *Journal of Econometrics 4*, 51–88.

Hendry, D. F. (2009). The methodology of empirical econometric modeling: Applied econometrics through the looking-glass. In T. C. Mills and K. D. Patterson (Eds.), *Palgrave Handbook of Econometrics*, pp. 3–67. Basingstoke: Palgrave MacMillan.

Hendry, D. F. (2018). Deciding between alternative approaches in macroeconomics. *International Journal of Forecasting 34*, 119–135, with 'Response to the Discussants', 142–146.

Hendry, D. F., S. Johansen, and C. Santos (2008). Automatic selection of indicators in a fully saturated regression. *Computational Statistics 33*, 317–335. Erratum, 337–339.

Hendry, D. F. and H.-M. Krolzig (2001). *Automatic Econometric Model Selection*. London: Timberlake Consultants Press.

Hendry, D. F. and H.-M. Krolzig (2005). The properties of automatic Gets modelling. *Economic Journal 115*, C32–C61.

Hendry, D. F., M. Lu, and G. E. Mizon (2009). Model identification and non-unique structure. See Castle and Shephard (2009), pp. 343–364.

Hendry, D. F. and G. E. Mizon (1993). Evaluating dynamic econometric models by encompassing the VAR. In P. C. B. Phillips (Ed.), *Models, Methods and Applications of Econometrics*, pp. 272–300. Oxford: Basil Blackwell.

Hendry, D. F., A. J. Neale, and F. Srba (1988). Econometric analysis of small linear systems using Pc-Fiml. *Journal of Econometrics 38*, 203–226.

Imbrie, J., E. A. Boyle, S. C. Clemens, and A. *et al.*. Duffy (1992). On the structure and origin of major glaciation cycles, 1, linear responses to Milankovitch forcing. *Paleoceanography 7*, 701–738.

Jaccard, S. L., E. D. Galbraith, A. Martínez-García, and R. F. Anderson (2016). Covariation of deep Southern Ocean oxygenation and atmospheric CO2 through the last ice age. *Nature 530*, 207–210. https://doi.org/10.1038/nature16514.

Johansen, S. and B. Nielsen (2009). An analysis of the indicator saturation estimator as a robust regression estimator. See Castle and Shephard (2009), pp. 1–36.

Jouzel, J., V. Masson-Delmotte, O. Cattani, G. Dreyfus, S. Falourd, and G. *et al.*. Hoffmann (2007). Orbital and millennial Antarctic climate variability over the past 800,000 years. *Science 317*, 793–797.

Kaufmann, R. and K. Juselius (2013). Testing hypotheses about glacial cycles against the observational record. *Paleoceanography 28*, 175–184.

Kaufmann, R. K. and K. Juselius (2010). Glacial cycles and solar insolation: The role of orbital, seasonal, and spatial variations. *Climate of the Past Discussions 6*, 2557–2591. https://doi.org/10.5194/cpd-6-2557-2010.

Koopmans, T. C. (1949). Identification problems in economic model construction. *Econometrica 17*, 125–144. Reprinted with minor revisions in Hood, W. C. and Koopmans, T. C. (eds.) (1953), *Studies in Econometric Method*. Cowles Commission Monograph 14, New York: John Wiley & Sons.

Koopmans, T. C. (Ed.) (1950a). *Statistical Inference in Dynamic Economic Models*. Number 10 in Cowles Commission Monograph. New York: John Wiley & Sons.

Koopmans, T. C. (1950b). When is an equation system complete for statistical purposes? See Koopmans (1950a), Chapter 17.

Koopmans, T. C. and O. Reiersøl (1950). The identification of structural characteristics. *The Annals of Mathematical Statistics 21*, 165–181.

Lamb, H. H. (1959). Our changing climate, past and present. *Weather 14*, 299–317.

Lamb, H. H. (1995). *Climate, History and the Modern World*. London: Routledge. Second edition (First ed., 1982).

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

Lisiecki, L. E. and M. E. Raymo (2005). A Pliocine-Pleistocene stack of 57 globally distributed benthic δ^{18}O records. *Paleoceanography 20*, https://doi.org/10.1029/2004PA001071.

Lüthil, D., M. Le Floch, B. Bereiter, T. Blunier, J.-M. Barnola, U. Siegenthaler, D. Raynaud, J. Jouzel, H. Fischer, K. Kawamura, and T. F. Stocker (2008). High-resolution carbon dioxide concentration record 650,000–800,000 years before present. *Nature 453*. https://doi.org/10.1038/nature06949.

Martinez-Garcia, A., A. Rosell-Melé, W. Geibert, R. Gersonde, P. Masqué, and V. *et al.*. Gaspari (2009). Links between iron supply, marine productivity, sea surface temperature, and CO_2 over the last 1.1 Ma. *Paleoceanography 24*, https://doi.org/10.1029/2008PA001657.

Mavroeidis, S. (2003). Weak identification of forward-looking models in monetary economics. Manuscript, University of Amsterdam.

Milankovitch, M. (1969). *Canon of insolation and the ice-age problem*. Washington, D.C: National Science Foundation. English translation by the Israel Program for Scientific Translations of *Kanon der Erdbestrahlung und seine Anwendung auf das Eiszeitenproblem*, Textbook Publishing Company, Belgrade, 1941.

Paillard, D. (2001). Glacial cycles: towards a new paradigm. *Reviews of Geophysics 39*, 325–346.

Paillard, D. (2010). Climate and the orbital parameters of the Earth. *Compte Rendus Geoscience 342*, 273–285.

Paillard, D., L. D. Labeyrie, and P. Yiou (1996). Macintosh program performs time-series analysis. *Eos Transactions AGU 77*, 379.

Parrenin, F., J.-M. Barnola, J. Beer, T. Blunier, and E. *et al.*. Castellano (2007). The EDC3 chronology for the EPICA Dome C ice core. *Climate of the Past 3*, 485–497.

Parrenin, F., J.-R. Petit, V. Masson-Delmotte, and E. *et al.*. Wolff (2012). Volcanic synchronisation between the EPICA Dome C and Vostok ice cores (Antarctica) 0-145 kyr BP. *Climate of the Past 8*, 1031–1045.

Phillips, P. C. B. (1989). Partially identified econometric models. *Econometric Theory 5*(2), 181–240.

Pistone, K., I. Eisenman, and V. Ramanathan (2019). Radiative heating of an ice-free Arctic Ocean. *Geophysical Research Letters 46*, 7474–7480.

Pretis, F. (2019). Econometric models of climate systems: The equivalence of two-component energy balance models and cointegrated VARs. *Journal of Econometrics*, https://doi.org/10.1016/j.jeconom.2019.05.013.

Pretis, F. and R. K. Kaufmann (2018). Out-of-sample Paleo-climate simulations: Testing hypotheses about the Mid-Brunhes event, the Stage 11 paradox, and orbital variations. Discussion paper, University of Victoria, Canada.

Richard, J.-F. (1980). Models with several regimes and changes in exogeneity. *Review of Economic Studies 47*, 1–20.

Rothenberg, T. J. (1971). Identification in parametric models. *Econometrica 39*, 577–592.

Rothenberg, T. J. (1973). *Efficient Estimation with A Priori Information*. Number 23 in Cowles Foundation Monograph. New Haven: Yale University Press.

Ruddiman, W. (2005). *Plows, Plagues and Petroleum: How Humans took Control of Climate*. Princeton: Princeton University Press.

Siddall, M., E. J. Rohling, A. Almogi-Labin, C. Hemleben, D. Meischner, and I. *et al.*. Schmelzer (2003). Sea-level fluctuations during the last glacial cycle. *Nature 423*, 853–858.

Snir, A., D. Nadel, I. Groman-Yaroslavski, Y. Melamed, M. Sternberg, and O. *et al.*. Bar-Yosef (2015). The origin of cultivation and proto-weeds, long before Neolithic farming. *PLoS ONE 10(7): e0131422*. https://doi.org/10.1371/journal.pone.0131422.

Staiger, D. and J. H. Stock (1997). Instrumental variables regression with weak instruments. *Econometrica 65*, 557–586.

Stock, J. H. and J. H. Wright (2000). GMM with weak identification. *Econometrica 68*, 1055–1096.

Vaks, A., A. J. Mason, and S. F. M. Breitenbach, *et al.* (2019). Palaeoclimate evidence of vulnerable permafrost during times of low sea ice. *Nature 577*, 221–225. https://doi.org/10.1038/s41586-019-1880-1.

Wang, J. and E. Zivot (1998). Inference on a structural parameter in instrumental variables regression with weak instruments. *Econometrica 66*, 1389–1404.

Zivot, E., R. Startz, and C. R. Nelson (1998). Valid confidence intervals and inference in the presence of weak instruments. *International Economic Review 39*, 1119–1144.

6 Simultaneous Equations Model of the Ice Ages variables with IIS

$$\begin{aligned}\widehat{Ice}_t = {} & \underset{(0.34)}{1.43} + \underset{(0.015)}{0.86}\ Ice_{t-1} - \underset{(0.002)}{0.020}\ Temp_{t-1} + \underset{(31)}{102}\ Ec_t - \underset{(32)}{101}\ Ec_{t-1} - \underset{(0.014)}{0.040}\ Ob_{t-1} \\ & - \underset{(1.30)}{5.07}\ EcOb_t + \underset{(1.36)}{5.05}\ EcOb_{t-1} - \underset{(1.03)}{4.97}\ EcPr_t - \underset{(0.08)}{0.19}\ 1_{\{-335\}} - \underset{(0.09)}{0.21}\ 1_{\{-276\}} \\ & - \underset{(0.08}{0.24}\ 1_{\{-268\}} - \underset{(0.09)}{0.17}\ 1_{\{-243\}} - \underset{(0.089)}{0.33}\ 1_{\{-229\}} - \underset{(0.09)}{0.22}\ 1_{\{-203\}} + \underset{(0.09)}{0.24}\ 1_{\{-191\}} \\ & - \underset{(0.09)}{0.27}\ 1_{\{-131\}} + \underset{(0.09)}{0.41}\ 1_{\{-129\}} - \underset{(0.09)}{0.20}\ 1_{\{-127\}} \end{aligned} \tag{10}$$

$$\begin{aligned}\widehat{CO}_{2,t} = {} & \underset{(32)}{218} + \underset{(0.018)}{0.85}\ CO_{2,t-1} + \underset{(0.18)}{1.34}\ Temp_{t-1} + \underset{(342)}{1400}\ Ec_t - \underset{(647)}{3070}\ Ec_{t-1} \\ & - \underset{(2.3)}{13}\ Ob_{t-1} + \underset{(23)}{71}\ EcOb_{t-1} + \underset{(0.047)}{0.23}\ ObSq_t + \underset{(4.2)}{24}\ 1_{\{-339\}} + \underset{(4.9)}{18}\ 1_{\{-335\}} \\ & + \underset{(4.2)}{15}\ 1_{\{-307\}} + \underset{(4.2)}{17}\ 1_{\{-251\}} + \underset{(4.8)}{16}\ 1_{\{-244\}} + \underset{(4.9)}{17}\ 1_{\{-243\}} - \underset{(4.9)}{19}\ 1_{\{-241\}} \\ & - \underset{(4.9)}{16}\ 1_{\{-225\}} + \underset{(4.9)}{12}\ 1_{\{-202\}} + \underset{(4.2)}{11}\ 1_{\{-199\}} + \underset{(4.8)}{11}\ 1_{\{-188\}} + \underset{(4.8)}{16}\ 1_{\{-178\}} \\ & + \underset{(4.9)}{17}\ 1_{\{-131\}} - \underset{(4.3)}{13}\ 1_{\{-127\}} \end{aligned} \tag{11}$$

$$
\begin{aligned}
\widehat{Temp_t} = &- \underset{(0.69)}{2.5} + \underset{(0.023)}{0.88}\ \mathit{Temp}_{t-1} + \underset{(0.003)}{0.0080}\ \mathit{CO}_{2,t} - \underset{(37)}{301}\ \mathit{Ec}_t + \underset{(2.5)}{23}\ \mathit{EcOb}_t \\
&- \underset{(1.9)}{9.8}\ \mathit{EcOb}_{t-1} + \underset{(7.1)}{26}\ \mathit{EcPr}_t + \underset{(0.67)}{2.12}\ 1_{\{-335\}} - \underset{(0.57)}{2.26}\ 1_{\{-295\}} - \underset{(0.57)}{1.88}\ 1_{\{-247\}} \\
&+ \underset{(0.66)}{2.24}\ 1_{\{-244\}} + \underset{(0.67)}{1.89}\ 1_{\{-243\}} - \underset{(0.66)}{1.85}\ 1_{\{-241\}} - \underset{(0.66)}{1.65}\ 1_{\{-225\}} \\
&- \underset{(0.58)}{1.18}\ 1_{\{-191\}} + \underset{(0.58)}{1.28}\ 1_{\{-189\}} - \underset{(0.66)}{2.00}\ 1_{\{-188\}} - \underset{(0.58)}{2.02}\ 1_{\{-187\}} \\
&+ \underset{(0.66)}{1.37}\ 1_{\{-178\}} - \underset{(0.57)}{2.28}\ 1_{\{-176\}} + \underset{(0.57)}{1.85}\ 1_{\{-172\}} - \underset{(0.58)}{2.81}\ 1_{\{-171\}} + \underset{(0.57)}{2.10}\ 1_{\{-163\}} \\
&- \underset{(0.57)}{2.03}\ 1_{\{-158\}} + \underset{(0.67)}{2.53}\ 1_{\{-131\}} - \underset{(0.58)}{2.13}\ 1_{\{-126\}} - \underset{(0.57)}{2.60}\ 1_{\{-109\}}
\end{aligned}
\tag{12}
$$