Anthropogenic Influences on Atmospheric CO_2

David F. Hendry and Felix Pretis*

*Institute for New Economic Thinking at the Oxford Martin School

and Economics Department, University of Oxford, UK

Abstract

We identify anthropogenic contributions to atmospheric CO_2 measured at Mauna Loa using a statistical automatic model selection algorithm (*Autometrics*). We find that vegetation, temperature and other natural factors alone cannot explain the trend or the variation in CO_2 growth. Industrial production components, driven by business cycles and economic shocks, are highly significant contributors.

JEL classifications: Q5, C5.
KEYWORDS: Climate change, CO_2 emissions, impulse-indicator saturation, *Autometrics*.

1 Introduction

We identify anthropogenic contributions to atmospheric CO_2 measured at Mauna Loa using the statistical automatic model selection algorithm *Autometrics*. Estimating the determinants of atmospheric CO_2 is traditionally a challenge due to the complex systems of data involved. CO_2 is a highly autocorrelated, non-stationary time series, and globally there exist a large number of potential carbon sources and sinks. There is mixed evidence in the literature on anthropogenic contributions to atmospheric CO_2: the long-term trend is widely attributed to human factors, while the main seasonal fluctuations are thought to be driven by the biosphere. However, the statistical measures applied are often somewhat unsatisfactory due to the complexities of dealing with large numbers of variables.

Over the long run of geological time, evidence of repeated glaciations, and of coal and oil deposits from extinct tropical forests reveal that atmospheric CO_2 has varied greatly, and manifestly without any anthropogenic influence, including very low levels and levels as high as 1000 parts per million (ppm): see e.g., Hoffman and Schrag (2000) (Hendry, 2011, provides a summary). In the more recent half million years of 'ice ages', natural fluctuations include highs and lows of 300 and 180 ppm from Antarctic ice sheet drilling (see Juselius and Kaufmann, 2009). Finally, in the last 10,000–12,000 years, humanity has transformed planet Earth, replacing forests by agriculture and creating an industrial world (see e.g., Ruddiman, 2005). Against the background of such movements, it is important to establish that recent levels of atmospheric CO_2 are not merely a natural event, but have an anthropogenic signature.[1]

*This research was supported in part by grants from the Open Society Institute and the Oxford Martin School. We thank Roger Fouquet for useful comments and Qin Duo for helpful notes on Chinese production. Forthcoming in Fouquet, R. (ed.) *Handbook on Energy and Climate Change*, Edward Elgar.

[1] 1 ppm by volume of CO_2 in the atmosphere is approximately equal to 2.13 giga tons of carbon: see Clark (1982) and http://cdiac.ornl.gov/pns/convert.html.

We introduce a new approach to modeling changes in atmospheric CO_2 using a model selection algorithm which allows for a larger number of potential variables than observations without *a priori* forcing any to be significant or to be excluded. Using this method, the main relevant explanatory variables are determined and their magnitudes estimated while irrelevant factors are dropped from the model.

The model controls for a number of natural carbon sources and sinks: vegetation measured by the Normalized Difference Vegetation Index (NDVI); temperature (measured as anomalies in the Northern hemisphere); weather phenomena (measured through the Southern Oscillation Index); as well as accounting for dynamic transport by including seasonal interaction terms. This allows an estimate of the anthropogenic contribution to CO_2 as measured by industrial output indices and fossil fuel use for different geographical areas. The resulting estimates describe the direct effects on CO_2 growth and the proportional contribution of each factor. We find that vegetation, temperature and other natural factors alone cannot explain either the trend or all the variation in CO_2 growth. Industrial production components, driven by business cycles and shocks, are highly significant contributors.

Section 2 provides an overview of related literature, 3 discusses model selection, impulse indicator saturation (IIS)–which we use to detect multiple breaks in the models–and the *Autometrics* algorithm. Section 4 describes the data used, section 5 outlines the estimation procedures and section 6 reports the main results. Section 7 concludes.

2 Literature review

There is a plethora of literature on atmospheric CO_2 and its link to anthropogenic factors. Key aspects in the literature are finding an appropriate measure for anthropogenic activity and sufficient controls for natural effects such as vegetation and oceanic absorption. Atmospheric CO_2 has been measured consistently and regularly since 1958, mostly due to the effort of Charles Keeling who initiated and supported the measurement at Mauna Loa, Hawaii (and later other measurement stations: see (Sundquist and Keeling, 2009). This led to the now well established and often cited "Keeling Curve", showing the increasing trend and highly seasonal pattern in atmospheric CO_2 (figure 1). We are primarily interested in identifying the anthropogenic contribution to atmospheric CO_2, so the following section reviews existing evidence and factors that need to be included in models.

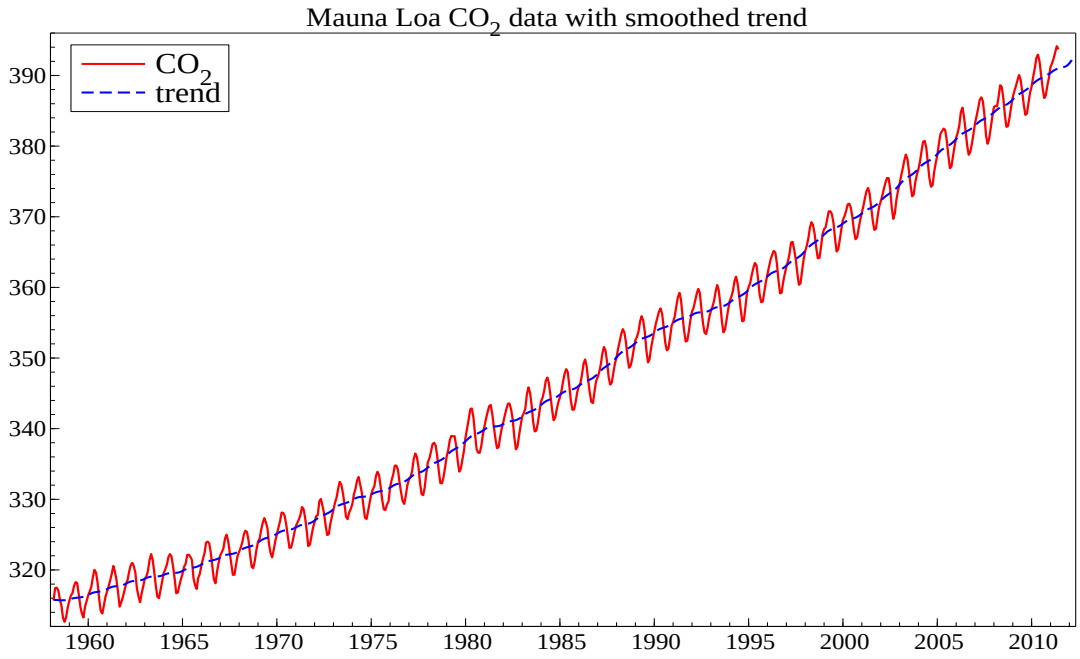
Sources and sinks

Anthropogenic sources

While atmospheric CO_2 has been consistently measured at multiple sites for a long time (see Keeling, Bacastow, Brainbridge, Ekdahl, Guenther and Waterman, 1976, and Sundquist and Keeling, 2009), the choice of anthropogenic variables is less straight forward. Three variables are regularly used: estimates of fossil fuel emissions, population, and cement production. In most cases, variables are measured on an annual basis and interpolated to monthly frequencies.

A standard measure for estimating fossil fuel emissions is the product of the amount of fuel produced, the proportion of the fuel that is oxidized, and the fuel carbon content (Marland and

Figure 1: Atmospheric CO_2 measured at Mauna Loa—the "Keeling Curve"



Rotty, 1984). Variations of this data are used in Erickson, Mills, Gregg, J. Blasing, Hoffmann, Andres, Devries, Zhu and Kawa (2008), Jones and Cox (2005), Randerson, Thompson, Conway, Fung and Field (1997), and Nevison, Mahowald, Doney, Lima, van der Werf, Randerson, Baker, Kasibhatla and McKinley (2008). Recent data using this methodology are available at an annual frequency by Marland, Boden and Andres (2011). In contrast, Hofman, Butler and Tans (2009) as well as Newell and Marcus (1987) focus on population as a measure of human industrial output. Granados, Ionides and Carpintero (2011) extend the model of population by including a measure of global GDP. Cement production is a further major component of anthropogenic emissions. CO_2 emissions in production through limestone calcination, kiln operation and power generation are estimated to make up approximately five percent of anthropogenic emissions (Worrell, Price, Martin, Hendriks and Meida, 2001).

These measures provide a good starting point to capture anthropogenic emissions in the long run, but due to the annual measurement frequency do not capture short-run fluctuations and seasonality. Population and GDP are too broad as measures to capture variability other than a trend. Measurement could be improved through supplementing overall fuel emissions by disaggregate individual fuel consumption. However, most importantly, a higher-frequency (monthly) anthropogenic measure is required to capture seasonal and short-term effects.

Terrestrial biosphere and transport

Aside from anthropogenic emissions, the terrestrial biosphere (vegetation) is a major factor in the carbon cycle's sources and sinks. When trying to model anthropogenic contributions to CO_2 it is therefore important to account for the most important factors in the natural world.

Atmospheric carbon dioxide falls and rises seasonally each year due to photosynthetic activity (during Summer) and respiratory release (during Winter) of CO_2 in the terrestrial biosphere

(Keeling, Chin and Whorf, 1996, Buermann, Lintner, Koven, Angert, Pinzon, Tucker and Fung, 2007). The intensity of these effects depends on the length of the growing season, a fertilization effect (feedback to plant growth from increased CO_2), and shifts in seasonal patterns (Kaufmann, Paletta, Tian, Myeni and D'Arrigo, 2008). The literature on atmospheric CO_2 suggests a strong link between the Earth's biosphere and the level of CO_2.

The Normalized Difference Vegetation Index (NDVI) provides a direct measure of photosynthetic activity. As Myeni, Hall, Sellers and Marshak (1995) describe, chlorophyll found in plants absorbs visible light for photosynthesis and reflects near infra-red light. The more active a plant is (indicated by higher density of green leaves), the more visible radiation is absorbed and the more near infra-red is reflected. Thus, the difference between the two measures increases with higher leaf density. Using satellite-based remote sensing the intensity of reflected visible and infra-red light can be measured. Using the ratio of the difference between the two measures, the NDVI is defined by:

$$NDVI = \frac{\rho_{Visible} - \rho_{NearIR}}{\rho_{Visible} + \rho_{NearIR}} \tag{1}$$

where $\rho_{Visible}$, and ρ_{NearIR} are measures for visible and near infra-red light respectively. NDVI is therefore an index ranging from -1 to $+1$, with values around 0 denoting non-vegetation objects, and values around > 0.7 indicating dense vegetation (Tucker, Pinzon, Brown and GIMMS/GSFC/NASA, 2010). The NDVI provides a direct measure for vegetation activity with high values close to 1 during the growing season (Summer in the Northern Hemisphere), and lower values closer to 0 during the less active season (Winter in the Northern Hemisphere). Kaufmann *et al.* (2008) investigate the link between NDVI and atmospheric carbon dioxide using econometric methods and find evidence NDVI values "Granger cause" CO_2. There is also evidence of a feedback effect of increased CO_2, leading to enhanced vegetation activity.

Due to transport airflows, the primary influence of vegetation on measured atmospheric CO_2 depends on the location of the measurement station. In the case of Mauna Loa, Hawaii, the seasonal variation due to the biosphere is driven by long-range transport from Eurasia during Winter and short-range transport from North America during Summer (see Buermann *et al.*, 2007, Taguchi, Murayama and Higuchi, 2003). Narrowing down the time frame, Levin, Graul and Trivett (1995) suggests that Eurasian airflows dominate from October to June while North American airflows are dominant from July to September.

In terms of long-term development of terrestrial vegetation, there has been a greening trend—an increase of the growing season in the Northern Hemisphere (Lucht, Prentice, Myeni, Sitch, Friedlingstein, Cramer, Bousquet, Buermann and Smith, 2002). This trend, however, was interrupted by the eruption of Mount Pinatubo in June 1991 which led to a decline in vegetation during 1992–1993. This poses the question whether volcanic influence on atmospheric CO_2 needs to be controlled for separately from the biosphere. Lucht *et al.* (2002) state that the main channel through which volcanoes affect atmospheric CO_2 is indirect through temperature while Hofman *et al.* (2009) propose that the Pinatubo eruption enhanced photosynthesis through scattered sunlight. Overall, Gerlach (2011) finds the direct effect of volcanic activity on measured atmospheric carbon dioxide to be small, thus accounting for vegetation and temperature is expected to be sufficient without considering volcanoes separately. However, sudden breaks due to large eruptions will be detected through our IIS procedure.

Oceanic absorption and El Niño

A large amount of carbon dioxide is transferred between the ocean and the atmosphere, where the ocean acts as both a source and a sink of atmospheric CO_2. There are two main factors determining oceanic carbon absorption and release (Bacastow (1976)). First, generally ocean absorption of CO_2 is governed by the difference in partial pressure between the atmospheric and oceanic CO_2:

$$\frac{dp}{dt} = -k \cdot (p - P) \tag{2}$$

where p is the partial pressure of atmospheric CO_2, P is the partial pressure of oceanic CO_2 beneath a surface layer and k, a variable capturing layer thickness and wind. These values are affected primarily by temperature and wind speed. Temperature is a key factor in oceanic CO_2 partial pressure (Copin-Montigut, 1988), as higher temperature increases the partial pressure and thus reduces absorption. Increased wind speed increases k, so leads to higher oceanic CO_2 uptake. The second effect is upwelling—dense cold water driven to the surface releases carbon dioxide stored in the ocean. Absorption and upwelling play opposite roles; which dominates is debated and depends on the geographical region.

The atmospheric fluctuations of air pressure differences known as Southern Oscillation affect oceanic absorption through the two channels described above. Southern Oscillation describes the change of air pressure differences between Tahiti and Darwin, Australia (see Troup, 1965, Bacastow, 1976, Keeling and Revelle, 1985). It is measured as an index (SOI) from the Australian Bureau of Meteorology (2011), and defined as:

$$SOI = 10 \cdot \frac{\Delta P_t - \overline{\Delta P}_t}{\sigma_{\Delta P_t}} \tag{3}$$

where ΔP_t is the difference in the average of mean sea level pressure between Tahiti and Darwin for month t. $\overline{\Delta P}_t$ is the long-run average of ΔP_t and $\sigma_{\Delta P_t}$ is the long-run standard deviation of ΔP_t for the given month respectively. Negative values of the SOI are generally referred to as El Niño years, while positive values correspond to episodes of La Niña. However, the effect on oceanic absorption is not so clear cut. Episodes of La Niña ($SOI > 0$) are associated with increased wind speeds (increase in k in equation (3)) thus making uptake easier. Nevertheless, increased wind also increases upwelling which leads to a release of oceanic CO_2. Bacastow, Keeling and Whorf (1985) suggests that easier absorption outweighs upwelling during episodes of La Niña ($SOI > 0$) resulting in higher absorption of carbon dioxide by the ocean when the SOI is positive. In turn, this implies less absorption during El Niño ($SOI < 0$) years.

On the contrary, Francey, Tanis, Allison, Enting, White and Trolier (1995) find the opposite—during La Niña years ($SOI > 0$) oceanic absorption is relatively lower because of the large upwelling effect. Keeling and Revelle (1985) side with Francey on the theoretical model that upwelling should outweigh increased absorption, but empirically find that less atmospheric CO_2 is absorbed during El Niño episodes ($SOI < 0$) which agrees with Bacastow's (1976) findings.

Another factor that is not often considered in the literature is carbon dioxide use by oceanic algae, as Ritschard (1992) mentions. Nevertheless, data on algae is limited as they are not covered by the NDVI satellite measures, and consequently are not considered in our study.

Looking at the bigger picture, most evidence suggests that the ocean has become a carbon sink for anthropogenic emissions (Christopher, Feely, Gruber, Key, Lee, Bullister, Wanninkhof, Wong, Wallace, Tilbrook, Millero, Peng, Kozyr, Ono and Rios, 2007). As increased CO_2 emissions into

the atmosphere increased atmospheric partial pressure of CO_2, absorption by the ocean increased due to the pressure difference. However, the absolute magnitude of this effect is not known, and they estimate approximately 48 percent of fossil fuel emissions are absorbed in the ocean; and atmospheric CO_2 would be approximately 55 ppm (parts per million) higher if there were no oceanic uptake. Orr, Maier-Reimer, Mikolajewicz, Monfray, Sarmiento, Toggweiler, Taylor, Palmer, Gruber, Sabine, Le Quere, Key and Boutin (2001) similarly find the ocean to be a net carbon sink but, as well as Nevison *et al.* (2008), suggest that most models over-estimate the proportion of carbon dioxide emissions absorbed by the ocean.

In order to capture the key factors determining oceanic uptake when modeling atmospheric CO_2, it is important to account for temperature (which crucially affects partial pressure) and control for Southern Oscillation, even though the effects of El Niño and La Niña are not fully understood.

Modeling methodology

The models for atmospheric CO_2 can broadly be classed into two categories: atmospheric transport and statistical models. Data for both is often decomposed into a long-run trend, cycle, and noise using Fourier decompositions or the Hodrick–Prescott (HP) filter.

Atmospheric transport describes spatial three dimensional models with vertical levels based on solving the fundamental equations for conservation of mass, momentum, and energy. For different latitude and longitude grid resolutions, these models simulate carbon emissions and global transport (for the set-up and methodology of these models see Hansen, Russel, Rind, Stone, Lacis, Lebedeff, Ruedy and Travis (1983), and Kawa, Erickson, Pawson and Zhu (2004). Within this group of models, Erickson *et al.* (2008), Nevison *et al.* (2008), Randerson *et al.* (1997) and Keeling, Whorf, Whalen and van der Plicht (1995) use annual emissions data to analyse anthropogenic effects.

Statistical approaches also vary in methodology. Thoning and Tans (1989), Keeling *et al.* (1976) and Enting (1987) use Fourier decompositions to study trend and seasonal cycle. Granados *et al.* (2011) and Granados, Ionides and Carpintero (2009) use cointegration and time-series regression to study links between population, GDP and HP-filtered CO_2 growth. However, there is no actual measure of anthropogenic emissions used in these studies. Hofman *et al.* (2009) use regression and graphical comparisons of carbon and population, while Jones and Cox (2005) regress growth rates of CO_2 on global emissions and cement production. Newell and Marcus (1987) look at the simple correlation between levels of carbon dioxide and global population.

Concerning modeling methods, atmospheric transport models and many statistical techniques are widely applied and well documented. However, they suffer from similar problems. Often inappropriate statistical techniques are applied without considering the time-series properties of the data. Correlation of time series alone is not an appropriate measure of dependence between them. The low frequency of measurement of emissions data is problematic and models are restricted by an initial choice of a small number of independent variables. Original data are rarely used: instead series are decomposed. This step is not necessary *a priori*, especially when explanatory variables that are seasonal themselves are available. Additionally, the regression analyses applied in many papers are not robust to outliers or structural breaks, do not always handle non-stationarity, and present few tests for mis-specification.

Summary of the main findings

The long-run trend in increasing carbon dioxide is clearly driven by anthropogenic factors, whereas the short-run seasonal fluctuations and changes in amplitude are mostly attributed to changes in the biosphere.

Long-term trend

The long-term trend in atmospheric CO_2 is fossil fuel induced. Pre-industrial levels of CO_2 are estimated to be around 260-280 ppm (see Wigley, 1983, and Hofman *et al.*, 2009) based on ice core, tree ring, and oceanic data. Consistent and repeated measurement, starting with Keeling's work in 1958, have documented the rise in CO_2 to a current level of approximately 390 ppm measured at Mauna Loa. The rate of increase of CO_2 is proportional to combustion of fossil fuels (see Keeling, 1973, Keeling *et al.*, 1976, Keeling *et al.*, 1995, Thoning and Tans, 1989). Using population as a proxy measure for emissions yields similar results (Hofman *et al.*, 2009, Newell and Marcus, 1987, Granados *et al.*, 2009, and Granados *et al.*, 2011).

Given the large departure from pre-industrial levels, Hansen, Sato, Kharecha, Beerling, Berner, Masson-Delmotte, Pagani, Raymo, Royer and Zachos (2008) investigate the question of a target level CO_2. They suggest that (at the time of their writing) the level of 385 ppm is too high to maintain climate conditions that current life has adapted to. Levels of 450 ppm in the Caenozoic era were associated with near ice-free conditions. Consequently, they propose a target level of at most 350 ppm.

Seasonal variation and amplitude

Seasonal fluctuations and changes in amplitude are mainly attributed to factors in the biosphere rather than industrial emissions. There are two effects described in the literature, one is the general pattern of seasonality, the second is an increase in the amplitude of this seasonality. In particular a perceived increase in the growing season is alleged to be the driving force behind increases in amplitude.

Many studies propose that the seasonal component of atmospheric CO_2 reflects the inter-annual uptake by plants. This is supported by the fact that the amplitude of this seasonality for a given season decreases towards the equator (Keeling *et al.*, 1976). In particular, Enting (1987) argues that vegetation is sufficient to account for most of the inter-annual variation and that economic data does not show the required seasonality. While the peak to trough ratio measured at Mauna Loa was approximately 0.8 for the time period Enting investigates, he suggests that industrial emissions are not sufficient to cause this seasonal change. However, as is obvious from many economic time series, there is high seasonality in production and therefore in emissions.

The amplitude of the seasonal effect has been increasing over time. Keeling *et al.* (1996), Randerson *et al.* (1997), Kohlmaier, Sire, Janecek, Keeling, Piper and Revelle (1989), and Bacastow *et al.* (1985) characterise the increase as a result of a lengthening growing season with only a very small effect directly from fossil fuel emissions. The effect from anthropogenic emissions in these studies ranges from 0.01 to 0.2 percent on the change of amplitude. Additionally, Keeling *et al.* (1995) find that changes in the overall growth rate of CO_2 are driven by changes in vegetation and temperature rather than changes in industrial emissions.

A major issue with many of the above mentioned studies is that anthropogenic emissions and production data are measured annually and therefore do not have the required frequency to be able to account for seasonal fluctuations. In a recent paper, Erickson *et al.* (2008) investigate this issue and find that economic data would suggest the highest anthropogenic fluxes occur at the same time as the respiration phase of plants (Winter in the Northern Hemisphere). Once models account for seasonality in fuel consumption this will then lead to a diminished effect of seasonality from the biosphere.

Contribution of this chapter

There are recurring problems with existing models of anthropogenic contributions to CO_2. Climate and atmospheric carbon fluxes are complex systems, nevertheless, many of the models are restricted by *a priori* selections of explanatory variables. The data used to account for anthropogenic emissions are often measured at too low a frequency to capture any seasonality. The main series of carbon dioxide is often decomposed into cycles and trends, something that is not necessary if the explanatory data is measured at a reasonable frequency. On the one hand, a significant number of the papers that approach the problem from a statistics or economic point of view do not sufficiently control for the biosphere or other natural factors. On the other hand, many models coming from a natural science background use statistical methods that are ill fitted given the time-series characteristics of the data. Modern econometric methods can provide an interdisciplinary solution to these problems.

To address these issues, we introduce an extended General to Specific (*Gets*) modeling approach based on automatic model selection and the theory of reduction. This allows for a large number of candidate explanatory variables, in particular, models can be estimated with more explanatory variables than observations. It is therefore possible to include many lags to capture time dynamics as well as a wide range of controls for natural factors and industrial output measures. As our main measure of anthropogenic productivity and emissions is industrial production (measured at monthly intervals), the data is analysed as a whole without requiring prior decompositions into trends and cycles. Models are also not restricted to a tight *a priori* selection of variables. Using IIS, the methods are robust to outliers and structural breaks, handle unit roots reasonably well, and provide a straight-forward method of testing for mis-specification.

Overall, the literature indicates a clear necessity to control for the biosphere, temperature, El Niño effects, and long (as well as short-run) anthropogenic measures. Intuitively, our approach is to utilise a large number of potential determinants controlling for the above mentioned factors, and then use automatic model selection techniques to determine which forces are significant. Starting with a theory-based, but very broad General Unrestricted Model (GUM), the initial system is reduced to a specific model. This is a comparatively agnostic and data-driven approach that imposes few restrictions on explanatory variables while being robust to sudden shifts (structural breaks).

3 Methodology

The carbon cycle, with many potential sinks and sources is a complex system which makes it near-impossible to correctly specify an appropriate model a priori. To model complex equations, we rely on general to specific modeling approaches (see Campos, Ericsson and Hendry, 2005). The

unknown data generating process (DGP) is the underlying structure that creates the data. Empirical modeling will always deal with a subset of variables of the DGP, thus an important factor is the local data generating process (LDGP)—the generating process in the space of the variables under analysis: see Hendry (2009). The approach, therefore, is to construct a set of data based on broad theoretical assumptions, which nests the LDGP, then within this set, reduce the model from its general form down to a specific representation. This is a two step procedure. One: define a set of N variables that include the LDGP as a sub-model. Two: starting with that general model as a good approximation of the overall properties of the data, reduce its complexity by removing insignificant variables, while checking that at each reduction the validity of the model is preserved. This is the basic framework of *Gets* modeling.

This section introduces theoretical concepts of model selection, their use in mis-specification testing, followed by the introduction of impulse indicator saturation (IIS) and its generalized version. All these concepts are then united and applied through the automatic search algorithm *Autometrics*. The algorithm combines these features through automated selection based on *Gets* while handling more variables than observations with IIS for detecting breaks and outliers, and mis-specification testing.

Model selection

The theory of reduction characterizes the operations implicitly applied to the DGP to obtain the local LDGP. Choosing to analyze a set of variables, denoted $\mathbf{y}_t, \mathbf{x}_{1,t}$, determines the properties of the LDGP, and hence of any models of $\mathbf{y}_t$ given $\mathbf{x}_{1,t}$ (with appropriate lags, non-linear transforms thereof, etc.). A congruent model is one that matches the empirical characteristics of the associated LDGP, evaluated by a range of mis-specification tests (see e.g., Hendry, 1995, and section 3). A model is undominated if it encompasses, but is not encompassed by, all other sub-models (see e.g., Mizon and Richard, 1986, Hendry and Richard, 1989, and Bontemps and Mizon, 2008).

Mis-specification testing

Using a large number of variables with IIS (discussed in the next section) also provides a new view of model evaluation: to avoid mis-specification and non-constancy, start as general as possible within the theoretical framework, using all the available data unconstrained by $N > T$, (where N is the number of variables and T the number of observations) retaining the theory inspired variables and only selecting over the additional candidates.

Even so, this approach does not obviate the need to test the specification of the auxiliary hypotheses against the possibility that the errors are not independent, are heteroskedastic (non-constant variance) or non-normal, that the parameters are not constant, that there is unmodeled non-linearity, and that the conditioning variables are not independent of the errors. When $N << T$, the first five are easily tested in the initial general model; otherwise their validity can be checked only after a reduction to a feasible sub-model. Congruence is essential not only to ensure a well specified final selection, but also for correctly-calibrated decisions during selection based on Gaussian significance levels, which IIS will help ensure.

More variables than observations: $N > T$

The model selection approach introduced here allows for more variables than observations to be used in modeling ($N > T$). For *Autometrics* this was first introduced through impulse indicator saturation, and has recently been extended to the general case.

Impulse indicator saturation

The numbers, timings and magnitudes of breaks in models are usually unknown, and are obviously unknown for unknowingly omitted variables, so a 'portmanteau' approach is required that can detect location shifts at any point in the sample while also selecting over many candidate variables. To check the null of no outliers or location shifts in a model, impulse-indicator saturation (IIS) creates a complete set of indicator variables $\left\{1_{\{j=t\}}\right\}$ equal to unity when $j = t$ and equal to zero otherwise for $j = 1, \ldots, T$, and includes these in the set of candidate regressors. Although this creates T variables when there are T observations, in the 'split-half' approach analyzed in Hendry, Johansen and Santos (2008), a regression first includes $T/2$ indicators. By dummying out the first half of the observations, estimates are based on the remaining data, so any observations in the first half that are discrepant will result in significant indicators. The location of the significant indicators is recorded, then the first $T/2$ indicators are replaced by the second $T/2$, and the procedure repeated. The two sets of significant indicators are finally added to the model for selection of the significant indicators. The distributional properties of IIS under the null are analyzed in Hendry *et al.* (2008), and extended by Johansen and Nielsen (2009) to both stationary and unit-root autoregressions.

Figure 2 illustrates the 'split-half' approach for $y_t \sim \mathsf{IN}\left[\mu, \sigma_y^2\right]$ for $T = 100$ selecting indicators at a 1% significance level (denoted α). The three rows correspond to the three stages: the first half of the indicators, the second half, then the selected indicators combined. The three columns respectively report the indicators entered, the indicators finally retained in that model, and the fitted and actual values of the selected model. Initially, although many indicators are added, only one is retained. When those indicators are dropped and the second half entered (row 2), none is retained. Now the combined retained dummies are entered (here just one), and selection again retains it. Since $\alpha T = 1$, that is the average false null retention rate.

We next illustrate IIS for a location shift of magnitude λ over the last k observations:

$$y_t = \mu + \lambda 1_{\{t \geq T-k+1\}} + \varepsilon_t \tag{4}$$

where $\varepsilon_t \sim \mathsf{IN}\left[0, \sigma_\varepsilon^2\right]$ and $\lambda \neq 0$. The optimal test in this setting would be a t-test for a break in (4) at $T - k + 1$ onwards, but requires precise knowledge of the location-shift timing, as well as knowing that it is the only break and is the same magnitude break thereafter. Figure (3) records the behavior of IIS for a mean shift in (4) of $10\sigma_\varepsilon$ occurring at $0.75T = 75$. Initially, many indicators are retained (top row), as there is a considerable discrepancy between the first-half and second-half means. When those indicators are dropped and the second set entered, all those for the period after the location shift are now retained. Once the combined set is entered (despite the large number of dummies) selection again reverts to just those for the break period. Under the null of no outliers or breaks, any indicator that is significant on a sub-sample would remain so overall, but for many alternatives, sub-sample significance can be transient, due to an unmodeled feature that occurs elsewhere in the data set. Thus there is an important difference between 'outlier detection' procedures and IIS.

Figure 2: Impulse indicator saturation under the null of no break

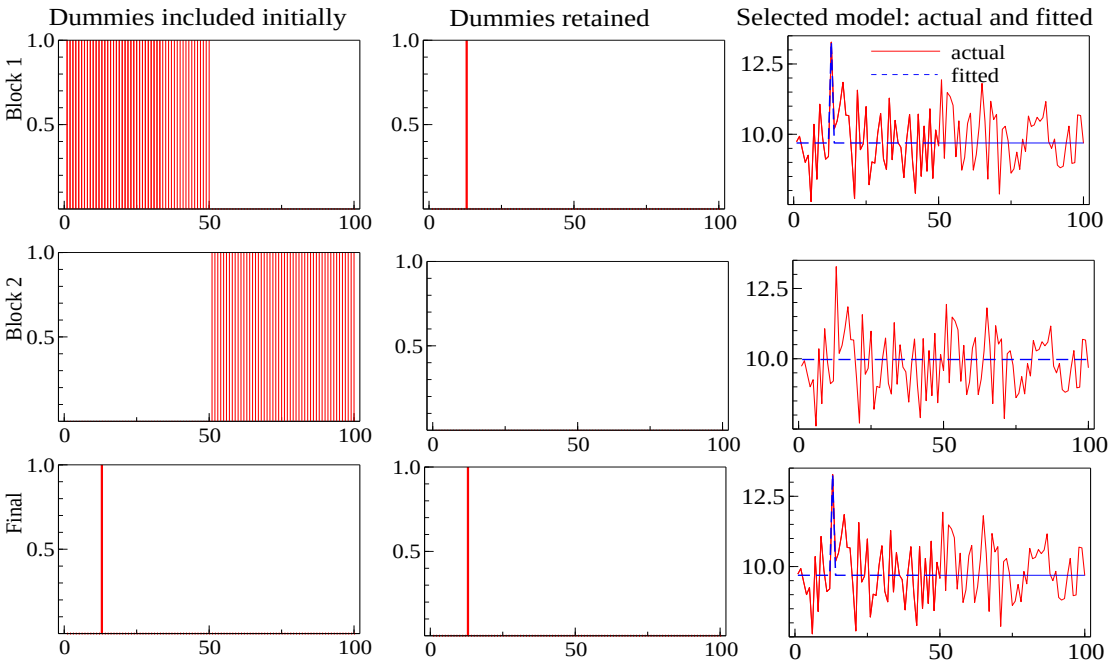


Figure 3: IIS stages for a location shift

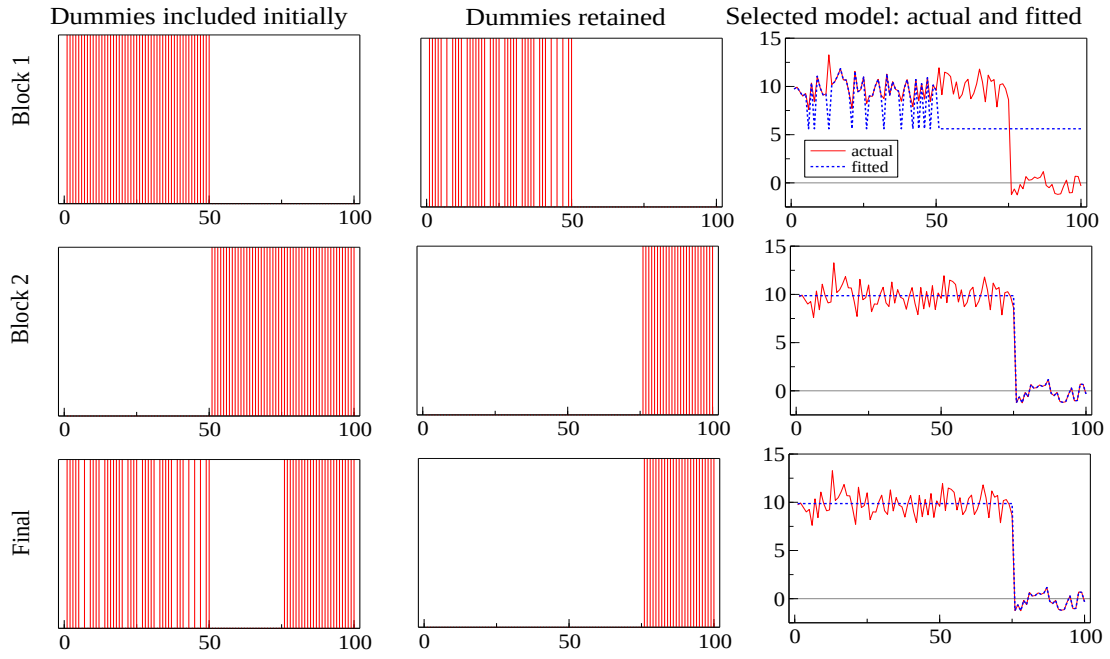
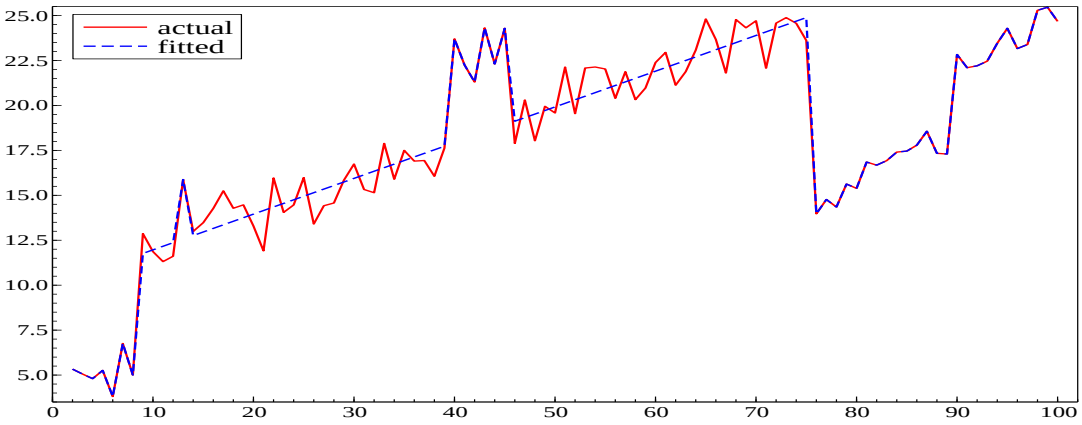


Figure 4: IIS outcomes for four location shifts



While IIS is perhaps surprising initially, many well-known statistical procedures are variants of IIS. The Chow (1960) test corresponds to sub-sample IIS over $T-k+1$ to T, but without selection, as Salkever (1976) showed, for testing parameter constancy using indicators. Recursive estimation is equivalent to using IIS over the future sample, and reducing the indicators one at a time. Johansen and Nielsen (2009) relate IIS to robust estimation, and show that under the null of no breaks, outliers or data contamination, the average cost of applying IIS is the loss of αT observations. Thus, at $\alpha = 0.01$, for $T = 100$ one observation is 'dummied out' by chance despite including 100 'irrelevant' impulse indicators in the search set and checking for location shifts and outliers at every data point. Retention of theory variables is feasible during selection with IIS, as is jointly selecting over the non-dummy variables, and IIS can be generalized to multiple splits of unequal size. While IIS entails more candidate variables than observations as $N+T > T$, selection is feasible as *Autometrics* undertakes expanding as well as contracting block searches (see next section). Non-linear model selection (including threshold models) is examined in Castle and Hendry (2011).

For a single location shift, Hendry and Santos (2010) show the detection power is determined by the magnitude of the shift τ; the length of the break interval $T - T^*$, which determines how many indicators need to be found; the error variance of the equation σ_η^2; and the significance level, α, where a normal-distribution critical value c_α, is used by the IIS selection algorithm. Castle, Doornik and Hendry (2011c) establish the ability of IIS to detect multiple location shifts and outliers, including breaks close to the start and end of the sample, as well as correcting for non-normality. Figure 4 shows the application of the general autometrics algorithm to a trending process with four breaks of varying magnitudes over $1,\ldots,10$; $40,\ldots,45$; $75,\ldots,90$; and $90,\ldots,100$, to illustrate the ability of IIS to capture multiple breaks, at both the start and end of the sample.

General case of $N > T$

The idea of generalizing using more variables than observations from IIS to all forms of independent variables is introduced by Hendry and Krolzig (2005) as well as Hendry and Johansen (2011). Suppose there are $N = \sum_{j=1}^{N} n_j$ total regressors such that $N > T$ and $n_j < T$ for all j. Consequently the total number of variables N exceeds the number of observations T but total

variables can be partitioned into blocks n_j each smaller than T. Their approach suggests randomly partitioning the set of variables into blocks of n_j, apply *Gets* to each block retaining the selected variables and crossing the groups to mix variables. The next step is to use the union of selected variables from each block to form a new initial model and repeat the process until the final union of selected variables is sufficiently small. *Autometrics* implements a variant of this algorithm to handle the general case of $N > T$.

Autometrics

Autometrics (see Doornik, 2009) is the latest installment in the automated *Gets* methodology and is available in the *OxMetrics* software package. The algorithm is based on the following main components:

1. GUM: The general unrestricted model (GUM) is the starting point of the search. The GUM should be specified based on broad theoretical considerations to nest the LGDP.

2. Pre-Search: prior to specific selection, a pre-search lag reduction is implemented to remove insignificant lags, speeding up selection procedures and reducing the fraction of irrelevant variables selected (denoted the gauge of the selection process). Pre-search is only applied if the number of variables does not exceed the number of observations ($N < T$)

3. Search Paths: *Autometrics* uses a tree search to explore paths. Starting from the GUM, *Autometrics* removes the least significant variable as determined by the lowest absolute t-ratio. Each removal constitutes one branch of the tree. For every reduction, there is a unique subtree which is then followed; each removal is back-tested against the initial GUM using an F-test. If back-testing fails, no sub-nodes of this branch are considered (though different variants of this removal exist). Branches are followed until no further variable can be removed at the pre-specified level of significance α. If no further variable can be removed, the model is considered to be terminal.

4. Diagnostic Testing: each terminal model is subjected to a range of diagnostic tests based on a separately chosen level of significance. These tests include tests for normality (based on skewness and kurtosis), heteroskedasticity (for constant variance using squares), the Chow test (for parameter constancy in different samples), and residual autocorrelation and autoregressive conditional heteroskedasticity. Parsimonious encompassing of the feasible general model by sub-models both ensures no significant loss of information during reductions, and maintains the null retention frequency of *Autometrics* close to α: see Doornik (2008). Both congruence and encompassing are checked by *Autometrics* when each terminal model is reached after path searches, and it backtracks to find a valid less reduced earlier model on that path if any test fails. This repeated re-use of the original mis-specification tests as diagnostic checks on the validity of reductions does not affect their distributions (see Hendry and Krolzig, 2003).

5. Tiebreaker: as a result of the tree search, multiple valid terminal models can be found. The union of these terminal models is referred to as the terminal GUM. As a tiebreaker to select a unique model, the likelihood-based Schwarz (1978) information criterion (SIC) is used, though other methods are also applicable, and terminal models should be considered individually.

In simulation experiments, models are primarily evaluated based on three concepts: gauge, potency and the magnitudes of the estimated parameters' root mean-square errors (RMSEs) around the DGP values (see Doornik and Hendry, 2009). Gauge describes the retention of irrelevant variables when selecting (i.e., variables that are selected but do not feature in the DGP). Potency measures the average retention frequency of relevant variables (variables that are selected and feature in the DGP). Low gauge (close to zero) and high potency (close to 1) are preferred, as are small RMSEs.

The main calibration decision in the search algorithm is the choice of significance level α at which selection occurs. Selection continues until retained variables are significant at α, though it can be the case that variables in the final model are also retained at a level above α if removal leads to diagnostic tests failing. α is approximately equal to the gauge of selection. Further, the choice of diagnostic tests and lag length selection for residual autocorrelation and autoregressive conditional heteroskedasticity need to be set.

In the general case of $N > T$ and IIS, *Autometrics* groups variables into two categories: selected and not selected (Doornik, 2010). Not currently selected variables are split into sub-blocks and the algorithm proceeds by alternating between two steps: first, the expansion step, selection is run over not-selected sub-blocks to detect omitted variables. Second, the reduction step, a new selected set is found by running selection on the system augmented with the omitted variables found in step one. This is repeated until the dimensions of the terminal model are small enough and the algorithm converges, so the final model is unchanged by further searches for omitted variables.

Autometrics has been applied successfully in a range of fields: see, for example, Hendry and Mizon (2011) on US food expenditure, Bardsen, Hurn and McHugh (2010) on unemployment in Australia, and Castle, Doornik and Hendry (2011a) for a comparison with other selection methods.

Nevertheless overall selections should be interpreted carefully. Successful identification of the underlying LGDP can be adversely affected by collinearity of the independent variables. Most simulations of *Autometrics* with large numbers of variables use orthogonal regressors, which makes selection easier. Furthermore, when $N > T$, in the block selection algorithm of *Autometrics*, adding or dropping a variable from the initial GUM may change the block partitioning of variables, so the selection is not invariant to the initial specification.

The next section covers the data used to construct the GUM in an attempt to nest the LGDP for atmospheric carbon dioxide. In the section following, *Autometrics* is then used to determine the anthropogenic contributions to CO_2.

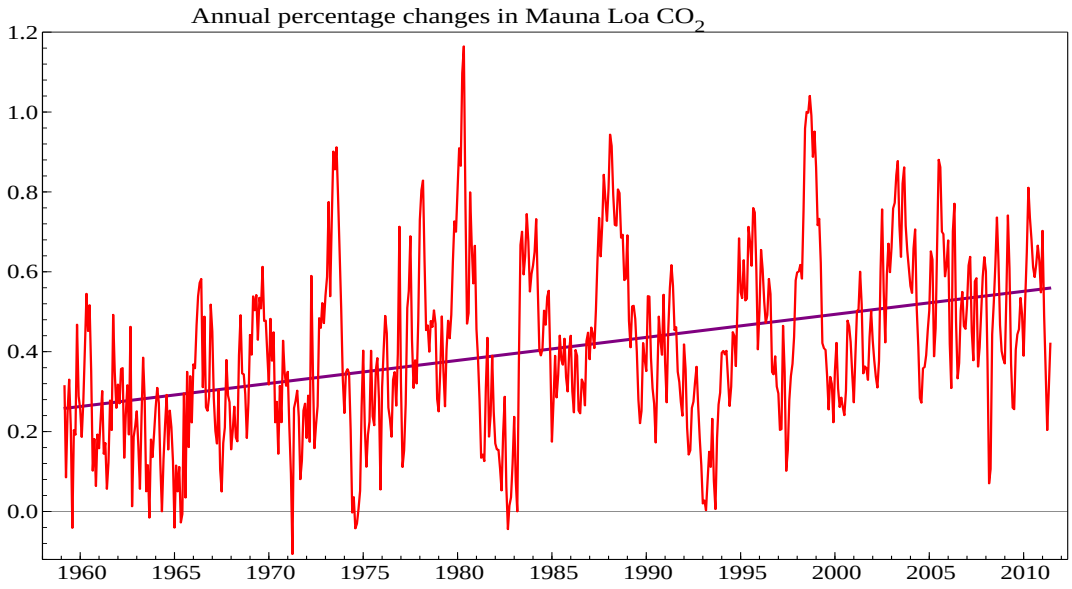
4 Data

CO_2

The atmospheric CO_2 data used here is taken from Keeling's measurements at Mauna Loa, available from Tans and Keeling (2011) (Scripps Institution of Oceanography). The time series of carbon dioxide in monthly averages runs from 1958:3 until 2011:7 at the time of writing. Simple inspection of the data shows that both the level (see figure 1) and the annual change (see figure 5) are increasing over time

The seasonal fluctuations are apparent in the data and as Buermann *et al.* (2007) poetically describe it, the regular seasonal cycle of CO_2 at Mauna Loa "records the breathing of the Northern

Figure 5: Annual changes in atmospheric CO_2



Hemisphere biosphere". In economic terms the level of atmospheric CO_2 can plainly be described as a stock variable. The total level of carbon dioxide can be approximated by the integral of the netflow to the atmosphere.

$$CO2_t = \int \text{Netflow}\, dt \tag{5}$$

where Netflow = Carbon Sources–Carbon Sinks. Therefore any analysis of the impact of anthropogenic emissions should be based on the change in atmospheric carbon dioxide:

$$\frac{dCO_2}{dt} = \text{Netflow} \tag{6}$$

Approximating this relationship from continous to discrete time:

$$\frac{dCO_2}{dt} \approx \Delta CO_{2,t} = CO_{2,t} - CO_{2,t-1} = \text{Netflow}_t \tag{7}$$

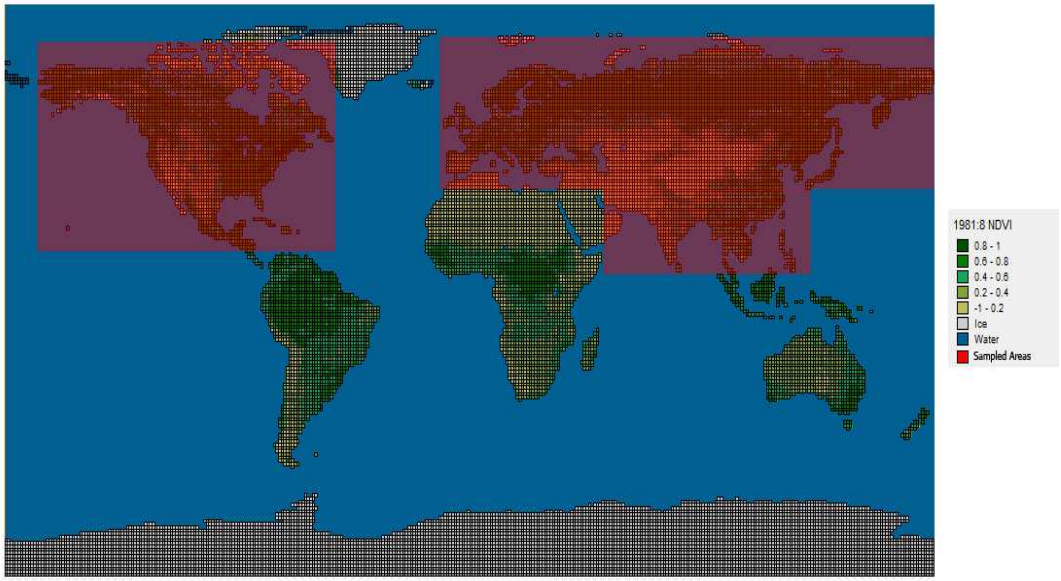
The dependent variable modeled is therefore $\Delta CO_{2,t}$. Equations (5)–(7) suggest that any relationship between the change in atmospheric carbon dioxide and netflow should be modeled in levels rather than any non-linear transformation thereof. The following section identifies variables that make up the netflow, both anthropogenic as well as natural sources and sinks.

Terrestrial biosphere

We use NDVI data to account for vegetation effects on carbon dioxide. Data are available for the NDVI from the Oak Ridge National Laboratory Distributed Archive Center (see Tucker *et al.*, 2010) ranging from 1981:7 until 2002:12 at spatial resolutions of 0.25, 0.5 and 1.0 degrees latitude and longitude.

CO_2 measured at Mauna Loa is driven by North American airflow during the Summer and airflow from Eurasia during the Winter. Therefore, the NDVI data is split into two main regions:

Figure 6: NDVI at 1° and sampled regions for 1981:8

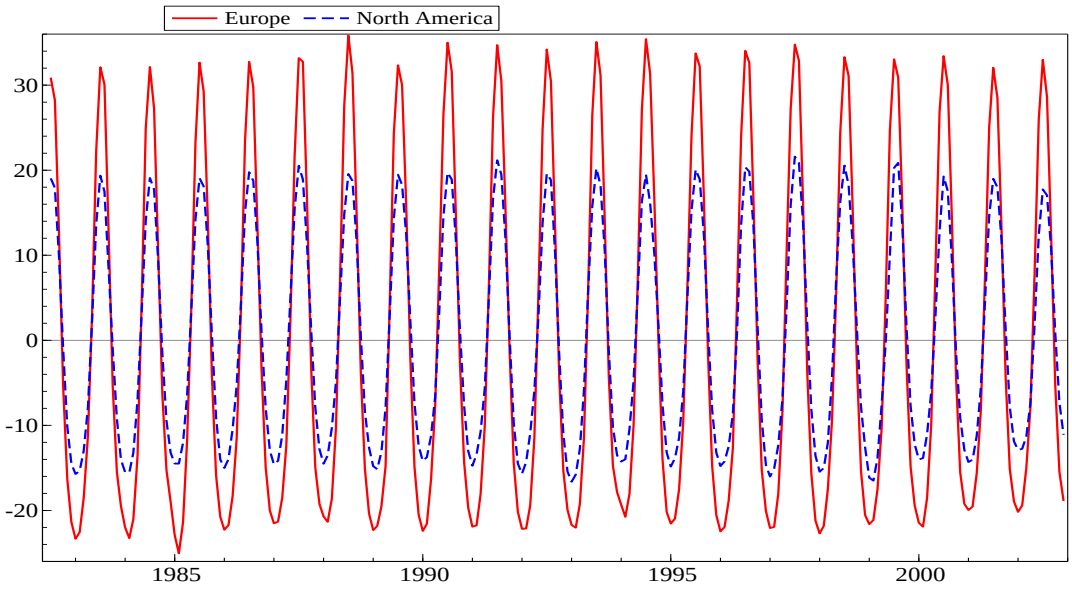


North America and Eurasia. Using 1° spatial resolution an algorithm then takes the average of every 3×3 observation grid on land within the two regions (excluding water, permanent ice and missing observations). To capture the main variation of vegetation, similar to the spatial analysis in Buermann *et al.* (2007), the North American region is defined by the rectangle ranging from 86°N/167°W to 14°N/48°W, and the Eurasian region by the rectangles ranging from 75°N/9°E to 36°N/51°E and 76°N/52°E to 7°N/358°E. Figure 6 shows the NDVI values for 1981:8 at 1° resolution with the regions defined by red rectangles. This generates 198 time series variables for North America and 567 for Eurasia.

Due to the nature of a common growing season in the Northern Hemisphere, the generated time series are highly collinear. Principal components (PCs) are used to reduce the number of variables, while retaining most of the variation in the data. Since this process captures the overall variation in vegetation, it should also reduce the problem of random noise, due to cloud cover at the time of satellite measurement. Although PCs are just linear transformations of the original times series, they have two potential advantages. First, PCs are mutually orthogonal, so adding or eliminating any one PC has little effect on the coefficient estimates of others, making the model more robust. Secondly, linear combinations of 'small' effects can be statistically significant (so retained during model selection) when individual time series would not be: see Castle, Doornik and Hendry (2011b) for a more detailed discussion. The contributions of individuals variables can be disentangled if needed.

For the following analysis, the first three principal components are entered for both North America and Eurasia. Cumulatively they explain 93.6 percent of the variation in North America and 88.5 percent of Eurasian variation. Table 1 summarizes the principal components that account for variation in the biosphere, and figure 7 shows the highly seasonal variation present in the biosphere as measured by PCs. The first principal component shows a higher amplitude in Eurasia compared to North America but the seasonal pattern is nearly identical. NDVI implicitly covers

Figure 7: First principal components for Eurasian and North American NDVI

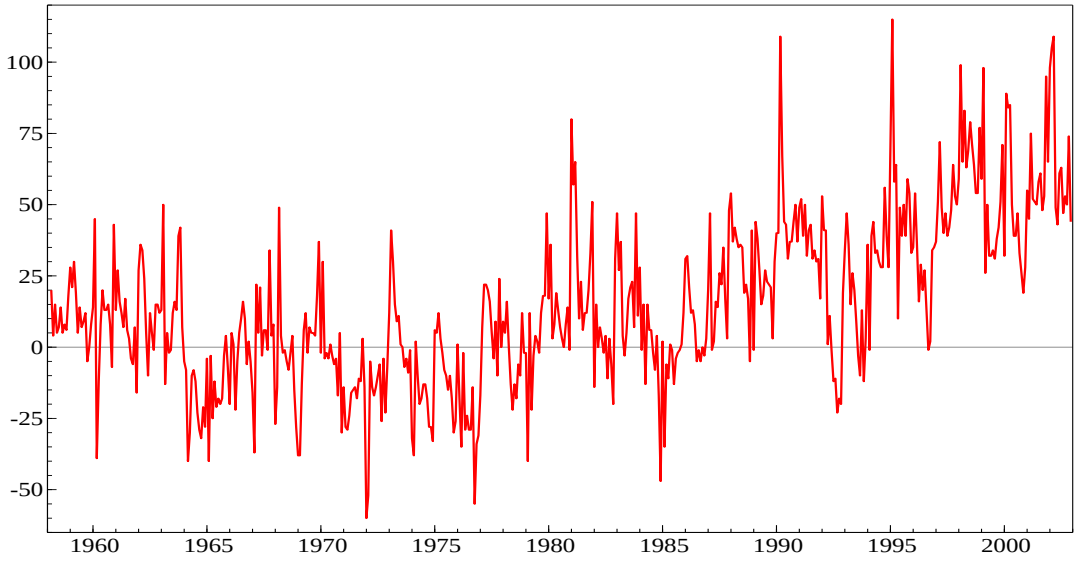


changes in land-use since it is a measure of photosynthetic activity for a particular area. A forest that is cut down would result in a change of NDVI from around $+1$ to closer to zero for that particular region. However, once NDVI is calculated for large regions and reduced in dimensionality (by PCs), changes in land use would have to occur on a grand scale to be identified in the time series. The principal components of NDVI should, therefore, be interpreted primarily as the variation in plant activity of photosynthesis and respiratory release.

Table 1: Principal components for vegetation, 1981:7–2002:12

North America	Proportion of variance	Cumulative
Principal Component 1	0.831	
Principal Component 2	0.064	0.895
Principal Component 3	0.042	0.937
Eurasia		
Principal Component 1	0.735	
Principal Component 2	0.108	0.843
Principal Component 3	0.042	0.885

Figure 8: Land and Sea temperature anomalies for the Northern Hemisphere



Oceanic indicators: Temperature and Southern Oscillation

The measure of temperature used here is the anomaly in land and sea surface temperature for the Northern Hemisphere. The temperature anomaly measure is expected to capture the main factors of ocean CO_2 absorption and is available from the NASA Goddard Institute for Space Studies (GISS) (2011) Surface Temperature Analysis from 1880:1–2011:9. The data is measured as an index in 0.01 degrees Celsius of deviations from the 1951–1980 base period (see Hansen and Lebedeff, 1987, and Hansen, Ruedy, Sato and Lo, 2010 for the detailed measurement methodology). Land measures are taken from multiple stations and are combined and corrected for urban and other non-climatic factors. Sea surface temperature measures are restricted to ice-free regions. As Hansen *et al.*, 2010 describe, temperature in the Northern Hemisphere has been increasing despite recent El Niño effects. Figure 8 shows temperature anomalies from 1958:3–2010:9.

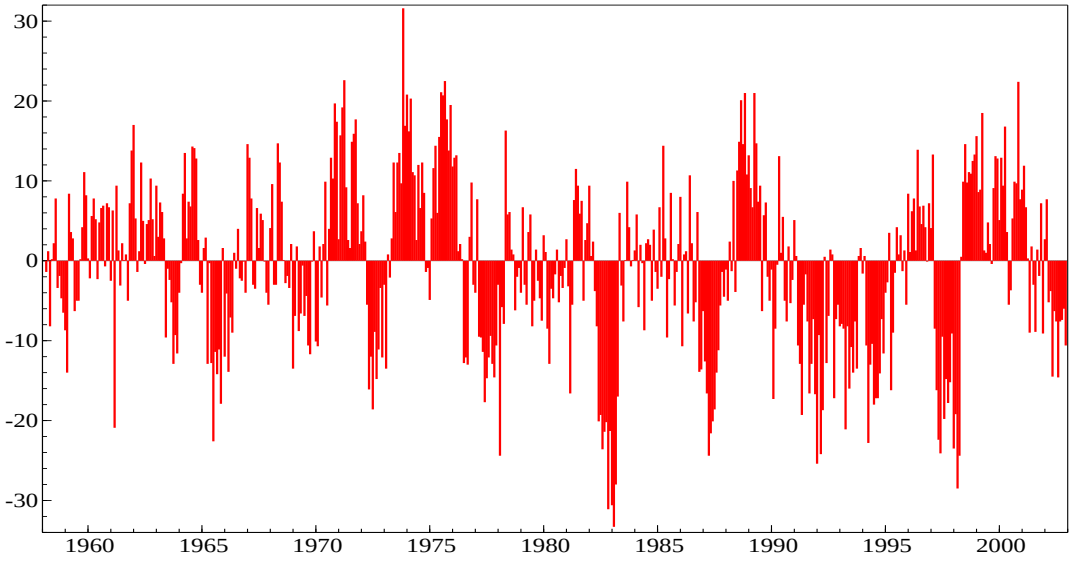
The feedback effect of CO_2 is one of the main concerns in climate change. The level of atmospheric CO_2 feeds to temperature which, in turn, affects the rate of growth of CO_2 particularly through oceanic absorption. There would be a potential problem of endogeneity if the level of CO_2 was modeled by temperature. However, lagged temperature measures should be predetermined for the growth of carbon dioxide.

To take account of weather phenomena through the Southern Oscillation we include the Southern Oscillation index (SOI). Data on the SOI is available from the Australian Bureau of Meteorology, (2011) from 1876:1 until 2011:9. Figure 4 shows the SOI for 1958 until 2011, with a noticeably strong episode of El Niño in 1997–1998 ($SOI < 0$).

Economic indicators

Anthropogenic contributions to atmospheric CO_2 are normally approximated by low-frequency economic indicators; for example annual GDP, population, an estimate of total CO_2 emissions or cement production (see section 2). This works reasonably well when trying to explain the long-run

Figure 9: Southern Oscillation index



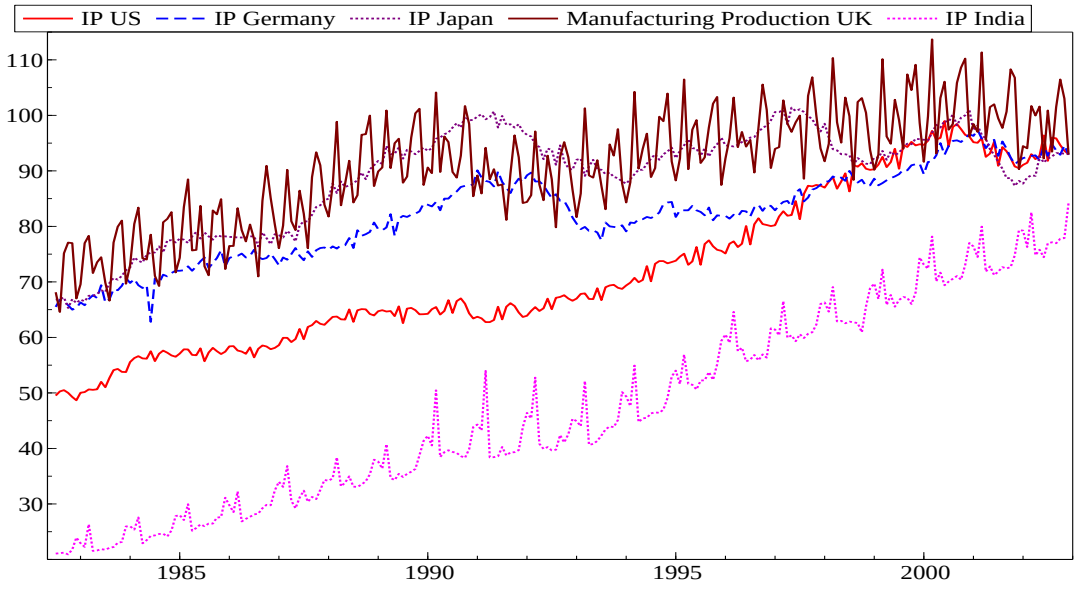
dynamics of carbon. However, using only low-frequency annual measures does not allow for estimation of any effect of anthropogenic emissions on the seasonal variation of CO_2. High-frequency (monthly) measures of anthropogenic output permit a richer analysis. Here we use a combination of multiple low-frequency (annual) and high-frequency (monthly) indicators. The annual data is included to provide a robustness check and potentially account for long-run growth. Monthly measures are included for short-run dynamics, which could explain the seasonal fluctuations as well as long-run growth. To capture atmospheric transport, variables are chosen to reflect North America as well as Europe/Asia.

High-frequency (monthly) measures

The main high-frequency indicators for anthropogenic contribution to CO_2 used here are monthly industrial production indices for multiple regions. Industrial production for North America is given by the US Industrial Production Index (2005=100) available from the Federal Reserve, (2011). The data are not seasonally adjusted and ranges from 1919:1 until 2011:5. The index measures real output in the sectors covering manufacturing, electric and gas utilities and mining and thus accounts for a large share in business-cycle fluctuations. To cover Europe and Asia, industrial production indices for the United Kingdom (UK), Germany, India and Japan are included. These measures function as a proxy for business-cycle fluctuations in the Eurasian region. Ideally Chinese and Russian production should also be included, however there are no data available on industrial production for both countries at the required frequency and time span.[2] UK industrial production is measured as an index (2005=100) of non-seasonally adjusted manufacturing activities. The data are obtained from the Office of National Statistics, (2011) and ranges from 1968:1–2011:6. German and Japanese industrial production is measured by the industrial production index (2005=100) cover-

[2]Chinese gross industrial output is only available from 1979–1999 with a change in measure thereafter. Inclusion of this series did not improve selection at the cost of a reduction in observations.

Figure 10: Industrial Production indices, 2005=100



ing manufacturing, mining and electricity, gas and water supply. The series are available only in seasonally-adjusted format from the OECD (2011) from 1960:1 until 2011:2, where for Germany after October 1990 the data account for the accession of the German Democratic Republic to West Germany. The Indian industrial production index (2005=100) covers manufacturing, mining and electricity (Government of India, Ministry of Statistics, (2011)) and ranges from 1981:4–2011:5. Table 2 and figure 10 summarize the high-frequency measures.

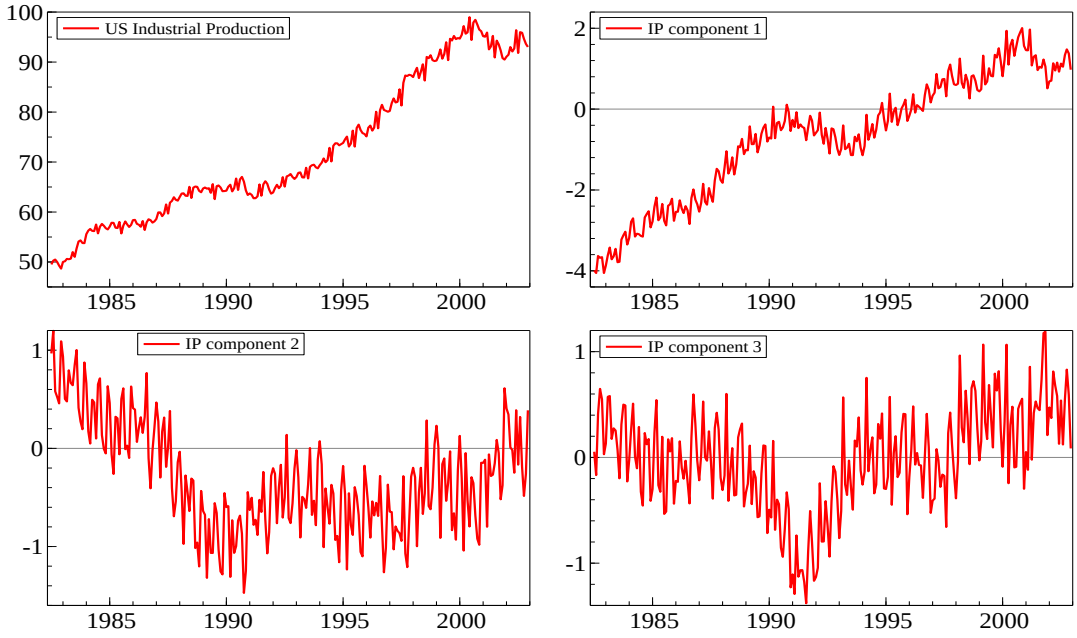
Table 2: High-frequency (monthly) variables

Measure	Description	Range	Source
US Industrial Production	Index 2005=100	1958:3–2011:5	US Federal Reserve
UK Industrial Production	Index 2005=100	1968:1–2011:6	ONS
Germany Industrial Production	Index 2005=100	1960:1–2011:2	OECD
India Industrial Production	Index 2005=100	1981:4–2011:5	Govt. of India
Japan Industrial Production	Index 2005=100	1960:1–2011:2	OECD

The seasonal adjustment of German and Japanese industrial production is visible in their dampened seasonal cycles. The higher seasonal variation in UK industrial production likely stems from it covering primarily manufacturing, which is more volatile to business cycles and seasonal factors than mining and energy production included in the other indices.

These industrial production series are highly collinear, so we again employ principal components to reduce the dimensionality and work with orthogonal variables, which improves robustness in selection as discussed above. Table 3 summarizes the first three industrial production components, which cumulatively explain approximately 97% of variation in production. Figure (11)

Figure 11: Industrial production principal components



compares US industrial production to the three components used in selection.

Table 3: Principal components for Industrial Production, 1981:4–2011:2

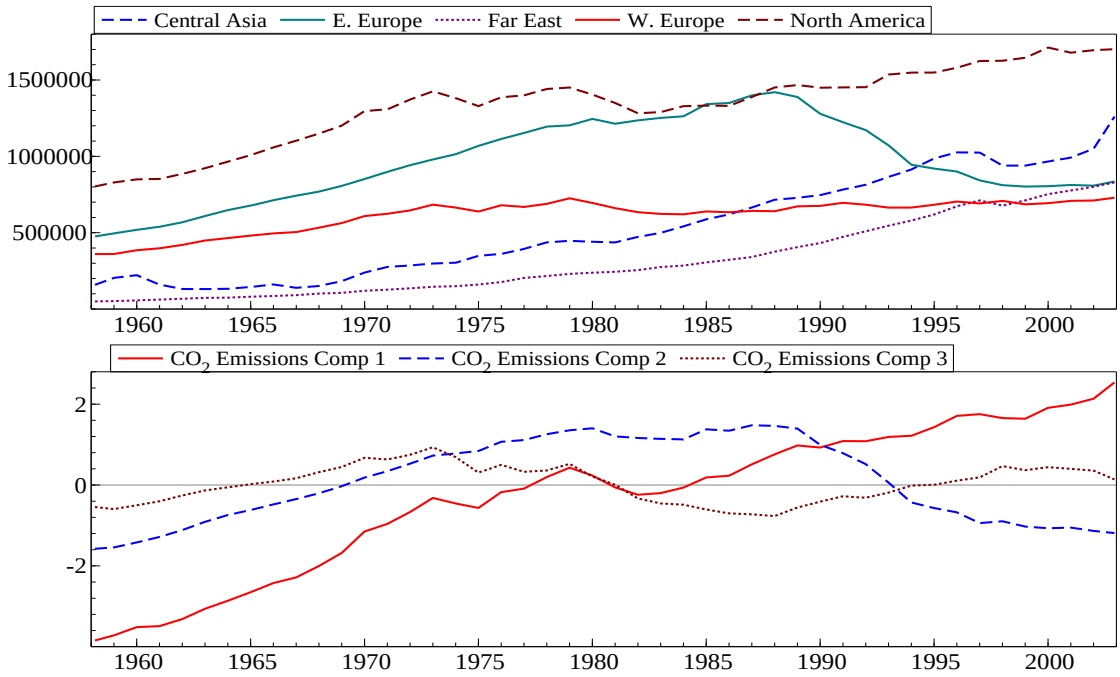
IP	Proportion of variance	Cumulative
Principal Component 1	0.809	
Principal Component 2	0.119	0.928
Principal Component 3	0.047	0.975

These are the first 3 anthropogenic high-frequency components included in the GUM. While industrial production reflects the intensity of economic activity associated with higher emissions, it does not account for changing emission intensity. More efficient processes could lead to an increase in industrial production without an equivalent increase in emissions. This is a crucial missing measure and difficult to control for: an attempt is made by including overall long-run emissions estimates in addition to production.

Low-frequency (annual) measures

Low-frequency anthropogenic measures are variables reported on an annual basis, and capture the long term of human-driven CO_2 emissions. These low-frequency variables include total estimated CO_2 emissions (in thousand metric tons of carbon) for North America, Western Europe, Eastern

Figure 12: CO_2 emissions and principal components



Europe, Central Asia and the Far East.[3] Emissions are estimated based on the burning of fossil fuels, cement manufacture and gas flaring by the US Department of Energy (see Marland *et al.*, 2011) from 1950–2007. Table 4 summarizes the annual measures.

Table 4: Low-frequency (annual) variables

Measure	**Description**	**Range**	**Source**
N. America CO_2 Emissions	000s tons carbon	1950-2007	US Dep. of Energy
W. Europe CO_2 Emissions	000s tons carbon	1950-2008	US Dep. of Energy
E. Europe CO_2 Emissions	000s tons carbon	1950-2008	US Dep. of Energy
Central Asia CO_2 Emissions	000s tons carbon	1950-2008	US Dep. of Energy
Far East CO_2 Emissions	000s tons carbon	1950-2008	US Dep. of Energy

In order to usefully combine these annual variables with the monthly measures listed above, all annual variables are linearly interpolated to monthly observations. In the case of moving from annual to monthly observations, this method estimates a straight line over 12 months between each annual observation. While this may be restrictive, if the seasonal structure of the variables is not known there is no *a priori* reason to prefer a different interpolation algorithm. In any case, seasonal dummy variables and IIS can 'pick up' any systematic or large deviations. As before, the low-frequency measures are reduced in dimensionality by calculating their principal components. The

[3]See Marland *et al.*, 2011 for a detailed listing of which countries are included. Eastern Europe includes Russia, Central Asia includes China, and India is covered by the Far East category.

first three components capture 99% of variation in the emissions series. Table 5 shows explained variation of each component and figure (12) displays interpolated CO_2 emissions as well as the three low-frequency PCs.

Table 5: Low-frequency (annual) components

CO_2 emissions	Proportion of variance	Cumulative
Principal Component 1	0.733	
Principal Component 2	0.222	0.955
Principal Component 3	0.039	0.994

The components of the interpolated annual variables are included to account for potential low-frequency movements, the industrial production short-term indicators are expected to be sufficient to capture inter-annual dynamics. These are the first three low-frequency anthropogenic components included in the GUM.

5 Estimation

Overview

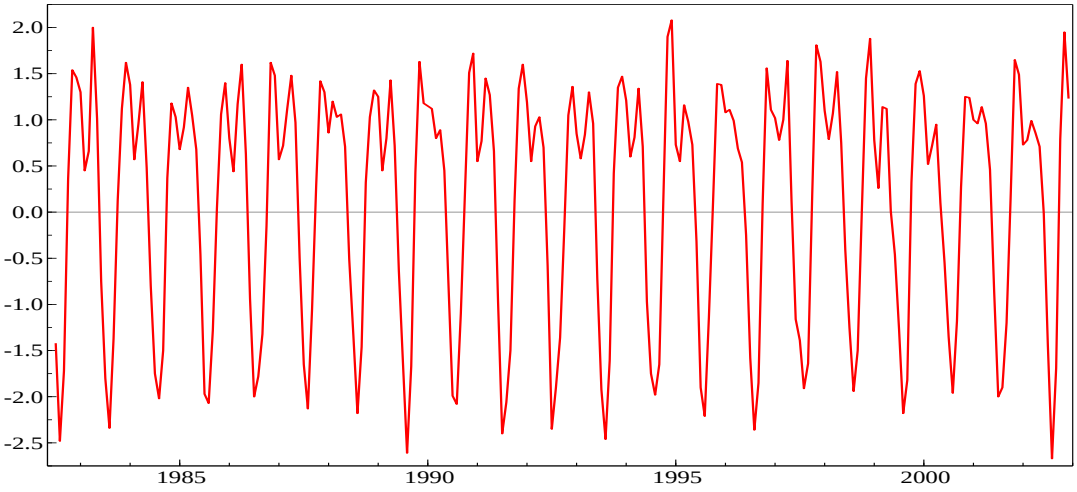
The dependent variable ΔCO_2 is modeled as a finite autoregressive-distributed lag model (ADL) (see e.g., Hendry, 1995):

$$\Delta CO_{2,t} = a(L)\Delta CO_{2,t-1} + \sum_{i=1}^{p} b_i(L)x_{i,t} + \mathbf{z}_t'\boldsymbol{\phi} + \epsilon_t \tag{8}$$

where L denotes the lag operator, and p is the number of explanatory variables x_i. The vector $\mathbf{z}$ consists of non-lagged deterministic terms such as a linear time trend, centred seasonal variables and impulse indicators. Let q denote the total number of explanatory variables that appear on the left-hand side of equation 8. Then the change in atmospheric CO_2 is modeled as a function of past values of the change in CO_2, current and past values of selected independent variables x_i, and deterministic components. Figure 13 graphs $\Delta CO_{2,t}$ for 1982:7–2002:12. The seasonal variation is so large that it is difficult to visually discern the slow but persistent growth.

There is a large number of potential explanatory variables x_i in modeling atmospheric CO_2. The model needs to account for all the above mentioned anthropogenic and natural factors as well as their lag reactions, as CO_2 is a highly autocorrelated time series. Adding IIS quickly moves the general unrestricted model to a situation with more explanatory variables than observations. This used to be a major difficulty in modeling, however, as outlined in section (3), can now be handled by estimation in blocks using *Autometrics* to select variables to retain in the final model in the form of (8). The estimation procedure operates as follows: first the theory-motivated GUM is specified, then estimation in blocks following the *Autometrics* algorithm selects down to individual terminal models. The union of terminal models is captured by the final GUM. Formally, the selection for

Figure 13: Dependent variable $\Delta CO_{2,t}$ over 1982:7–2002:12



the final model can use the likelihood-based SIC, though as each terminal model represents a valid representation, final model selection can be based on other theoretical considerations.

Formulation of the GUM

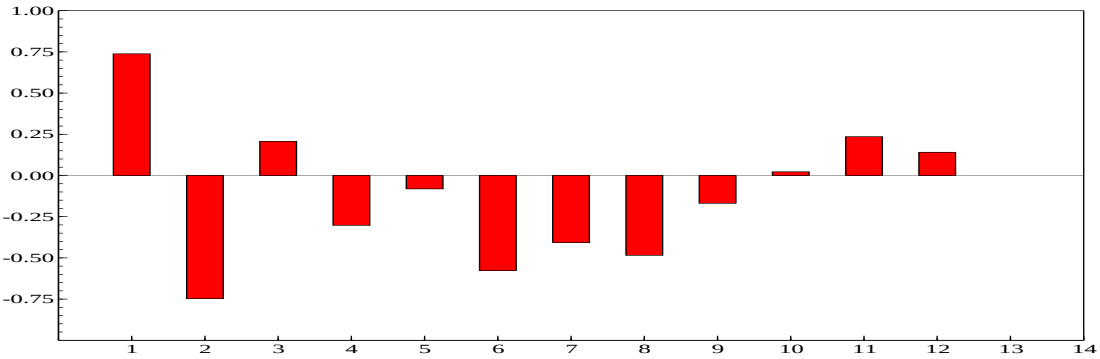
The dependent variable that is being modeled is ΔCO_2 (the change in atmospheric carbon dioxide measured at Mauna Loa). While *Autometrics* has been shown to be effective at recovering the data generating process in large models, the algorithm is not perfect and is sensitive to initial specification. As a robustness check, therefore, we estimate multiple variations of initial sample specifications. The sub-sample GUMs 1 and 2 are selected for different initial variables, though they always include the main variables of interest: short-term dynamics captured by monthly industrial production and the control variables for natural factors—temperature, SOI and vegetation. Non-anthropogenic emissions are captured through potential oceanic release of CO_2 (SOI and temperature) and the respiratory release phase in vegetation (NDVI close to zero). The models include general control variables of a constant, a linear time trend and 11 centred seasonal dummy variables (annual frequency -1 dummies with mean zero in the long run), which are subject to selection so not automatically included in terminal models.

To capture inter-seasonal transport dynamics, interaction terms for Winter/Summer are included with each vegetation measure. Summer is defined as May–October and Winter is defined as November–April. Thus, binary variables (Winter and Summer Weights) are interacted with all region-specific NDVI variables and included in the GUM. This could be extended to smoothed weights following a sine/cosine pattern, but binary weights are expected to capture the main seasonal effect of different atmospheric transport.

Sample 1 includes all variables measured at a monthly frequency: natural control variables as well as the first three components of industrial production. Sample 2 includes all short-term dynamics as in sample 1 and adds the first three low-frequency interpolated fossil fuel emission components.

Since ΔCO_2 is a highly autocorrelated series when measured monthly, a long lag length is allowed. Based on the partial-autocorrelation function (PACF) of ΔCO_2, the longest lag length

Figure 14: Partial-autocorrelation function for ΔCO_2



is selected to be 12, as longer lags fall below the 95% critical level. Figure 14 shows the PACF for ΔCO_2. Additionally, 12 lags of each independent variable are added to GUM 1. Lag lengths for GUM 2 are specified based on selection from GUM 1 to compensate for a larger number of independent variables (see section 6).

Table 6: GUM: Sample 1

Variables included	lag length
Temperature	12
NDVI PC1 Eurasia (Eur) + Winter Interaction	12
NDVI PC2 Eurasia (Eur) + Winter Interaction	12
NDVI PC3 Eurasia (Eur) + Winter Interaction	12
NDVI PC1 North America (NA) + Summer Interaction	12
NDVI PC2 North America (NA) + Summer Interaction	12
NDVI PC3 North America (NA) + Summer Interaction	12
SOI	12
Industrial Production Comp. 1	12
Industrial Production Comp. 2	12
Industrial Production Comp. 3	12
Constant	yes
Trend	yes
Centred Seasonal Variables	yes
Impulse Indicators	yes
Total variables	492

As controls for the terrestrial biosphere are included in every GUM, it is the NDVI measure that defines the maximum number of available observations. Carbon dioxide is measured at Mauna Loa from 1958:3 until present, but NDVI data is only available from 1981:7 until 2002:12, so the vegetation control limits the maximum number of observations to 258. A maximum lag length of 12 then leads to the estimated sample size being 246 observations ($T = 246$).

Impulse indicator saturation includes one binary variable for each observation. In the case of a model with 12 lags, this means that 246 individual binary variables are added to the GUM. Tables 6 and 7 provide an overview of the variables making up the sample-specific GUMs.

Table 7: GUM: Sample 2

Variables included	lag length
Temperature	6
NDVI PC1 Eurasia (Eur) + Winter Interaction	12
NDVI PC2 Eurasia (Eur) + Winter Interaction	12
NDVI PC3 Eurasia (Eur) + Winter Interaction	12
NDVI PC1 North America (NA) + Summer Interaction	12
NDVI PC2 North America (NA) + Summer Interaction	12
NDVI PC3 North America (NA) + Summer Interaction	12
SOI	6
Industrial Production Comp. 1	6
Industrial Production Comp. 2	6
Industrial Production Comp. 3	6
CO_2 Emissions Comp. 1	6
CO_2 Emissions Comp. 2	6
CO_2 Emissions Comp. 3	6
Constant	yes
Trend	yes
Centred Seasonal Variables	yes
Impulse Indicators	yes
Total variables	483

Non-stationarity

Non-stationarity is a common feature of time-series data. Broadly, a non-stationary process is defined as a process whose distribution changes over time, for example the mean or variance of the process are non-constant (see Hendry and Juselius, 2000, for a detailed discussion). The trending level of CO_2 measured at Mauna Loa is non-stationary as its mean is increasing with time. There are various forms of non-stationarity. The time series could be integrated processes (a time series is said to be integrated of order r, or I(r), if differencing the series r times yields a stationary process), or alternatively there could be structural breaks/shifts in coefficients or levels. Non-stationarity through structural breaks need not be removed by differencing. IIS is used to detect structural breaks (as well as mis-specification): if a large number of impulse indicators are selected, the model may be mis-specified, data badly mis-measured; or there are breaks in the data.

Non-stationary data is not a problem in automatic model selection so long as it is approached correctly. Selection in *Autometrics* is primarily based on t- and F-statistics that under non-stationarity can follow non-standard distributions. Sims, Stock and Watson, 1990 show that the limit distributions of these test statistics are standard if the original equation can be re-written in terms of coefficients on mean-zero stationary variables. That is to say, the actual re-parametrization is not required, merely the existence of a linear re-parametrization in mean-zero variables is sufficient for the test statistics in the original equation to have standard distributions. So long as the equation can be written that way, selection based on t- and F-statistics will be valid. However, there are potential problems during selection when a path is considered in which this re-parametrization is not possible. Such a difficulty is hard to avoid, so is handled here by using tight significance levels that allow for possible non-standard distributions. As figure 5 showed, $\Delta_{12}CO_2$ manifests a strong stochastic trend so is clearly not stationary: in monthly data, that trend is 'swamped' by the seasonal variation. Due to this strong seasonality, augmented Dickey–Fuller (Dickey and

Fuller, 1981) type tests would not successfully identify a unit root. Seasonal unit-root tests are somewhat unreliable as other determinants, such as breaks, are not included, but based on the apparent stochastic trend in $\Delta_{12}CO_2$, it is safe to say that ΔCO_2 is integrated of order one, I(1). Unit roots can also not be rejected for most of the independent time-series variables that are included as potential determinants. The order of integration of the dependent variable is then the same as the order of integration of the independent variables. Given that the model can be written as coefficients on stationary mean-zero variables (see Banerjee, Dolado, Galbraith and Hendry, 1993), we proceed by estimating the model in levels with tight significance levels to account for selection effects where an I(0) re-parametrization is not possible.

Algorithm specification

The *Autometrics* algorithm is used to estimate and select within the GUMs described in tables 6 and 7. The algorithm is calibrated to the following parameters. Selection is done at a 0.1% significance level, considerably tighter than the conventional 5% or 1% used in the literature. In models starting with K irrelevant variables, this implies that on average $0.001K$ irrelevant variables are retained in the terminal model. The current model is reduced by removing the least significant variable until variables cannot be dropped at the 0.1% level. At the termination of each path, models are backtested to the initially specified GUM when feasible. Diagnostic tests, defined formally below (see Doornik and Hendry, 2009), are conducted at a 1% level, for normality, heteroskedasticity, coefficient constancy (set to a 70% sample split), residual autocorrelation and autoregressive conditional heteroskedasticity, both based on 8 lags. The specified GUM includes more variables than observations with IIS, so diagnostic tests are only applied to terminal models, and if satisfactory, conventional standard errors are used.

6 Results

Overview

The specifications given in tables 6 and 7 are selected from general unrestricted models using *Autometrics* by the algorithm outlined in section 5. Sample 1 includes all natural controls and monthly anthropogenic components; sample 2 includes all natural controls as well as monthly and annual anthropogenic components. The crucial feature of *Autometrics* is determining the selection of variables, rather than their estimated coefficients, although bias corrections have been implemented. Below, we also note the relative importance of individual variables through decompositions of explained variance.

The selection algorithm results in 14 terminal models for sample 1, and 16 terminal models for sample 2. It may surprise that so many congruent undominated different representations can be found at such a tight significance level as 0.1%, but this merely reflects the high collinearity both between the different series and over time as represented by their lagged values. Most of the terminal models are minor variations on others as the final GUMs had 21 and 30 variables respectively.[4] The final models are selected from the set of terminal models by the smallest SIC value. Equations (9) and (10) show the selected final models for sample 1 and sample 2 respectively.

[4]Detailed results are available from the authors on request.

$$
\begin{aligned}
\widehat{\Delta CO}_{2,t} = & \underset{(0.053)}{0.24}\ \Delta \mathrm{CO}_{2,t-1} - \underset{(0.049)}{0.67}\ \Delta \mathrm{CO}_{2,t-2} + \underset{(0.05)}{0.20}\ \Delta \mathrm{CO}_{2,t-3} \\
& - \underset{(0.037)}{0.32}\ \Delta \mathrm{CO}_{2,t-4} + \underset{(0.057)}{0.24}\ \mathrm{IP}_{1_{t-1}} - \underset{(0.056)}{0.30}\ \mathrm{IP}_{1_{t-4}} \\
& - \underset{(0.034)}{0.20}\ \mathrm{IP}_{2_{t-4}} + \underset{(0.035)}{0.15}\ \mathrm{IP}_{3_t} + \underset{(0.0005)}{0.003}\ \mathrm{Temp}_{t-4} \\
& - \underset{(0.0013)}{0.006}\ \mathrm{SOI}_{t-5} - \underset{(0.007)}{0.042}\ \mathrm{NDVI}_{1,\,\mathrm{Eur}_{t-1}} + \underset{(0.003)}{0.019}\ \mathrm{NDVI}_{1,\,\mathrm{Eur}_{t-10}} \\
& - \underset{(0.006)}{0.020}\ \mathrm{NDVI}_{1,\,\mathrm{Eur}_{t-12}} + \underset{(0.007)}{0.026}\ \mathrm{w\_NDVI}_{3,\,\mathrm{Eur}_{t-8}} \qquad (9)
\end{aligned}
$$

$$
\widehat{\sigma} = 0.212 \;\; T = 246 \;\; n = 14 \;\; \mathsf{SIC} = -0.006 \;\; \mathsf{F}_{\mathsf{ar}}(8, 224) = 1.55
$$
$$
\chi^2_{\mathsf{nd}}(2) = 1.47 \;\; \mathsf{F}_{\mathsf{reset}}(2, 230) = 1.42 \;\; \mathsf{F}_{\mathsf{arch}}(8, 230) = 1.07 \;\; \mathsf{F}_{\mathsf{het}}(28, 217) = 1.42
$$

Let $\mathsf{F}_{\mathsf{name}}$ denote an approximate Lagrange-multiplier F-test, then F_{ar} tests for autocorrelation of order k (Godfrey, 1978), $\mathsf{F}_{\mathsf{het}}$ tests for heteroscedasticity (White, 1980), $\mathsf{F}_{\mathsf{arch}}$ tests for k^{th}-order autoregressive conditional heteroskedasticity (ARCH: Engle, 1982), $\mathsf{F}_{\mathsf{reset}}$ tests for functional-form mis-specification (White, 1980), and $\chi^2_{\mathsf{nd}}(2)$ tests for non-normality (Doornik and Hansen, 2008).

$$
\begin{aligned}
\widehat{\Delta CO}_{2,t} = & \underset{(0.051)}{-0.57}\ \Delta \mathrm{CO}_{2,t-2} - \underset{(0.050)}{0.22}\ \Delta \mathrm{CO}_{2,t-4} + \underset{(0.062)}{0.27}\ \mathrm{IP}_{1_{t-1}} \\
& - \underset{(0.063)}{0.34}\ \mathrm{IP}_{1_{t-4}} - \underset{(0.039)}{0.28}\ \mathrm{IP}_{2_{t-4}} + \underset{(0.035)}{0.16}\ \mathrm{IP}_{3_t} - \underset{(0.21)}{0.74}\ \mathrm{I}_{1990(7)} \\
& + \underset{(0.0006)}{0.004}\ \mathrm{Temp}_{t-4} - \underset{(0.001)}{0.007}\ \mathrm{SOI}_{t-5} - \underset{(0.008)}{0.047}\ \mathrm{NDVI}_{1,\,\mathrm{Eur}_t} \\
& - \underset{(0.006)}{0.044}\ \mathrm{NDVI}_{1,\,\mathrm{Eur}_{t-3}} + \underset{(0.006)}{0.033}\ \mathrm{NDVI}_{1,\,\mathrm{Eur}_{t-11}} - \underset{(0.009)}{0.042}\ \mathrm{NDVI}_{1,\,\mathrm{Eur}_{t-12}} \\
& + \underset{(0.008)}{0.030}\ \mathrm{NDVI}_{2,\,\mathrm{Eur}_{t-11}} - \underset{(0.007)}{0.029}\ \mathrm{w\_NDVI}_{2,\,\mathrm{Eur}_{t-7}} - \underset{(0.006)}{0.022}\ \mathrm{w\_NDVI}_{2,\,\mathrm{Eur}_{t-8}} \\
& + \underset{(0.007)}{0.035}\ \mathrm{w\_NDVI}_{3,\,\mathrm{Eur}_{t-8}} - \underset{(0.008)}{0.039}\ \mathrm{s\_NDVI}_{1,\,\mathrm{NA}_{t-8}} + \underset{(0.007)}{0.034}\ \mathrm{s\_NDVI}_{1,\,\mathrm{NA}_{t-9}} \qquad (10)
\end{aligned}
$$

$$
\widehat{\sigma} = 0.199 \;\; T = 246 \;\; n = 19 \;\; \mathsf{SIC} = -0.048 \;\; \mathsf{F}_{\mathsf{ar}}(8, 219) = 1.62
$$
$$
\chi^2_{\mathsf{nd}}(2) = 0.044 \;\; \mathsf{F}_{\mathsf{reset}}(2, 225) = 0.042 \;\; \mathsf{F}_{\mathsf{arch}}(8, 230) = 0.57 \;\; \mathsf{F}_{\mathsf{het}}(36, 208) = 1.03
$$

First: as is to be expected from the theory, controls for natural factors are selected in both final models. Temperature anomalies enter the model with a positive coefficient, likely capturing the effect of oceanic uptake such that CO_2 increases with higher temperatures. Vegetation controls through the principal components of NDVI are selected in both models, as is the control for Southern Oscillation. However, a key finding is that in both terminal models, natural controls are insufficient to account for the variation in the change of atmospheric CO_2. Anthropogenic factors captured through components of industrial output indices are consistently selected in both models. Selection of these components is highly consistent over the two models as the selected production

components are identical in models 1 and 2. Selection of these is robust to the addition of emissions components which are not selected in sample 2, suggesting that the high frequency measures provide a better approximation for anthropogenic emissions measured at Mauna Loa.

Second: most selected variables enter the model in lagged form. Only the third principal component of production (in sample 1 and 2) and the first component of Eurasian NDVI (in sample 2) have an estimated immediate effect on the growth of CO_2. Most anthropogenic emissions and vegetation growth "lead" measured atmospheric CO_2 by suggested time periods of 1 to 12 months. Relative to the initial sizes of the GUMs, few variables are retained, yet relative to the tight significance levels, many more are retained than could be attributed to chance (less than 1 on average).

Third: the final models appear to be well specified. As a result of the algorithm, all models pass tests for normality, heteroskedasticity, residual autocorrelation and autoregressive conditional heteroskedasticity. The number of selected indicators from IIS is low. There is only one indicator selected for 1990:7. This suggest that the model is correctly specified and there appear to be no real structural breaks or shifts in the change of atmospheric CO_2. No deterministic variables are selected: no constant, time trend or centred seasonals appear in the final models. This suggests that changes in CO_2 are well approximated by the selected final variables covering anthropogenic and natural factors. If constants are included post-selection (both not statistically different from zero in models 1 and 2), R^2 can be used as a rough measure of goodness of fit. Both final models exhibit a high goodness of fit ($R^2 = 0.974$ and $= 0.978$ respectively). This is not a straight result of selection as *Autometrics* does not directly maximize the goodness of fit. Moreover, R^2 measures should not be attributed much weight when assessing models, as there are preferred likelihood-based measures that also account for the number of parameters included.

GUM: Sample 1

Sample 1 in equation (9) covers all variables measured at a monthly frequency. *Autometrics* in sample 1 with 246 observations estimated 345 models reducing the number of explanatory variables from an initial 492 (split into initial 6 blocks) down to 14 in the final model. The final model passes the stationarity test on the residual, unit roots ranging from lags 1 to 12 are rejected at the 1% level using an ADF test. There are no impulse indicator variables selected in the final GUM. Together this provides evidence for a well specified model that forms a stationary relationship.

Selection: neither the constant, the linear time trend nor centred seasonal variables feature in the final model, so that the seasonality and increase in the growth of CO_2 are explained by the anthropogenic and natural factors. All selected variables (apart from IP_{3_t}) enter the model as lags, suggesting a delay in the effect of CO_2 emissions/absorption and measurement at Mauna Loa. The longest lag on an anthropogenic component is 4 months.

The anthropogenic sources that are selected are all three principal components for industrial production at lag lengths ranging from immediate t to $t-4$. Component 1 is selected at $t-1$ and $t-4$ with opposite signs, suggesting that it enters the model mainly as a difference. As these variables are principal components of production indices, the coefficients are not straight-forward to interpret. The key result is the consistent selection, relative importance and lag length of these, rather than the signs of individual estimated coefficients.

The non-seasonally weighted principal components for NDVI are mostly selected, but only the Eurasian region is included. Given that the growth cycle is relatively similar in North America and Eurasia this should not be over-interpreted. It is likely that the PCs for North American and

Eurasian vegetation capture a very similar pattern and are to a considerable extent interchangeable. The negative coefficient on the first PC of vegetation (at $t-1$, as well as in the long-run solution below) supports the theory that increased vegetation activity slows down CO_2 growth. The near equal magnitude, opposite signs on $t-10$ and $t-12$ suggest a difference, a pattern also seen in (10). The coefficient on temperature is positive at a lag of $(t-4)$, likely capturing reduced oceanic absorption under higher temperature. Southern Oscillation enters the model at a lag of five months with a negative coefficient. Thus, during El Niño episodes ($SOI < 0$) growth in CO_2 appears to increase, in line with findings of other papers (see Bacastow, 1976, Keeling and Revelle, 1985).

To quantify and assess the relative importance of each regressor, we decompose the total explained variation in ΔCO_2. Decomposition is not straight forward when independent variables are correlated. We use two measures, partial R^2 and hierarchical partitioning. The partial R^2 provides a measure of the marginal contribution to explained variation for a given variable, while holding other factors constant. In hierarchical partitioning the explained variance is decomposed by calculating the average contribution of each variable over all potential orderings of the variables (see Kruskal, 1987, and method LMG in Groemping, 2007).[5] Individual variance contributions sum to unity and can be interpreted as percentages. This yields values for individual variables that sum to the total explained variance (R^2). Table 8 ranks the selected variables by the partial R^2 and also reports relative importance based on hierarchical partitioning.

Table 8: GUM: Sample 1 relative importance

Variable	**Partial R^2**	**hierarchical partitioning**
$\Delta CO_{2,t-2}$	0.4472	0.0661
$\Delta CO_{2,t-4}$	0.2461	0.0643
$NDVI_{1,\,Eur_{t-10}}$	0.1583	0.0257
$IP_{2_{t-4}}$	0.1259	0.0036
$NDVI_{1,\,Eur_{t-1}}$	0.1259	0.2795
$Temp_{t-4}$	0.1225	0.0067
$IP_{1_{t-4}}$	0.1105	0.0117
SOI_{t-5}	0.0836	0.0014
$\Delta CO_{2,t-1}$	0.0796	0.1954
IP_{3_t}	0.0756	0.0074
$IP_{1_{t-1}}$	0.0727	0.0157
$\Delta CO_{2,t-3}$	0.0619	0.0372
$w\_NDVI_{3,\,Eur_{t-8}}$	0.0578	0.0200
$NDVI_{1,\,Eur_{t-12}}$	0.0483	0.2648

Based on this decomposition, the single largest (non-autoregressive) contribution comes from the Eurasian NDVI principal component of vegetation followed by industrial production. The anthropogenic components cumulatively explain a large fraction of the variation in atmospheric CO_2. Perhaps surprisingly, both the SOI and temperature are ranked low based on hierarchical partitioning.

[5] An intercept is forced to be retained in selection of these models, though it is not statistically different from zero for model 1 and 2.

GUM: Sample 2

Sample 2 in (10) covers all variables measured at a monthly frequency as well as interpolated annual components for long-term anthropogenic emissions. Lag selection for sample 2 is based on selection in sample 1. Selection in model 1 results in a maximum lag of 4 on anthropogenic components and 5 for temperature and SOI. The longest lag for vegetation was selected at $t - 12$. Therefore, GUM 2 starts with a maximum lag of 6 for anthropogenic components, and a maximum lag of 12 for vegetation. *Autometrics* in sample 2 with 246 observations estimated 571 models reducing the number of explanatory variables from an initial 483 (split into initial 6 blocks) down to 19 in the final model. Unit roots are rejected for the model residuals at the 1% level using ADF tests covering up to 12 lags. There is only a single impulse indicator selected (1990:7), suggesting no major breaks or mis-specification.

Selection: sample 2 results in a slightly higher number of selected variables in the final model relative to sample 1. In terms of robustness, selection is highly consistent relative to sample 1. Anthropogenic components, as well as temperature and Southern Oscillation are selected identically to model 1. No deterministic terms are selected.

Anthropogenic emissions are captured solely by the principal components for industrial production: the low-frequency fossil fuel measures are not selected. These findings suggest that high-frequency industrial production provides a better measure for anthropogenic factors than interpolated fuel emissions. This may appear surprising given that anthropogenic emissions directly measure emitted carbon dioxide: however, this result likely stems from the annual frequency of carbon dioxide emissions that miss any seasonal component.

In terms of natural controls, the estimated coefficient on temperature is positive and that on Southern Oscillation is negative, both as in the previous model. The selection of vegetation variables moves towards seasonally-weighted Eurasian indicators but also adds North American NDVI measures. The selection of seasonally-weighted vegetation measures supports the theory of atmospheric transport. The shift in selection of vegetation variables is likely due to the collinearity in the NDVI measure resulting from a similar growth cycle in North America and Eurasia. The selected indicator for 1990:7 (July) suggests there was a reduction in growth relative to other years. Although the cause for this is not apparent in the data, it should be noted that such an indicator creates a step shift in the level of CO_2, so may correspond to China's resurgent growth thereafter.

Table 9 ranks the selected variables by relative importance based on the partial R^2. Anthropogenic components rank similarly to model 1, with $\text{IP}_{2_{t-4}}$ being the second largest (non-autoregressive) contributor to explained variation after Eurasian vegetation. Though, the measure of relative importance matters – based on hierarchical partitioning anthropogenic measures are ranked lower than vegetation, while still above temperature and SOI.

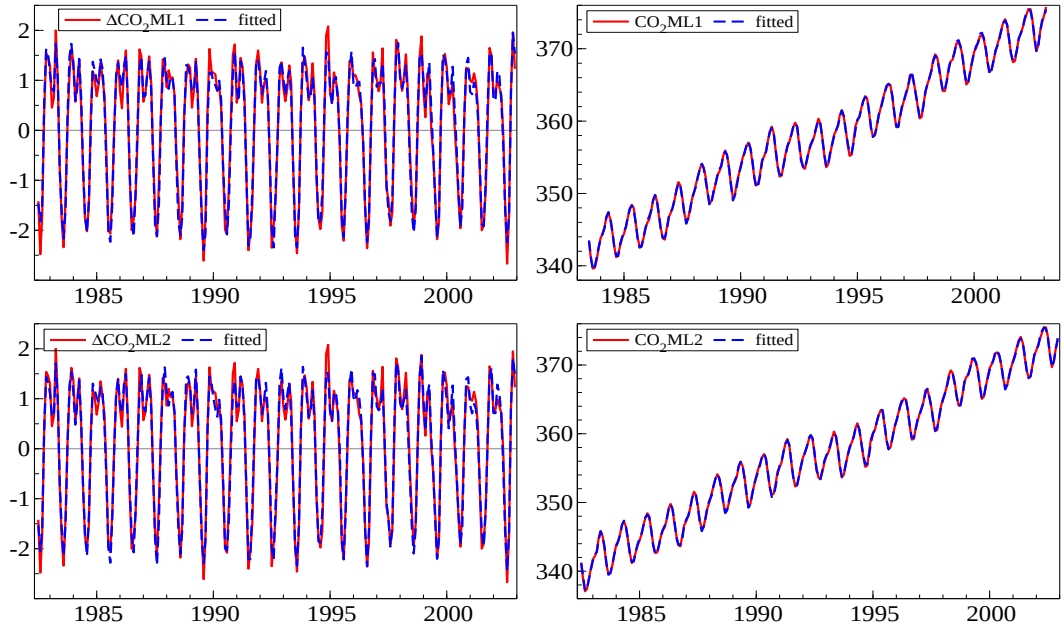
Comparisons between sample 1 and sample 2

GUM sample 1 and GUM sample 2 are estimated as robustness checks. Comparing *Autometric* selection in sample 1 and 2 there is highly consistent selection of key variables. Industrial production, temperature, vegetation and Southern Oscillation are selected identically in both models. Anthropogenic emissions are consistently selected in the form of high-frequency industrial production rather than low-frequency annual emissions. Conversely, selection of regional NDVI varies slightly between models, which could be the result of a similar growing season and pattern in North America and Eurasia as measured through NDVI data. Differences in selection from sample

Table 9: GUM: Sample 2 relative importance

Variable	**Partial R^2**	**hierarchical partitioning**
$\Delta CO_{2,t-2}$	0.3582	0.0298
$NDVI_{1,\ Eur_{t-3}}$	0.2147	0.029
$IP_{2_{t-4}}$	0.1814	0.0040
$NDVI_{1,\ Eur_{t}}$	0.1376	0.175
$Temp_{t-4}$	0.1349	0.0055
SOI_{t-5}	0.1149	0.0013
$IP_{1_{t-4}}$	0.1109	0.0067
$NDVI_{1,\ Eur_{t-11}}$	0.1098	0.095
$w\_NDVI_{3,\ Eur_{t-8}}$	0.0974	0.0122
$s\_NDVI_{1,\ NA_{t-8}}$	0.0936	0.055
$s\_NDVI_{1,\ NA_{t-9}}$	0.0869	0.0464
IP_{3_t}	0.084	0.0034
$NDVI_{1,\ Eur_{t-12}}$	0.0813	0.174
$\Delta CO_{2,t-4}$	0.0805	0.0778
$IP_{1_{t-1}}$	0.0792	0.0086
$w\_NDVI_{2,\ Eur_{t-7}}$	0.077	0.078
$w\_NDVI_{2,\ Eur_{t-8}}$	0.0605	0.1125
$NDVI_{2,\ Eur_{t-11}}$	0.0555	0.081
$I_{1990(7)}$	0.0507	0.0018

Figure 15: Graphs of fitted and actual values in differences and levels



1 to sample 2 may seem surprising given that all vegetation variables in sample 1 are identical to those in sample 2. However, the PCs are only orthogonal within regions, and are highly correlated between as figure 7 shows. The result of different selection may be due to computational short-cuts, namely dropping branches in the tree search after a model failed back-testing, and block partitioning in *Autometrics*, which adversely affect consistent selection.

Overall, based on *Autometrics* selection, natural factors such as vegetation, temperature, Southern Oscillation are necessary but not sufficient in explaining changes in atmospheric CO_2 measured at Mauna Loa. Industrial production variables are consistently selected. Most estimated effects from selected variables affect ΔCO_2 with a lag, and there seem to be few or no structural breaks in the relationships being modelled. As a further robustness check, future work will involve applying the estimation method to other measurement stations such as Barrow, Alaska (Keeling, Piper, Bollenbacher and Walker, 2008). Atmospheric carbon dioxide at Barrow has been measured from 1974:2 until 2007:12 and displays higher amplitude and higher autocorrelation due to its location relative to Mauna Loa.

Using relatively few assumptions, automatic model selection with *Gets* modelling can provide a tool to successfully model complex relationships. Starting from a broad GUM that nests the LDGP theoretically (accounting for natural and anthropogenic sinks and sources), the analysis proceeds with an agnostic approach to determine the key factors in changes in atmospheric CO_2. First, automatic model selection reduces the GUM to terminal models. Shortcomings of automatic model selection are computational issues in *Autometrics* selection because of short cuts and block partitioning. However, the terminal models appear congruent (acceptable diagnostic tests, few to no indicators or seasonal dummies selected) and are also supported by theoretical conclusions from the broader literature—natural effects are selected with the expected signs on coefficients. Key in the results is that additionally to the natural determinants selected, all terminal models include a large number of anthropogenic factors. The explained variance is decomposed to establish the relative importance of the variables. Using hierarchical partitioning and ranking based on partial correlations, the anthropogenic contribution to explained variation is quantifiable and found to be high.

From changes to levels

The models estimated as GUM sample 1 and GUM sample 2 describe the change in atmospheric carbon dioxide, ΔCO_2. While most of the analysis has been focused on the change in CO_2, it is possible to recover the estimated level from the models 1 and 2. Level estimates for model 1 and model 2, based on $\widehat{CO}_{2,t} = \widehat{\Delta CO}_{2,t} + CO_{2,t-1}$ are given in figure 15 and show the close fit. The next step is to attribute the components of the long-run explanation. To do so, we derive the relation after all dynamics from lagged variables have been solve out (the 'long-run solution': see Hendry, 1995). In a simple ADL(1) of the form:

$$y_t = \lambda_1 x_{1,t} + \lambda_2 y_{t-1} + e_t \tag{11}$$

where $|\lambda_2| < 1$, the long-run conditional expected value is:

$$\mathsf{E}[y \mid x] = \lambda_1 x_1/(1-\lambda_2) = \beta_1 x_1 \tag{12}$$

Then, based on the theoretical specifications given in equations (5)–(7), the solved estimated model is in the approximate form given in equation (13):

$$\Delta CO_{2,t} = \beta_1 x_{1,t} + \beta_2 x_{2,t} + \cdots + \beta_q x_{q,t} + \epsilon_t \tag{13}$$

This can be re-written as:

$$CO_{2,t} = CO_{2,t-1} + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \cdots + \beta_q x_{qt} + \epsilon_t \tag{14}$$

Recursive substitution for $CO_{2,t-1}, CO_{2,t-2}, ...$ in equation (14) yields:

$$CO_{2,t} = CO_{2,0} + \beta_1 \sum_{j=0}^{t} x_{1,t-j} + \beta_2 \sum_{j=0}^{t} x_{2,t-j} + \cdots + \beta_q \sum_{j=0}^{t} x_{q,t-j} + e_t \tag{15}$$

We divide the variables into two different groups: let s be the number of variables $x_{i,t}$ that have a stationary cumulative sum (non-trending) $\sum_{j=0}^{t} x_{i,t-j} \sim \mathsf{I}(0)$, so that $q - s$ variables have non-stationary cumulative sums (trending) $\sum_{j=0}^{t} x_{i,t-j} \sim \mathsf{I}(r)$ where $r > 0$. Equation 15 can then be expressed as:

$$CO_{2,t} = CO_{2,0} + \mathbf{x}'_{s,t}\boldsymbol{\beta}_s + \mathbf{x}'_{q-s,t}\boldsymbol{\beta}_{q-s} + e_t \tag{16}$$

where $\mathbf{x}_{s,t}$ and $\mathbf{x}_{q-s,t}$ are $s \times 1$ and $(q-s) \times 1$ column vectors respectively with $\sum_{j=0}^{t} x_{i,t-j}$ as their row elements. Equation (16) implies that the trending level of CO_2 is a function of the cumulative sums of the stationary ($\mathbf{x}_{s,t}$) and non-stationary ($\mathbf{x}_{q-s,t}$) variables in our model. Stationary variables in $\mathbf{x}_{s,t}$ by nature cannot drive the trend. Only explanatory variables with non-stationary cumulative sums, $\mathbf{x}_{q-s,t}$, determine the trend. This provides a straight-forward method of evaluating the underlying factors of the long-run trend.

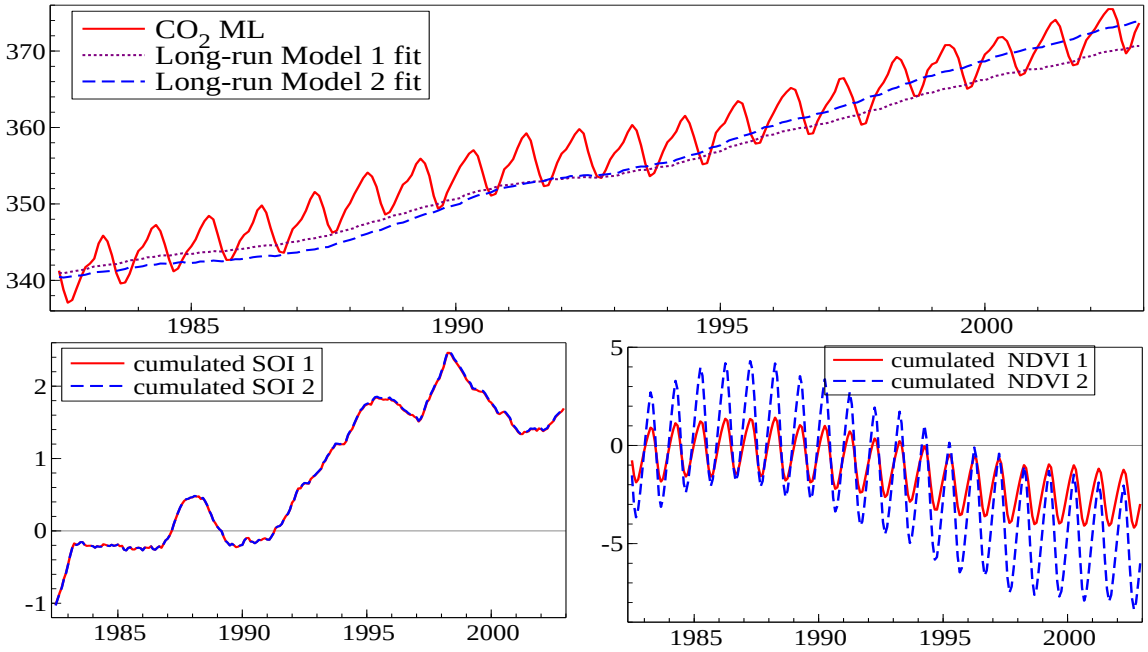
Out of the selected variables in models 1 and 2, only a sub-set exhibit trending cumulative sums, which are all the anthropogenic factors and the temperature anomaly. Both natural controls of NDVI and SOI remain approximately stationary around zero over time. Importantly, neither final model includes a deterministic intercept or trend, which on summation would become a linear or a quadratic time trend. However, summed variables do not have a straight-forward interpretation in the case of PCs of industrial output and temperature. The trending temperature anomaly is likely a mutually supporting feedback effect, as mentioned in section 4. Overall, the trend in the levels of CO_2 is derived from the trends in the independent variables in both estimated models, so is driven primarily by the PCs and partly by temperature. Specifically, the empirical equivalents of (16) for model 1 are:

$$\mathbf{x}'_{s,t}\boldsymbol{\beta}_s = -0.0038 \sum_{j=0}^{t} SOI_j + (-0.0275 \sum_{j=0}^{t} \text{NDVI}_{1,\,\text{Eur}_j} + 0.0169 \sum_{j=0}^{t} w\_\text{NDVI}_{3,\,\text{Eur}_j}) \tag{17}$$

$$\mathbf{x}'_{q-s,t}\boldsymbol{\beta}_{q-s} = -0.037 \sum_{j=0}^{t} \text{IP}_{1_j} - 0.127 \sum_{j=0}^{t} \text{IP}_{2_j} + 0.097 \sum_{j=0}^{t} \text{IP}_{3_j} + 0.0018 \sum_{j=0}^{t} \text{Temp}_j \tag{18}$$

Figure 16 shows the resulting coefficient-weighted cumulative sums of vegetation (NDVI_{EUR} principal components), SOI, and the combined anthropogenic components (IP_1 to IP_3 and temperature trend) for both models 1 and 2, together with the recorded level of CO_2. The industrial production components and temperature approximate the level of CO_2 well, marking a slight slow down in the trend around 1991–1993. Both cumulative stationary components vary over a small range so contribute little to the long-run changes. Even though the model is estimated in net inflows to atmospheric carbon dioxide, in re-parametrized form it can explain the long-run trend—and attributes it primarily to anthropogenic emissions. This is an outcome of the data analysis and is not enforced.

Figure 16: Level CO_2 and cumulative sums of anthropogenic and natural factors



7 Conclusions

We identified anthropogenic contributions to atmospheric CO_2 measured at Mauna Loa using an automatic model selection algorithm. Traditionally, estimation of anthropogenic effects on carbon dioxide relied on *a priori* selection of variables, which may not be appropriate in complex relations, low-frequency measures of anthropogenic emissions, and decompositions of time series. Using *Autometrics* in a general to specific modeling approach allows for model selection with more variables than observations, stringent mis-specification testing and a more agnostic way of modeling complicated interactions. The algorithm is applied to model changes in atmospheric CO_2 controlling for natural as well as anthropogenic sinks and sources without *a priori* restrictions of the determinants. While not completely robust to initial specification, we find that natural factors such as vegetation, temperature and Southern Oscillation are necessary, but not sufficient in explaining variation of atmospheric CO_2. Industrial production components measured monthly are highly significant and consistently selected in the estimated models. A higher proportion of variation should be attributed to anthropogenic sources than has been previously the case in the literature. Producing congruent models, our methodology introduces *Gets* modeling through *Autometrics* as a useful tool in modeling complicated climate relationships.

References

Australian Bureau of Meteorology (2011). Southern oscillation index archive. http://www.bom.gov.au/climate/current/soihtm1.shtml.

Bacastow, R. B. (1976). Modulation of atmospheric carbon dioxide by the southern oscillation. *Nature*, **261**, 116–118.

Bacastow, R. B., Keeling, C. D., and Whorf, T. P. (1985). Seasonal amplitude increase in atmospheric CO2 concentration at Mauna Loa, hawaii 1959-1982. *Journal of Geophysical Research*, **90**, 10,529–10,540.

Banerjee, A., Dolado, J. J., Galbraith, J. W., and Hendry, D. F. (1993). *Co-integration, Error Correction and the Econometric Analysis of Non-Stationary Data*. Oxford: Oxford University Press.

Bardsen, G., Hurn, S., and McHugh (2010). Asymmetric unemployment rate dynamics in Australia. Creates research paper 2010-2, Aarhus University, Denmark.

Bontemps, C., and Mizon, G. E. (2008). Encompassing: Concepts and implementation. *Oxford Bulletin of Economics and Statistics*, **70**, 721–750.

Buermann, W., Lintner, B. R., Koven, C. D., Angert, A., Pinzon, J. E., Tucker, C. J., and Fung, I. Y. (2007). The changing carbon cycle at Mauna Loa observatory. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 4249–4254.

Campos, J., Ericsson, N. R., and Hendry, D. F. (2005). Editors' introduction. In Campos, J., Ericsson, N. R., and Hendry, D. F. (eds.), *Readings on General-to-Specific Modeling*, pp. 1–81. Cheltenham: Edward Elgar.

Castle, J. L., Doornik, J. A., and Hendry, D. F. (2011a). Evaluating automatic model selection. *Journal of Time Series Econometrics*, **3 (1)**, DOI: 10.2202/1941–1928.1097.

Castle, J. L., Doornik, J. A., and Hendry, D. F. (2011b). Model selection in equations with many 'small' effects. Discussion paper 528, Economics Department, Oxford University.

Castle, J. L., Doornik, J. A., and Hendry, D. F. (2011c). Model selection when there are multiple breaks. *Journal of Econometrics*, forthcoming.

Castle, J. L., and Hendry, D. F. (2011). Automatic selection of non-linear models. In Wang, L., Garnier, H., and Jackman, T. (eds.), *System Identification, Environmental Modelling and Control*, pp. 229–250. New York: Springer.

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

Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, **28**, 591–605.

Christopher, S. L., Feely, R. A., Gruber, N., Key, R. M., Lee, K., Bullister, J. L., Wanninkhof, R., Wong, C. S., Wallace, D. W. R., Tilbrook, B., Millero, F. J., Peng, T. H., Kozyr, A., Ono, T., and Rios, A. F. (2007). The oceanic sink for anthropogenic CO2. *Science*, **305**, 367–371.

Clark, W. C. (ed.)(1982). *Carbon Dioxide Review*. New York: Oxford University Press.

Copin-Montigut, C. (1988). A new formula for the effect of temperature on the partial pressure of CO2 in seawater. *Marine Chemistry*, **25**, 29–37.

Dickey, D. A., and Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, **49**, 1057–1072.

Doornik, J. A. (2008). Encompassing and automatic model selection. *Oxford Bulletin of Economics and Statistics*, **70**, 915–925.

Doornik, J. A. (2009). Autometrics. In Castle, and Shephard (2009), pp. 88–121.

Doornik, J. A. (2010). Econometric model selection with more variables than observations. Working paper, Economics Department, University of Oxford.

Doornik, J. A., and Hansen, H. (2008). An omnibus test for univariate and multivariate normality. *Oxford Bulletin of Economics and Statistics*, **70**, 927–939.

Doornik, J. A., and Hendry, D. F. (2009). *Empirical Econometric Modelling using PcGive: Volume I*. London: Timberlake Consultants Press.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity, with estimates of the variance of United Kingdom inflation. *Econometrica*, **50**, 987–1007.

Enting, I. G. (1987). The interannual variation in the seasonal cycle of carbon dioxide concentration at Mauna Loa. *Journal of Geophysical Research*, **92**, 5497–5504.

Erickson, D. J., Mills, R. T., Gregg, J., J. Blasing, T. J., Hoffmann, F. M., Andres, R. J., Devries, M., Zhu, and Kawa, S. R. (2008). An estimate of monthly global emissions of anthropogenic CO2: Impact on the seasonal cycle of atmospheric CO2. *Journal of Geophysical Research*, **113**, –.

Federal Reserve (2011). Industrial production and capacity utilization. http://www.federalreserve.gov/releases/g17.

Francey, R. J., Tanis, P. P., Allison, C. E., Enting, I. G., White, J. W. C., and Trolier, M. (1995). Changes in oceanic and terrestrial carbon uptake since 1982. *Nature*, **373**, 326–330.

Gerlach, T. M. (2011). Changes in oceanic and terrestrial carbon uptake since 1982. *EOS, Transactions, American Geophysical Union*, **92**, 201–202.

Godfrey, L. G. (1978). Testing for higher order serial correlation in regression equations when the regressors include lagged dependent variables. *Econometrica*, **46**, 1303–1313.

Granados, J. A. T., Ionides, E. L., and Carpintero, O. (2009). A threatening link between world economic growth and atmospheric CO2 concentrations. Working Paper.

Granados, J. A. T., Ionides, E. L., and Carpintero, O. (2011). Climate change and the world economy: Short-run and long-run determinants of amtospheric CO2. Working Paper.

Groemping, U. (2007). Estimators of relative importance in linear regression based on variance decomposition. *American Statistician*, **61:2**, 139–147.

Hansen, J., and Lebedeff, S. (1987). Global trends of measured surface air temperature. *Journal of Geophysical Research*, **92**, 13,345–13,372.

Hansen, J., Ruedy, R., Sato, M., and Lo, K. (2010). Global surface temperature change. *Review of Geophysics*, **48**, 2010RG000345.

Hansen, J., Russel, G., Rind, D., Stone, P., Lacis, A., Lebedeff, S., Ruedy, R., and Travis, L. (1983). Efficient three-dimensional global models for climate studies: Models I and II. *Monthly Weather Review*, **111**, 609–662.

Hansen, J., Sato, M., Kharecha, P., Beerling, D., Berner, R., Masson-Delmotte, V., Pagani, M., Raymo, M., Royer, D. L., and Zachos, J. C. (2008). Target atmospheric CO2: Where should humanity aim?. *The Open Atmospheric Science Journal*, **2**, 217–231.

Hendry, D. F. (1995). *Dynamic Econometrics*. Oxford: Oxford University Press.

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

Hendry, D. F. (2011). Climate change: Possible lessons for our future from the distant past. In Dietz, S., Michie, J., and Oughton, C. (eds.), *The Political Economy of the Environment*, pp.

19–43. London: Routledge.

Hendry, D. F., and Johansen, S. (2011). The properties of model selection when retaining theory variables. Discussion paper, Economics Department, University of Copenhagen.

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

Hendry, D. F., and Juselius, K. (2000). Explaining cointegration analysis: Part I. *Energy Journal*, **21**, 1–42.

Hendry, D. F., and Krolzig, H.-M. (2003). New developments in automatic general-to-specific modelling. In Stigum, B. P. (ed.), *Econometrics and the Philosophy of Economics*, pp. 379–419. Princeton: Princeton University Press.

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

Hendry, D. F., and Mizon, G. E. (2011). Econometric modelling of time series with outlying observations. *Journal of Time Series Econometrics*, **3 (1)**, DOI: 10.2202/1941–1928.1100.

Hendry, D. F., and Richard, J.-F. (1989). Recent developments in the theory of encompassing. In Cornet, B., and Tulkens, H. (eds.), *Contributions to Operations Research and Economics. The XXth Anniversary of CORE*, pp. 393–440. Cambridge, MA: MIT Press.

Hendry, D. F., and Santos, C. (2010). An automatic test of super exogeneity. In Watson, M. W., Bollerslev, T., and Russell, J. (eds.), *Volatility and Time Series Econometrics*, pp. 164–193. Oxford: Oxford University Press.

Hoffman, P. F., and Schrag, D. P. (2000). Snowball Earth. *Scientific American*, **282**, 68–75.

Hofman, D. J., Butler, J. H., and Tans, P. P. (2009). A new look at atmospheric carbon dioxide. *Atmospheric Environment*, **43**, 2084–2086.

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

Jones, C. D., and Cox, P. M. (2005). On the significance of atmospheric CO2 growth rate anomalies in 2002-2003. *Journal of Geophysical Research*, **32**, –.

Juselius, K., and Kaufmann, R. (2009). Long-run relationships among temperature, CO2, methane, ice and dust over the last 420,000 years: Cointegration analysis of the Vostok ice core data. *IOP Conference Series: Earth and Environmental Science*, **6**, –.

Kaufmann, R. K., Paletta, L. F., Tian, H. Q., Myeni, R. B., and D'Arrigo, R. D. (2008). The power of monitoring stations and a CO2 fertilization effect: Evidence from causal relationships between NDVI and carbon dioxide. *Earth Interactions*, **12**, 1–23.

Kawa, S. R., Erickson, D. J., Pawson, S., and Zhu (2004). Global CO2 transport simulations using meteorological data from NASA data assimilation system. *Journal of Geophysical Research*, **109**, D18312.

Keeling, C. D. (1973). Industrial production of carbon dioxide from fossil fuels and limestone. *Tellus*, **25:2**, 174–198.

Keeling, C. D., Bacastow, R. B., Brainbridge, A. E., Ekdahl, C. A., Guenther, J. P. R., and Waterman, L. S. (1976). Atmospheric carbon dioxide variations at Mauna Loa observatory, Hawaii. *Tellus*, **6**, 538–551.

Keeling, C. D., Chin, J. F., and Whorf, T. P. (1996). Increased activity of northern vegetation

inferred from atmospheric CO2 measurements. *Nature*, **382**, 146–149.

Keeling, C. D., and Revelle, R. (1985). Effects of el nino/southern oscillation on the atmospheric content of carbon dioxide. *Meteoritics*, **20:2**, 437–450.

Keeling, C. D., Whorf, T. P., Whalen, M., and van der Plicht, J. (1995). Interannual extremes in rate rise of atmospheric carbon dioxide since 1980. *Nature*, **375**, 666–670.

Keeling, R. F., Piper, S. C., Bollenbacher, A. F., and Walker, S. J. (2008). Atmospheric co2 records from sites in the sio air sampling network. In Trends: A Compendium of Data on Global Change. Carbon Dioxide Information Analysis Center, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A. http://cdiac.ornl.gov/trends/co2/sio-bar.html.

Kohlmaier, G. H., Sire, E. O., Janecek, A., Keeling, C. D., Piper, S. C., and Revelle, R. (1989). Modelling the seasonal contribution of a CO2 fertilization effect of the terrestrial vegetation to the amplitude increase in atmospheric CO2 at Mauna Loa observatory. *Tellus*, **41B**, 487–510.

Kruskal, W. (1987). Relative importance by averaging over orderings. *American Statistician*, **41:1**, 6–10.

Levin, I., Graul, R., and Trivett, N. B. A. (1995). Long-term observations of atmospheric CO2 and carbon isotopes at continental sites in Germany. *Tellus*, **47B**, 23–34.

Lucht, W., Prentice, I. C., Myeni, R. B., Sitch, S., Friedlingstein, P., Cramer, W., Bousquet, P., Buermann, W., and Smith, B. (2002). Climate control of the high-latitude vegetation greening trend and Pinatubo effect. *Science*, **296**, 1687–1689.

Marland, G., Boden, T. A., and Andres, R. (2011). Global, regional, and national fossil fuel CO2 emissions. In Trends: A Compendium of Data on Global Change. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A. http://cdiac.ornl.gov/trends/emis/overview.html.

Marland, G., and Rotty, R. M. (1984). Carbon dioxide emissions from fossil fuels: a procedure for estimation and results for 1950-1982. *Tellus*, **36B**, 232–261.

Ministry of Statistics Government of India (2011). Index of industrial production (IIP). http://mospi.nic.in/.

Mizon, G. E., and Richard, J. F. (1986). The encompassing principle and its application to non-nested hypothesis tests. *Econometrica*, **54**, 657–678.

Myeni, R. B., Hall, F. G., Sellers, P. J., and Marshak, A. L. (1995). The interpretation of spectral vegetation indexes. *Transactions on Geoscience and Remote Sensing*, **33:2**, 481–486.

NASA Goddard Institute for Space Studies (GISS) (2011). GISS surface temperature analysis. http://data.giss.nasa.gov/gistemp/.

Nevison, C. D., Mahowald, N. M., Doney, S. C., Lima, I. D., van der Werf, G. R., Randerson, J. T., Baker, D. F., Kasibhatla, P., and McKinley, G. A. (2008). Contribution of ocean, fossil fuel, land biosphere, and biomass burning carbon fluxes to seasonal and interannual variability in atmospheric CO2. *Journal of Geophysical Research*, **113**, –.

Newell, N. D., and Marcus, L. (1987). Carbon dioxide and people. *Society for Sedimentary Geology*, **2:1**, 101–103.

OECD (2011). Composite leading indicators: MEI. http://stats.oecd.org/Index.aspx.

Office of National Statistics (2011). Primary production. http://www.statistics.gov.uk/hub/business-energy/.

Orr, C. J., Maier-Reimer, E., Mikolajewicz, U., Monfray, P., Sarmiento, J. L., Toggweiler, J. R., Taylor, N. K., Palmer, J., Gruber, N., Sabine, C. L., Le Quere, C., Key, R. M., and Boutin, J. (2001). Estimates of anthropogenic carbon uptake from four three-dimensional global ocean models. *Global Biogeochemical Cycles*, **15:1**, 43–60.

Randerson, T. J., Thompson, M. V., Conway, T. J., Fung, I. Y., and Field, C. B. (1997). The contribution of terrestrial sources and sinks to trends in the seasonal cycle of atmospheric carbon dioxide. *Global Biogeochemical Cycles*, **11:4**, 535–560.

Ritschard, R. L. (1992). Marine algae as a CO2 sink. *Water, Air and Soil Pollution*, **64**, 289–303.

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

Salkever, D. S. (1976). The use of dummy variables to compute predictions, prediction errors and confidence intervals. *Journal of Econometrics*, **4**, 393–397.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, **6**, 461–464.

Sims, C. A., Stock, J. H., and Watson, M. W. (1990). Inference in linear time series models with some unit roots. *Econometrica*, **58**, 113–144.

Sundquist, E. T., and Keeling, R. F. (2009). The Mauna Loa carbon dioxide record: Lessons for long-term earth observations. *Geophysical Monograph Series*, **183**, 27–35.

Taguchi, S., Murayama, S., and Higuchi, K. (2003). Sensitivity of inter-annual variation of CO2 seasonal cycle at Mauna Loa to atmospheric transport. *Tellus*, **55B**, 547–554.

Tans, P., and Keeling, R. (2011). Mauna Loa, monthly mean carbon dioxide. Scripps Institution of Oceanography. (scrippsco2.ucsd.edu/) and NOAA/ESRL http://www.esrl.noaa.gov/gmd/ccgg/trends/.

Thoning, K. W., and Tans, P. P. (1989). Atmospheric carbon dioxide at Mauna Loa observatory 2. Analysis of the NOAA GMCC data, 1974-1985. *Journal of Geophysical Research*, **94**, 8549–8565.

Troup, A. J. (1965). The 'Southern Oscillation'. *Quarterly Journal of the Royal Meteorological Society*, **91**, 490–506.

Tucker, C. J., Pinzon, J., Brown, M., and GIMMS/GSFC/NASA (2010). ISLSCP II GIMMS Monthly NDVI, 1981-2002. In Hall, Forrest G., G. Collatz, B. Meeson, S. Los, E. Brown de Colstoun, and D. Landis (eds.). ISLSCP Initiative II Collection. Data set. http://daac.ornl.gov/ from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A.

White, H. (1980). A heteroskedastic-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, **48**, 817–838.

Wigley, T. M. L. (1983). The pre-industrial carbon dioxide level. *Climate Change*, **5**, 315–320.

Worrell, E., Price, L., Martin, N., Hendriks, C., and Meida, L. O. (2001). Carbon dioxide emissions from the global cement industry. *Annual Review of Energy and the Environment*, **26**, 303–329.