Quantifying the influence of anthropogenic surface processes and inhomogeneities on gridded global climate data

Ross R. McKitrick
Department of Economics
University of Guelph
Guelph ON Canada N1G 2W1
rmckitri@uoguelph.ca
519-824-4120 X52532

Patrick J. Michaels
Cato Institute
1000 Massachusetts Avenue NW
Washington DC 20001
pmichaels@cato.org

DECEMBER 2007

IN PRESS: JOURNAL OF GEOPHYSICAL RESEARCH-ATMOSPHERES

Quantifying the influence of anthropogenic surface processes and inhomogeneities on gridded global climate data

Abstract

Local land surface modification and variations in data quality affect temperature trends in surface-measured data. Such effects are considered extraneous for the purpose of measuring climate change, and providers of climate data must develop adjustments to filter them out. If done correctly, temperature trends in climate data should be uncorrelated with socioeconomic variables that determine these extraneous factors. This hypothesis can be tested, which is the main aim of this paper. Using a new data base for all available land-based grid cells around the world we test the null hypothesis that the spatial pattern of temperature trends in a widely-used gridded climate data set is independent of socioeconomic determinants of surface processes and data inhomogeneities. The hypothesis is strongly rejected ($P = 7.1 \times 10^{-14}$), indicating that extraneous (nonclimatic) signals contaminate gridded climate data. The patterns of contamination are detectable in both rich and poor countries, and are relatively stronger in countries where real income is growing. We apply a battery of model specification tests to rule out spurious correlations and endogeneity bias. We conclude that the data contamination likely leads to an overstatement of actual trends over land. Using the regression model to filter the extraneous, nonclimatic effects reduces the estimated 1980-2002 global average temperature trend over land by about half.

Index terms: Atmosphere; Land/atmosphere interactions; Instruments and techniques; Climate change and variability

Quantifying the influence of anthropogenic surface processes and inhomogeneities on gridded global climate data

1 Introduction

It has long been known that local economic conditions and demographic changes leave measurable traces in meteorological records. Climate data, as used for measuring global warming and detecting a CO_2 influence, originates with meteorological records, but it then undergoes a modeling step, the aim of which is to identify and remove all such extraneous signals, in principle yielding an estimate of the air temperature trend in a location, had there never been any human settlement there (see, e.g. Mitchell 1953, Peterson 2003). Typical usage of climate data assumes this filtering to have taken place, such that contaminating signals due to socioeconomic factors leave only small, unsystematic, zero-mean and zero-trend noise in climatic data series. If true, the spatial pattern of observed climatic trends should be uncorrelated with socioeconomic measures that account for variations in extraneous, nonclimatic signals in the underlying meteorological data. This hypothesis can be tested, which is the main aim of this paper. Our data and model allow us to test for a range of extraneous signals in climatic records which are categorized into modifications to the local environment (anthropogenic surface processes) and observational difficulties (or data inhomogeneities). We reject the hypothesis that the spatial pattern of temperature trends in global climate data is independent of extraneous effects ($P = 7.1\times10^{-14}$). We present evidence that our results are not due to reverse causality (endogeneity bias) or spurious correlations. The economic imprints are present in both rich and poor countries but are strongest in countries experiencing real income growth. The effects are significant at the global level and likely add a sizable upward bias to trends in the global mean temperature anomaly. Our results suggest that as much as half of measured post-1980 land-based "global warming" may be attributable to contamination of the basic data.

Over 50 years ago, referring to the use of long-term weather records for measuring climate change, J. Murray Mitchell Jr. (1953, page 244) cautioned: "The problem remains one of determining what part of a given temperature trend is climatically real and what part the result of observational difficulties and of artificial modification of the local environment." These two types of bias continue to affect the measurement of climate change. Observational difficulties, or data inhomogeneities (such as station moves and closure, record discontinuities, equipment change and changes to the time of observation) are known to have affected records of mean temperature (e.g. Baker 1975, Schal and Dale 1977, Karl and Williams 1987, Willmott et al. 1991, Peterson 2003). Modification of the land-surface, including urbanization and other economic activity, has been shown to affect local, regional and possibly global meteorology, and thus locally-measured temperature data (e.g. Feddema et al. 2005a,b; Pielke et al 2002; McKendry 2003; de Laat and Maurellis 2004, 2006; McKitrick and Michaels 2004). For some local meteorological purposes these extraneous effects may not matter, but for applications in which weather data are used to construct measures of long term climate change and detect anthropogenic influences, a modeling step is required to measure and filter them out. The variety of methods in common use will be discussed in the next section. The difficulties in filtering extraneous effects continue to be noted in empirical climate studies (e.g. Peterson 2003) but both the Third and Fourth Assessment Reports of the Intergovernmental Panel on Climate Change (IPCC 2001, 2007) assert that such biases at the global level are extremely small. Also, climate change attribution studies (e.g. Tett et al. 1999) presuppose that trends in gridded climate data are only attributable to climatic "forcings" such as solar flux, atmospheric dust and greenhouse gas concentrations, on the assumption (stated or implied) that the spatial configuration of local modifications to the land surface or determinants of observational difficulties are not significant or systematic features of the gridded data.

If this assumption is true, then the spatial pattern of gridcell temperature trends should be uncorrelated with variables like Gross Domestic Product, population density, average income, and other local, nonclimatic factors. The presence of such correlations, on the other hand, would indicate that gridded surface climate data contain extraneous biases, thus measured climatic trends may be inaccurate and attempts to identify the climatic influences of greenhouse gases might misattribute the causes of apparent trends. Alternatively, if the spatial pattern of greenhouse warming just happens to match the spatial pattern of socioeconomic development, it suggests conventional signal detection methodology would be unable to identify which one explains the observed changes. However, this possibility is critiqued in Sections 4.4 and 4.6 below.

In this study we develop a new data base encompassing all available land-based grid cells around the world, matched to detailed local economic and social conditions, as well as fixed geographical factors. Our data and model allow us to test for a range of extraneous signals in climatic records which are categorized into land surface changes and observational difficulties (or data inhomogeneity). We find clear evidence that the spatial pattern of temperature trends in global climate data is significantly associated with each type of extraneous effect.

It is sometimes customary to refer to all non-atmospheric, extraneous effects as “nonclimatic.” This terminology is not entirely satisfactory, since land use change can also be considered a climatic influence (e.g. Feddema et al. 2005a,b). Mitchell (1953) proposed a taxonomy in which effects due to station movement, instrumental change, etc are denoted “Apparent” effects, those attributable to local environmental change (pollution, urbanization) are “Real-local” and effects due to atmospheric composition, solar flux etc. are denoted “Real-climatic.” The latter category is what gridded temperature anomaly data are said to measure; the first two are assumed to have been filtered out. To keep the terminology simple, we refer herein to apparent and real-local effects as “extraneous” biases.

2. Extraneous Biases in Climate Data

The urban heat island (UHI) effect is not the only source of data contamination, but has been the focus of particularly extensive investigation. A survey is in McKendry (2003). UHI effects have been documented in, for example: South Africa (Balling and Hughes 1996), Vienna (Böhm 1998), China (Jones et. al. 1990), Alaska (Magee, Curtis and Wendler 1999), Japan (Fujibe 1995), India (Hingane 1996), Illinois (Chagnon 1999), Korea (Chung et al., 2004), Turkey (Karaca et al. 1995), Poland (Klysik and Fortuniak 1999), Singapore and Kuala Lumpur (Tso 1995) etc. See Parker (2004) and Peterson (2003) for contrasting arguments. Guidelines exist for setting up climate monitoring stations so as to minimize the influence of siting on the recorded temperature data, but it is rare for the guidelines to be reliably met, even in the US (Davey and Pielke Sr., 2005). Typical adjustment models for urbanization are based on rural-urban comparisons (e.g. Jones et al. 1990) if sufficient data are available, or, most commonly, empirical parameterizations based on regressions against local population growth (McKendry 2003). UHI effects have been shown to arise even at very low levels of population, i.e. in towns with less than 10,000 people (Karl et. al. 1988). Changnon (1999) used a unique 64-year record of below-ground temperatures collected in rural Illinois to show that an upward bias was present in nearby weather stations that had been designated "rural" and assumed to be free of UHI problems. Böhm (1998) identified a substantial UHI in Vienna temperature records from 1951 to 1995 even though the city population had remained constant over the interval. Kalnay and Cai (2003) applied a technique for using 6 hour-ahead weather forecasts constrained by atmospheric observations from weather balloons to produce estimated surface temperature trends unaffected by land-use effects. In a study of the continental US, the comparison to standard surface temperature data suggested land-use changes, even in rural areas, had an effect on temperature records two to four times larger than previously thought.

Other methods for identifying and removing extraneous signals related to local land use include satellite-based measurement of surface energy flux to determine the urbanization component of regional temperature trends (Gallo and Owen 2002, Streutker 2003) and satellite measurement of night-time lighting (Hansen et al. 2001). However these approaches have not been widely applied, mainly because the necessary data are only available for the US.

The study of extraneous biases in surface temperature data has broadened out beyond the simple population-based approach, in recognition that there are changes that do not necessarily involve large population increases that can nevertheless affect regional temperatures (as in the example from Vienna from Böhm (1998) mentioned above): such as changes in agricultural activity and vegetation types, soil moisture, local air pollution levels, groundwater diversion, etc. These are referred to as “anthropogenic surface processes” (de Laat and Maurellis 2006). Economic variables have recently been introduced in some empirical climatological studies as a way of quantifying these processes, and have been shown to have significant explanatory power in regional and global climate data (e.g. Kalnay and Cai 2003). de Laat and Maurellis (2004, 2006) proposed interpreting carbon dioxide emissions as a proxy for local industrial activity, and thereby as an index of local extraneous warming influences on atmospheric temperature trends. This interpretation implies a particular spatial pattern of enhanced warming trends not predicted by climate models in response to greenhouse gas increases, but which they found to be clearly present in global temperature data collected both at the surface and the lower atmosphere. In McKitrick and Michaels (2004) we regressed the spatial pattern of trends from 93 countries on a matrix of local climatic variables and socioeconomic indicators such as income, education, and energy use. Some of the nonclimatic variables yielded significant coefficients. We then repeated the analysis on the IPCC gridded data covering the same locations and found approximately the same coefficients emerged, albeit diminished in size, with many individual indicators remaining significant. An error in the original regression program was found and corrected—see Erratum listed in citation—with little effect on the results. We concluded that the IPCC gridded data is contaminated by extraneous socioeconomic signals, a finding that is confirmed and strengthened in the present paper.

Temperature records are also potentially susceptible to discontinuities if a climate station is moved, malfunctions, or is de-staffed, or if the time of day at which the observations are taken changes (Baker 1975, Schaal and Dale 1977, Willmott et al. 1991 etc.). Collectively these effects are called “inhomogeneities.” Establishing a climate data series of uniform quality requires quantifying and removing the inhomogeneities. In some cases, written records exist of a station's history, revealing dates at which discontinuities may have emerged. Comparison of nearby stations can help identify and quantify sudden inhomogeneities at one site that might arise from equipment changes or construction near the instruments. But this is only feasible if there are many stations suitably close together, which is for the most part only true in parts of the US and Europe. Also it only removes short-term discontinuities and does not correct long term biases affecting multiple stations, such as those arising from regional urbanization (Mitchell and Jones 2005).

The challenge of producing quality global climate data arises in part because high-quality meteorological data is very costly to collect (see, e.g., Linacre 1992) and therefore changes in local and national economic conditions may induce inhomogeneities. The number of reliable monitoring sites around the world has fallen dramatically since the mid-1970s. The Global Historical Climatology Network reached a peak of 6,000 unique contributing sites in the late 1960s, but the number fell to fewer than 3,000 as of the late 1990s, with the most dramatic drop in the early 1990s (Peterson and Vose 1997) when the number of stations fell by nearly half in four years. The drop coincided with the collapse of the Soviet Union and a major international recession, and was not spatially uniform. A dramatic visualization of the loss of monitoring sites in the early 1990s is available at http://climate.geog.udel.edu/~climate/index.shtml. In its 2001 *Third Assessment Report* the IPCC warned that "unless networks [of climate monitoring equipment] are significantly improved, it may be difficult or impossible to detect climate change in many regions of the globe" ((IPCC 2001, Technical Summary, Page 78)).

Of the global climate data sets produced from available historical weather and climate data (Hansen et al. 2001, Peterson and Vose 2002, Jones and Moberg 2003) the "gridded" series from the Climate Research Unit (CRU) at the University of East Anglia (http://www.cru.uea.ac.uk/) are perhaps the best known, and are used for IPCC reports. The gridded data are disseminated by the IPCC as its reference climate data set (See http://ipcc-ddc.cru.uea.ac.uk/obs/cru_climatologies.html). The IPCC (2007) has downplayed concerns about extraneous biases by focusing on urbanization effects, estimating the influence as at most 0.006 C/decade globally (IPCC 2007, page 5). IPCC (2001) refers to Easterling et al. (1997) and Jones et al. (1990), both of which are confined to discussing UHI effects. Easterling et al. (1997) compared trends in global averages of climate data and reported minimal differences between pooled (rural and urban) results versus rural-only results. However, their definition of "rural" included cities up to 50,000 in population, which is large enough to exhibit a UHI. Jones et al. (1990) ran a similar comparison on three regions: Eastern Australia, Eastern China and Western USSR. Their definition of "rural" included towns of up to 10,000 in the USSR and up to 100,000 in China. They found relatively strong urban warming in China relative to the rural and pooled series, and in the USSR they found stronger relative cooling post-1930 in the rural stations. Eastern Australia yielded no differences. They also reported earlier results of strong relative warming in the contiguous USA. Although the conclusions of each paper were phrased optimistically, neither study suffices to alleviate concerns about extraneous effects, including general anthropogenic surface processes, in gridded IPCC temperature data.

Parker (2004) argued that UHI signals in IPCC temperature data do not have explanatory power at the global level, based on the similarity in trends between urban samples taken on calm nights versus windy nights. However elevated windspeed has been disputed as a factor in reducing UHI effects (see discussion in McKendry 2003), so the similarity in trends may simply indicate that the nonclimatic effects exert a similar influence under both conditions (see also Pielke Sr. and Matsui 2005). Peterson (2003) looked at more general data contamination issues by applying adjustments to US Historical Climatology Network data for variations in elevation, latitude, instrumental continuity and time of observation. He found that these sufficed to remove an observed difference in means between urban and rural temperatures (differences in trends were not reported). The time-of-observation bias had the largest effect, accounting for two-thirds of the initial rural-urban mean difference. One implication of the Peterson (2003) findings is that multiple sources of extraneous bias (not merely population growth) must be removed to homogenize temperature records. The closure of so many weather stations around the world since the 1980s raises the possibility that few countries, especially outside the developed world, have the staff or money to engage in such quality control efforts.

To summarize, both surface processes and inhomogeneities must be successfully filtered from temperature records to yield data products suitable for measuring global climate trends. If done correctly this would imply a lack of local correlations between observed temperature trends and socioeconomic trends. But we will show that such correlations clearly exist, supporting the conclusion that the filtering methods are not successful.

Another interesting implication of these issues concerns the attribution of climate change to greenhouse gas emissions. The roles of surface processes and inhomogeneities are ignored in attribution studies (e.g. Tett et. al 1999) on the assumption that they have already been removed from climate data. Measured temperature changes are regressed on a matrix of model-generated "forcing vectors" that predict the climatic responses to various combinations of solar irradiance, volcanic dust, greenhouse gases and sulfate aerosols. The test is whether observed data are consistent with the climate having an assumed sensitivity to greenhouse gas levels, and are inconsistent with zero sensitivity. Critical to the methodology are the assumptions that the climate model used to generate the forcings is substantially "true" and that the temperature data are free of extraneous nonclimatic patterns that might be confounded with the pattern of climatic changes resulting from greenhouse gases and sulfate emissions (Allen and Tett 1999). If the latter assumption is not true, components of observed climate change arising from, e.g., land surface processes may be wrongly attributed to greenhouse gas accumulation in the atmosphere (as pointed out in, e.g. Pielke Sr. et al. 2002).

The next section outlines the empirical model and the data set used in this paper. Subsequent sections present results and discussion.

3 Model and Data

3.1 A Model of Climate Measurement Distortions

Suppose there are i = 1, …, n locations around the world at which temperature is measured. In each location i a climatic trend T_i over the interval τ=[1979:1—2002:12] in °C/decade is sought, but what is actually measured is an observed trend θ_i:

$$\theta_i = T_i + f(S_i) + g(I_i) \quad (1),$$

where f and g are functions of unknown form, S_i represents surface processes and I_i represents inhomogeneities. Surface processes are represented using the percentage changes over the time interval τ in four socioeconomic variables: local population p_i, per capita income m_i, total Gross Domestic Product (GDP) y_i and coal consumption c_i. Inhomogeneities, or factors affecting data quality, are represented using three socioeconomic variables: GDP density g_i as of 1979, the average level of educational attainment e_i as late in the interval τ as possible, and the number of missing months in the observed temperature series x_i over the interval τ. Educational attainment is measured herein as the sum of national literacy and national post-secondary education rates. It is included not as an indicator of skill of the specific staff responsible for handling meteorological data, but as a measure of the difficulty of recruiting and retaining trained technical staff in general in that country. GDP density is national Gross Domestic Product per square kilometer. Countries with low GDP density (large land areas relative to their total national income) may have a measurement advantage if the low density arises due to high agricultural intensity. Some agricultural-based economies, even in low-income countries, have made a point of high quality weather data collection in support of their food-producing industry. However low GDP density is also a disadvantage if the country has a lot of uninhabited land to monitor relative to its resources. By using the GDP density at 1979 we capture the measurement conditions going into the interval rather than as they would have developed over the interval, and we ensure the measure is "predetermined" in an econometric sense. Possible endogeneity bias is discussed in Section 4.4. Other details on data sources are below.

In general, surface processes and inhomogeneities may introduce cold or warm biases into the data, and no *a priori* restrictions are imposed. For example, Feddema et al. (2005a,b) estimate that global land surface changes since before industrialization have yielded a net cooling effect on the climate system. The terms in f and g in (1) are observable, and a linear functional form will be assumed (though a RESET test will be applied to check for a nonlinear alternative: see section 4.3 below). To put equation (1) into a form useful for estimation would require observations of T_i, which are not available. Instead we assume T_i is a function of atmospheric data $\mathbf{A}_i$ that can represent the surface climatic temperature trend up to a multiplicative constant, so as to condition the estimated coefficients in (1):

$$T_i = h(\mathbf{A}_i) \equiv \beta_0 + \beta_1 TROP_i + \beta_2 PRESS_i + \beta_3 DRY_i + \beta_4 DSLP_i + \beta_5 WATER_i + \beta_6 ABSLAT_i \qquad (2).$$

The uniqueness up to a multiplicative constant arises since replacing T_i in equation (2) with kT_i, where k is an arbitrary constant to be estimated, would yield the same estimation results from equation (3) below, but k would not be identifiable. In effect, equation (2) assumes it has a value of 1, but the conclusions herein would be unchanged if k took a different value.

$TROP_i$ is the time trend of Microwave Sounding Unit (MSU)-derived temperatures in the lower troposphere in the same grid cell as θ_i over the same time interval, based on Spencer and Christy (1990) and published by the Global Hydrology and Climate Centre at the University of Alabama (GHCC 2005). Our interpretation assumes the Spencer-Christy data are substantially free of extraneous biases due to surface conditions, but de Laat and Maurellis (2004, 2006) have presented evidence that even MSU data exhibit some contamination by socioeconomic activity (see also Section 4.6). We comment on the implications of potential contamination of MSU data below in Section 6. We selected the Christy and Spencer series as it is a well-known data product that has been validated against independent data from weather balloons and other meteorological sources in overlapping regions (Pielke Sr. et al. 2004). We have not re-examined these results using other MSU-based tropospheric data series, but we do not expect any of the results reported herein to be contingent on the choice of MSU product. We use MSU version 5.2, released September 2005, reflecting corrections for all known errors due to orbital drift, instrument heating and diurnal averaging. The MSU data are expressed as monthly averages and are divided into grid cells that can be matched with IPCC data grid cells.

Geographic variables are defined as follows. $PRESS_i$ is the mean sea level air pressure in grid cell *i*. The source of the pressure data is the climatology of Jenne (1974), which is the most recent global data base of mean pressure readings we were able to find. DRY_i is a dummy variable denoting when a grid cell is characterized by predominantly dry conditions (which is indicated by the mean dewpoint being below 0 °C). $DSLP_i$ is $DRY_i \times PRESS_i$. Surface warming due to greenhouse gases is hypothesized to occur faster in regions with relatively dry air and high atmospheric pressure (Staley and Jurica 1970, Michaels et al. 2000) so pressure enters (2) as a linear spline function with a different intercept and slope in dry regions versus moist regions. $WATER_i$ is a dummy variable indicating the grid cell contains a major coastline. $ABSLAT_i$ denotes the absolute latitude of the grid cell. This is included to account for latitudinal changes in the rate of surface warming. It is sometimes conventional to use the cosine of latitude, which adjusts for declining grid cell size towards the poles, but this makes only trivial differences in the results. (Data and STATA code allowing readers to reproduce all our results, and experiment with different specifications, are archived at the journal web site.)

Equation (2) takes the observed temperature trend from the lower layers of the atmosphere above the surface, which are presumed to be closely coupled to surface trends but largely unaffected by the extraneous distortions in the surface record, and allows for location-specific geographical factors to account for differences between the trend aloft and that at the surface. Using (1) and (2) we can write out an estimating equation as follows:

$$\theta_i = \beta_0 + \beta_1 TROP_i + \beta_2 PRESS_i + \beta_3 DRY_i + \beta_4 DSLP_i + \beta_5 WATER_i + \beta_6 ABSLAT_i$$
$$+ \beta_7 p_i + \beta_8 m_i + \beta_9 y_i + \beta_{10} c_i + \beta_{11} e_i + \beta_{12} g_i + \beta_{13} x_i + u_i \qquad (3).$$

where u_i is the regression residual. While (3) cannot identify T_i, except under fortuitous circumstances which are not themselves testable, it allows us to test specific hypotheses regarding the independence of observed temperature trends from surface processes and determinants of inhomogeneities. Potential multicollinearity in (3) will be discussed in the results section below.

3.2 Other Data Sources

The variable names, definitions and summary statistics are shown in Table 1. The observed surface temperature trend θ_i consists of linear (Ordinary Least Squares) trends through monthly temperature anomalies (not subject to annual averaging) within 5x5 degree grid cells over 1979:1 to 2002:12 in 469 land-based grid cells in the ‘crutem2v’ data set available through the IPCC

Data Distribution Centre (http://ipcc-ddc.cru.uea.ac.uk/). Because of the need for a trend across 23 years we required each cell to have data for at least ninety percent of the years, where a year is considered intact if at least 8 months are available. This left 451 usable locations. 11 cells are in Antarctica, where there is no economy to speak of, several countries share jurisdiction over different research sites, and there is an anomalously high rate of missing values, probably due to the extreme conditions in which data are collected, so these were also removed. Hence there were 440 observations in the final data set. Of these, 348 (79 percent) were from the Northern Hemisphere and 92 were from the Southern Hemisphere. The imbalance is partly due to the fact that there is more land in the Northern Hemisphere, but also reflects the relative sparseness of continuous data in many parts of South America and Africa (see Figure 4 below). The $TROP_i$ variable is an OLS time trend through monthly data for grid cell *i* over the same interval.

3.2.1 Surface Process Data

Each grid cell was assigned to a country. Where a grid cell contained a border the country was considered the one with the most land area in the grid cell. Annual real (inflation adjusted) GDP for 1979, 1989 and 1999 for each country was obtained primarily from Easterly and Sewadeh (2003) or the Central Intelligence Agency (CIA) World Fact Book web site http://www.odci.gov/cia/publications/factbook/index.html. Conversions from local currency to US dollars was done using the purchasing power parity method.

There were small adjustments made to the economic data for some countries to provide consistency in quantities where direct measures were unavailable. In most cases the adjustment took the form of using an available observation for one or two years after the desired year, and adjusting it backwards.

Population data are obtained from Easterly and Sewadeh (2003) and the percent change p_i is measured from 1979 to 1999. Income growth m_i is the percentage change in real GDP per capita from 1979 to 1999. GDP growth y_i is defined as the percentage change in real GDP from 1979 to 1999. National coal consumption data were obtained from the US Energy Information Administration (http://www.eia.doe.gov/emeu/iea) and the coal growth measure is the percentage growth of short tons of coal consumed between 1980 and 2000.

Population and GDP density varies considerably within countries, as well as between countries. Hence national averages will not capture all the important variations that may influence the temperature data. However, the trade-off we face is between encompassing the full range of variables we want to include versus matching the grids of measurement of climatic and economic data. Since national governments bear primary responsibility for climate data collection, the nationally-defined economic measures will capture important information about the availability of resources to monitor the whole country's climate. Also, the substantial variation among countries implies that some of the effects of interest are definitely measured by the data we have available, albeit at a more coarse resolution than we would like. In the concluding section we will discuss the possibilities for future research arising from the development of some new socioeconomic data bases at the gridcell level.

3.2.2 Determinants of Inhomogeneities

We measure the abundance of human capital using data on international educational attainment. McKitrick and Michaels (2004) used national literacy rates as an indicator of the ease of maintaining a staff of trained meteorological technicians to operate weather stations. For the present study we have updated the literacy data to the 1999 (or closest year) national literacy rate (UNESCO 2003) and augmented it with estimates of the percentage completing post-secondary education (PSE), obtained from UNESCO (2003). The two measures are summed together to yield e_i. Qualitatively similar results would be obtained if we used either literacy or PSE alone, but by using the sum it controls for changes in one or the other. In the current sample, literacy averages 90 percent and PSE averages 16.6 percent. Literacy ranges from a low of 11 percent in Niger to over 99 percent throughout the industrialized countries. PSE ranges from less than one percent in many African countries up to 45 percent in the United States.

Land area estimates (excluding water) for each country were obtained from the CIA World Fact Book (CIA 2003). GDP density g_i is measured as \$million/km^2. The 1979 value is used to help ensure the right-hand side variables are predetermined with respect to the dependent variable, but see the further discussion of endogeneity in Section 4.4. A country with a low GDP density has relatively fewer national resources for monitoring its domestic land surface. g_i varies widely across the sample, from less than 0.01 in parts of Africa up to 4.5 in Japan and 4.8 in Taiwan. Canada, China and the US have comparable land masses (9.2, 9.3 and 9.5 million square km, respectively). But GDP density in China is 0.16 million\$/square km while in the US it is three times higher, at 0.47 million\$/square km and in Canada it is only one-third as large, at 0.05 million\$/square km. The global sample mean is 0.41, just below the US level, and most countries fall in the range between Canada and the US.

The variable x_i is the indicator of technical problems in maintaining continuous weather records. It is measured as the number of months over the period 1979-2002 in which an observation was missing for a grid cell. After removing the Antarctic stations only 95 out of 440 remaining cells (22 percent) had at least one missing month, and only 5 (1% of the sample) had more than 12 months missing. The distribution of missing data shows no pattern across months.

4 Estimation and Testing

4.1 Model Results

Equation (3) was estimated using Generalized Least Squares (GLS) as follows. Re-write (3) in matrix notation as follows:

$$\boldsymbol{\theta} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \tag{4}$$

where $\boldsymbol{\theta}$, $\boldsymbol{\beta}$ and $\mathbf{u}$ are *n*-vectors (dependent variable, coefficients and residuals, respectively) and $\mathbf{X}$ is the *n*x*k* matrix of independent variables. Coefficients were obtained using the least squares estimator $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\theta}$. The GLS variance-covariance estimator is

$$Var(\hat{\boldsymbol{\beta}}_{GLS}) = \hat{\mathbf{V}}(\mathbf{X}'\boldsymbol{\Omega}\mathbf{X})\hat{\mathbf{V}} \tag{5}$$

where $\mathbf{V} = (\mathbf{X}'\mathbf{X})^{-1}$, and the *n*x*n* matrix $\boldsymbol{\Omega}$ is the covariance matrix of $\mathbf{u}$. Following White (1980), a robust estimator of $\mathbf{X}'\boldsymbol{\Omega}\mathbf{X}$ can be obtained by replacing $\boldsymbol{\Omega}$ with a diagonal matrix formed with the squared residuals from (4), even if this is itself an inconsistent estimator of $\mathbf{\Omega}$ (see Davidson and MacKinnon 2004 p. 198), as long as observations are independent. In our data base some of the socioeconomic variables are constant within the 81 countries in our sample, resulting in possible non-independence (clustering) of errors within country groups. Denote the country groups as $C(1),\ldots,C(81)$. To allow within-cluster non-independence the estimator (5) is re-written as

$$Var(\hat{\boldsymbol{\beta}}^{c}_{GLS}) = \hat{\mathbf{V}}\left(\sum_{j=1}^{81}\boldsymbol{\xi}'_j\boldsymbol{\xi}_j\right)\hat{\mathbf{V}} \qquad (6)$$

where $\boldsymbol{\xi}_j = \left(\sum_{\kappa\in C(j)} u_\kappa \mathbf{x}_\kappa\right)$, $\kappa \in C(j)$ denotes the elements of cluster *j*, and $\mathbf{x}_\kappa$ is the κ-th row of $\mathbf{X}$ (StataCorp. 2003, pp. 274-275). OLS parameter estimates and the variances from (6) were estimated using STATA 8.0 (StataCorp. 2003). Estimates for equation (3) and various submodels are presented in Table 2.

Coefficient standard errors at the global level (Table 2, *SURF*) were also checked by the bootstrap method using 500 repetitions. Confidence intervals were quite stable. In all cases where a parameter is significant under GLS its confidence interval did not expand to encompass zero under bootstrap resampling.

The coefficient on $TROP_i$ is positive and significant as expected, and has a value of approximately 0.9 in all models, reflecting the expected correlation between temperature trends at the surface and those in the atmospheric layer just above the surface. The remaining geographical variables are mostly insignificant. In a regression of the surface trends just on the geographic variables (Model G1), the other variables besides $TROP_i$ are insignificant, the R^2 is 0.45 and the log-likelihood is 105.0. In the full model (*SURF*, column 1) the R^2 score rises to 0.53 and the log-likelihood rises to 139, indicating that there is a fraction of variability in the surface temperature data unexplained by the atmospheric temperature trend sub-model, for which the socioeconomic indicators provide significant explanatory power. The joint F test on the socioeconomic indicators is highly significant ($P = 7.1 \times 10^{-14}$).

It is noteworthy in the *SURF* column that population is significant and the coefficient is large. The IPCC gridded data are supposed to have been pre-filtered for the influence of population growth, and if this contamination had successfully been removed the coefficient would be zero. The coefficient size, if extrapolated linearly, indicates that a 100% increase in population (p_i = 1.00) would add 0.38 °C/ decade to the observed trend in a gridcell. Model G2 introduces population growth as the only nonclimatic factor. It is positive and significant, though smaller, but the overall R^2 increases very little, indicating this variable alone has limited explanatory power. Population growth alone cannot explain the role of other economic factors, which need to be controlled separately.

In Model G3 the inhomogeneity factors g_i, e_i and x_i are introduced, and two of the three are significant. Increased educational attainment (e_i) is associated with less measured warming (identical to the finding in McKitrick and Michaels 2004), while higher GDP density (g_i) is associated with more measured warming (the effect was insignificant in McKitrick and Michaels 2004). Adding in a squared GDP density variable did not improve the model: both g_i and its square became insignificant. Missing data counts (x_i) had a slightly positive effect but the effect is insignificant. In Model G4 the surface process measures are introduced and all four are significant. Positive GDP growth is associated with lower measured warming trends; population, income and coal use add to warming trends.

The surface process measures are clearly significant. The negative coefficient on GDP growth (y_i) suggests that an increase in GDP is associated with a cooling trend. This is consistent with the findings of Feddema et al. (2005a,b) regarding overall land-surface modification since industrialization. However y_i cannot be interpreted on its own in this model because of the way it interacts with average income m_i and population p_i. If y_i is used on its own (as is p_i in G2) then the coefficient becomes positive (0.037) and significant (t = 2.82). GDP growth is defined as y_i=GDP(1999)/GDP(1979)-1, and similarly for m_i and p_i. Since income is just GDP/population the three variables factorize as $(1+y_i)=(1+p_i)(1+m_i)$, implying $m_i + p_i + y_i = 2m_i + 2p_i + m_i p_i$. If y_i is replaced by (m_i p_i) the estimated coefficient is identical, though the coefficients on m_i and p_i change (not shown). Hence the surface process coefficients should be examined jointly, and the individual effects should be interpreted with some care.

The p-values for joint hypothesis tests (using standard F statistics) are listed in the bottom rows of Table 3. The test P(I) is the prob-value of the test that the inhomogeneity variables g_i, e_i and x_i are jointly zero. The test P(S) is the prob-value of the test that the four surface process growth rates are jointly zero, and P(all) tests whether all the nonclimatic factors (p_i through x_i) are jointly zero. Every entry indicates a significant rejection of the hypothesis, and the overall conclusion is unambiguously that the socioeconomic data have significant explanatory power on the spatial pattern of trends in the surface climate data. The hypothesis that the temperature data are independent of socioeconomic influences can be confidently rejected.

Multicollinearity can be a concern in a regression model with many explanatory variables, however the usual indication of its presence is a combination of insignificant coefficient t-statistics and significant joint F or model F scores. In our case the joint and model F scores are significant, but the socioeconomic variables are almost always individually significant as well. Hence if some variance inflation occurs due to partial correlations among regressors, it is not sufficient to obscure the basic results. Of the 78 correlation coefficients among regressors, 71 were less than 0.5. Only two were above 0.9, that between DRY_i and $DSLP_i$ ($\rho = 1.00$) and that between m_i and y_i ($\rho = 0.958$). The first pair is not important since they are part of a spline function and are identical by construction in the overlap segment. The second pair are related by factorization, as noted above, and the interpretation is primarily in their joint significance. The variance inflation factors of m_i and y_i were 110.8 and 124.9 respectively. For the remaining nine variables, the variance inflation factors were all less than 10.0; indeed eight were less than 5.0 and six were less than 2.0, indicating that the model has sufficient data to identify independent effects of the included regressors.

Another model run (G5) used all variables except *TROP*. The remaining geography variables became more significant, especially latitude, but the socioeconomic measures hardly changed and all the individual and joint hypothesis tests remained highly significant.

4.2 Influential Outliers

In this and the next few sections we consider tests of specification and endogeneity, to test whether the model is merely generating fluke correlations. We begin with a test for the role of influential outliers. The global model was re-run as follows. For each observation, the corresponding diagonal element of the OLS hat matrix was evaluated, and the observation removed if the value exceeded twice the mean of the hat matrix diagonal elements (Kmenta 1986, pp. 424-426). This resulted in removal of 29 observations, leaving a sample size of 411. There was no obvious spatial pattern to the 29 outliers (see Figure 1), though there is some indication of a cluster in the North Sea region. A comparison of summary statistics between the samples suggests that the outlier regions have much higher rates of missing data and relatively high growth in coal consumption. The coefficients of the model without outliers were quite similar to the *SURF* results in Table 2, though the growth in coal use was no longer significant. The vector of coefficients was compared to that of the Table 2 *SURF* results using a Hausman-type chi-squared statistic. The joint variance-covariance matrix was estimated and the model coefficients were compared, yielding a $\chi^2(14)$ score of 18.82, which is insignificant (P = 0.17), indicating that we do not reject the hypothesis that there are no systematic differences in the coefficients between the models with and without outliers. Consequently it is unlikely that the results in Table 2 are merely due to uncharacteristic outlier observations.

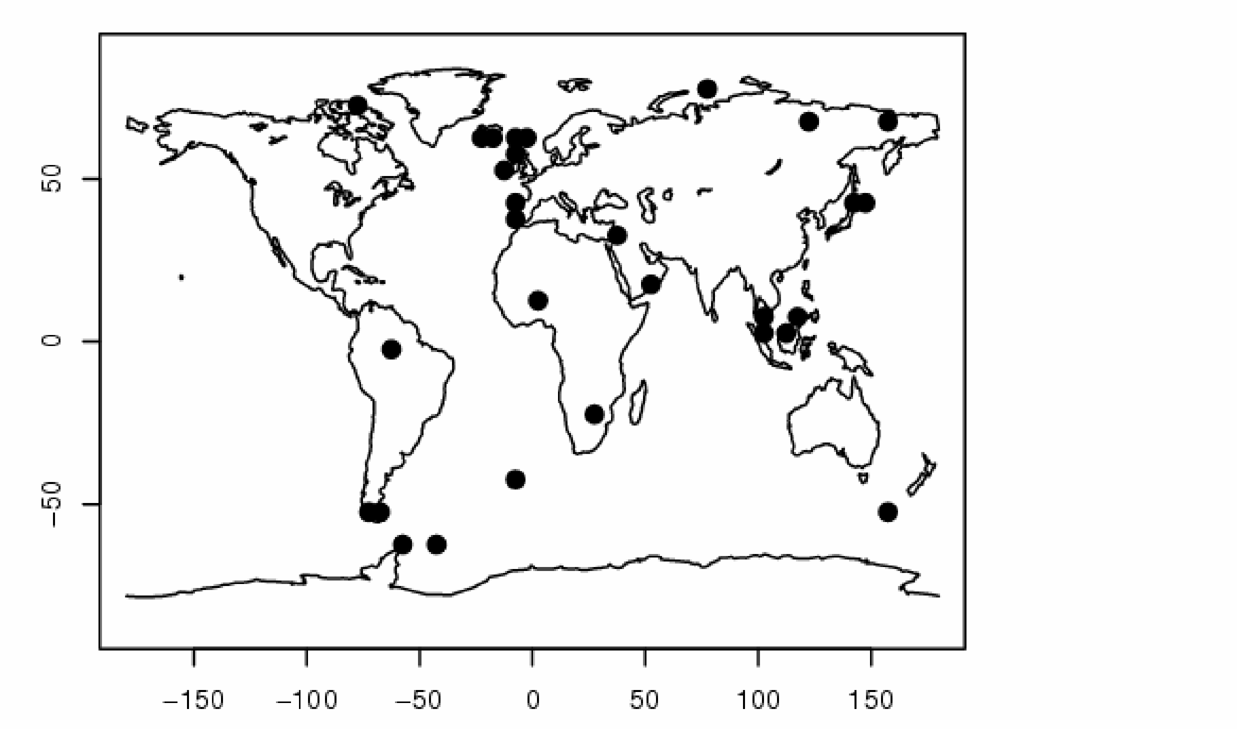

Figure 1. Global distribution of influential outliers removed from sample for calculation of results in Section 4.2.

4.3 Regression Error Specification Test Against General Nonlinear Alternative

A regression error specification (RESET) test was applied to check for biases due to unmodeled nonlinear structural components in the error term. The RESET test evaluates whether the dependent variable is a nonlinear function of the explanatory variables, in which case the linear model would be a misspecification. The test is run as follows. Predicted values $\hat{\theta}_i$ were obtained from fitting the *SURF* model and the regression was re-run using the same model augmented with $\hat{\theta}_i^2$ on the right-hand side. The *t*-statistic on $\hat{\theta}_i^2$ is an exact test of the null hypothesis that there is no nonlinear structure in the residuals, including any monotonic function of the right-hand side variables up to a quadratic (Davidson and MacKinnon 2004 pp. 653—655). The coefficient was -0.0225 and the *t*-statistic was -0.06 ($P = 0.956$), clearly failing to reject the null, indicating support for the linear model specification in (3).

4.4 Endogeneity

Endogeneity (also called simultaneity) bias arises in a regression model if the regressors are themselves partly determined by the value of the dependent variable. This implies that they are not orthogonal to the random error terms, violating the assumptions of classical linear regression and yielding biased and inconsistent coefficients. It could arise in this model if the right-hand side variables were not predetermined with respect to temperature trends, e.g. if economic agents were forward-looking with respect to climate change and adjusted productive activity in a region based on anticipated temperature changes. We find the concern about endogeneity implausible for three reasons.

First, as was noted by Schelling (1992), among others, very little economic activity in developed countries is affected by the weather. Agriculture, fishing and forestry are, but greenhouse warming does not involve predictions of uniformly deleterious outcomes (e.g. Mendelsohn et al. 2000), and in any case these sectors make up small fractions of the world's economies, typically less than five percent in developed countries.

Second, region-specific climate change prediction was not available in 1979 and is not even reliably available today. But suppose agents did have rational expectations and accurate forecasts as of 1979. Then we would expect to see the largest economic adjustments coinciding with the regions of the largest forecasted climate change from conventional global climate models, such as those used for the IPCC reports. But the pattern of large economic changes is uncorrelated with the regional pattern of predicted greenhouse warming (see de Laat and Maurellis 2006, Fig. 1). Consequently, it is highly unlikely that changes in economic activity can be explained by expectations of regional warming, since it occurs in places other than where the warming is expected.

Third, a Hausman test provides no support for a charge of inconsistency. A Hausman test compares two versions of the regression model, one in which the estimates are efficient but potentially inconsistent, and one in which the estimates are consistent but inefficient. "Consistency," in statistical terms, means that the expected value of an estimate converges to the true value as the sample size approaches infinity. "Efficiency" means that the estimated variance is the lowest among the class of unbiased estimators. The two vectors of coefficient estimates are compared, with the null hypothesis that there is no systematic difference between them. The variables that might be susceptible to endogeneity are the surface process measures (p_i, m_i, y_i, c_i). A Hausman test was implemented as follows. First, an efficient estimator was obtained using OLS on equation (3). Second, we regressed each surface process variable on pre-determined explanatory variables and obtained the model-predicted values:

$$\begin{aligned} q_i^j &= \alpha_0 + \alpha_1 g_i + \alpha_2 g_i^2 + \alpha_3 e_i + \alpha_4 e_i^2 + \alpha_5 g_i e_i + \alpha_6 x_i + \alpha_7 popden79_i + \alpha_8 GDPden79_i \\ &+ \alpha_9 coal80_i + \alpha_{10} GDP79_i + \alpha_{11} pop79_i + \alpha_{12} Soviet_i + \alpha_{13} slp_i + \alpha_{14} dry_i + \alpha_{14} slp_i \\ &+ \alpha_{15} water_i + \alpha_{16} abslat_i + v_i \qquad (7) \end{aligned}$$

where q_i^{1-4} represents the four surface variables, $popden79_i$ is 1979 population density, $GDPden79_i$ is 1979 GDP density, $coal80_i$ is 1980 national coal consumption, $GDP79_i$ is 1979 national GDP, $pop79_i$ is 1979 total population, $Soviet_i$ is a dummy variable for membership in the former Soviet Union and v_i is a regression residual. Since the righthand side variables are all predetermined as of 1979, and all temperature data (either surface or tropospheric) is left out of (7), the OLS predicted values $\hat{q}_i^{1-4}$ are strictly exogenous with respect to post-1979 temperature. Third, equation (3) was re-run with (p_i, m_i, y_i, c_i) replaced by $\hat{q}_i^{1-4}$, obtaining consistent estimators. The variance-covariance matrix comparing the efficient and consistent estimators was obtained and the Hausman $\chi^2(14)$ score was 3.83, which has a P-value of 0.9964, indicating no grounds whatsoever for finding a difference between the efficient and consistent estimators. Consequently, on both *ex ante* and *ex post* grounds we can rule out endogenous temperature effects on the right-hand side of (3) as the explanation of our main results.

4.5 Out of Sample Prediction

A rigorous test of a regression model is its ability to predict data not included in the estimation. This is especially useful for testing whether both dependent and independent variables are jointly determined by omitted "third" factors, resulting in fluke regression coefficients. Our test consisted of randomly removing 30% of the observations, then running the regression (3) on the remaining 70% of the grid cells and using the resulting coefficients to predict the withheld sample, yielding the vector $SUR\hat{F}_i$. Then, for the 30% withheld sample, the actual grid cell trends were regressed on the predicted trends. A perfect predictor would yield a 45° line (zero intercept,

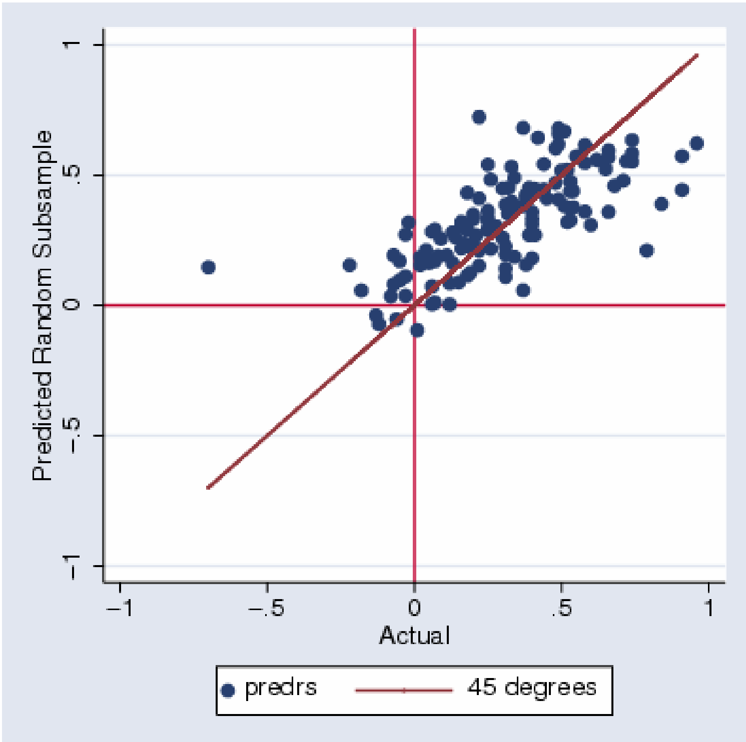

Figure 2. Horizontal axis: observed values of temperature trends in randomly-selected subset of data comprising 30% of original data set. Vertical axis: predicted values for same locations from regression equation (3) applied to global data with randomly-selected 30% of data withheld. Line shown is 45^{o} (not regression fit).

unit slope) between predicted and actual observations. In 500 repetitions the constant term was typically near zero (mean approximately 0.01), the slope coefficient was typically near one (mean approximately 0.96) and the R^2 indicated a high level of explained variance (mean approximately 0.50). An F test that the regression is a 45 ^{o} degree line was consistently not rejected (mean P approximately 0.37). An example is shown in Figure 2. Consequently, in repeated tests of out-of-sample prediction, the closeness of a scatter of predicted and actual data to a 45^{o} line and the high significance level of the predictions, gives us confidence that equation (3) is a valid empirical model.

4.6 Tropospheric Model

If the surface regression results are simply spurious spatial correlations based on a coincidental similarity to the spatial pattern of the Earth's general atmospheric circulation, then we would find the same right hand side coefficients and significance levels if the dependent variable $SURF_i$ were replaced with $TROP_i$. But if the surface processes are genuine effects they ought to be substantially weaker in the tropospheric data compared to the surface (though not necessarily zero). The results of this regression are in Table 3.

The surface processes (p_i, m_i, y_i, c_i) are, as expected, smaller and in three cases insignificant. The coal use effect is about half the size but remains significant. The four measures are jointly insignificant (P = 0.2114). The inhomogeneity measures show an ambiguous change. Educational attainment vanishes, GDP density retains its size (though not its significance) and x_i, the missing data score, becomes larger and significant. Obviously, problems in measuring surface data would not affect satellite records, so this indicates that x_i is serving as a proxy for something else that would have an atmospheric interpretation, in the five percent of grid cells with missing data. In particular, greater rates of missing data at the surface correlate with lower temperature trends aloft. We regressed x_i on the five variables *dry* through *abslat*, and there was an interesting contrast between the moist and dry regions. The results were

$$\hat{x}_i = 135.02 - 0.132 \times slp_i + (.) \qquad \text{(moist regions)}$$

$$\hat{x}_i = -9.97 + 0.0.010 \times slp_i + (.) \qquad \text{(dry regions)}$$

where (.) denotes the water and abslat variables (which are both significant). In the moist regions, missing data is less prevalent the higher the air pressure and the relationship is significant

(P=0.003). In dry regions the relationship vanishes and the pressure coefficient is insignificant (P=0.726). Low air pressure is associated with the more storm-prone regions, hence x_i may be acting as a marker for moist, storm-prone regions. In that case the negative coefficient on x_i in the second column of Table 3 may be a spurious effect reflecting the fact that air temperature trends are smaller over moist, unstable regions such as the tropics. However, to the extent that this raises a question about the interpretation of x_i in Table 2 it doesn't matter much, since at the global level the variable is insignificant in all specifications (but see next section).

5. Economic Subsamples

Some further detail emerges in subsamples defined on economic grounds. The sample was split into rich and poor locations, and growing/declining groups. 'Rich' was defined as having above-median income based on the 1999 data (217 of 440 observations); 'growing' was defined as 1999 real per capita income exceeding that in 1979 (335 of 440 observations). The results are in Table 4.

When divided into rich and poor subgroups, a distinction emerges between inhomogeneity effects and surface processes, whereby the former are uniformly significant across income groups while the latter become uniformly insignificant. It is noteworthy that in the global sample, missing data is insignificant, but this masks significant, contrasting effects between rich and poor regions. However, as indicated in Section 4.6, x_i is potentially confounded with a local climate extremity.

The disappearance of the surface process effects within subgroups suggests that they have a step-function-like character, such that the significant effects in the pooled global sample derive from differences between rich and poor groups, whereas the within-group effects appear to be insignificant.

An even sharper contrast emerges between growing and declining economies. In growing regions (76% of the sample) the inhomogeneity effects are close to those in the global sample, while three of the four surface process effects are roughly double the size of those in the global sample, and all are significant, individually and jointly. But in declining economies, neither inhomogeneity nor surface process effects exert significant effects on the temperature trends. The absence of inhomogeneity effects is somewhat unexpected, though the overall GDP density effect (g_i) is larger and nearly significant in the declining region. The disappearance of the education effect may indicate that constraints on human capital are offset by the low opportunity cost of labour of all kinds (including skilled labour) during periods of decline.

Overall, the global results appear to be particularly associated with growing economies. Since only 24 percent of the grid cells are in countries that experienced real declines in income, the growth effects are sufficiently widespread to affect the global results. While the relative strength of surface process effects in growing countries accords well with intuition, the relatively strength of inhomogeneity effects in rich countries compared to poor countries does not. We would have expected resource constraints to have stronger effects in poor countries, though we note that the inhomogeneity effects are jointly significant in both regions.

6. Identifying Nonclimatic trends

Focusing on the results at the global level, we can reject the hypothesis that adjustments to climatic data are successful in removing the extraneous influences of socioeconomic conditions in the regions of origin. While it is not possible to use the coefficients from (3) to identify the vector of 'true' climatic trends T_i, it is possible to try and simulate ideal climatic measurement conditions.

Peterson (2003) shows that US data can, in principle, be adjusted to remove extraneous biases of significant size. On this basis we postulate that countries with public sector resources and general public skill levels comparable to those in the US would be, in principle, able to provide uncontaminated climatic data. We therefore generated an adjusted vector of predicted values $\hat{\theta}_i^{ADJ}$ under the assumptions that all countries have GDP density and educational levels equivalent to those in the USA and that all other surface and inhomogeneity effects were set equal to zero:

$$\hat{\theta}_i^{ADJ} = \hat{\beta}_0 + \hat{\beta}_1 TROP_i + \hat{\beta}_2 PRESS_i + \hat{\beta}_3 DRY_i + \hat{\beta}_4 DSLP_i + \hat{\beta}_5 WATER_i + \hat{\beta}_6 ABSLAT_i$$
$$+ \hat{\beta}_{11} \times 144.2 + \hat{\beta}_{12} \times 0.36762 \qquad (8)$$

The resulting average temperature trend using (8) is 0.17 °C/decade, a drop of just under one-half of the observed sample average grid cell trend of 0.30 °C/decade, and below the MSU average of 0.23 °C/decade. If the data are weighted by relative grid cell size (using the cosine of latitude) the effect is a bit larger. The weighted average grid cell trend is 0.27 °C/decade, the weighted MSU average is 0.20 °C/decade and (8) yields a weighted average of 0.13°C/decade, a drop of just over one-half. Additionally, the sample density is lowest in regions like Africa and South America, the majority of whose grid cells show a warm bias. If these cells were weighted relatively more heavily to adjust for the extent of missing data, the drop in the global average trend would be even larger.

Frequency histograms are shown in Figure 3 The effect of removing the local distortions as estimated by the model is to bring the shape of the surface data distribution more closely into line with that of the satellite-measured lower troposphere data, primarily by removing the large upper tail. While we do not assert that the 'true' average land-based climatic warming trend is 0.17 °C/decade, our analysis does suggest that nonclimatic effects are present in the gridded temperature data used by the IPCC and that they likely add up to a net warming bias at the global level that may explain as much as half the observed land-based warming trend. This result mirrors that in McKitrick and Michaels 2004, as well as the findings in deLaat and Maurellis (2004, 2006) and Kalnay and Cai (2005), all of whom found the overall effect of surface processes to be a positive bias to observed temperature trends. Since this analysis takes the tropospheric record to be 'clean', whereas the results in Section 4.6 suggest that even the MSU series may reflect anthropogenic surface process effects, our findings should be viewed as a lower bound, or conservative estimate of the magnitude of the global data contamination.

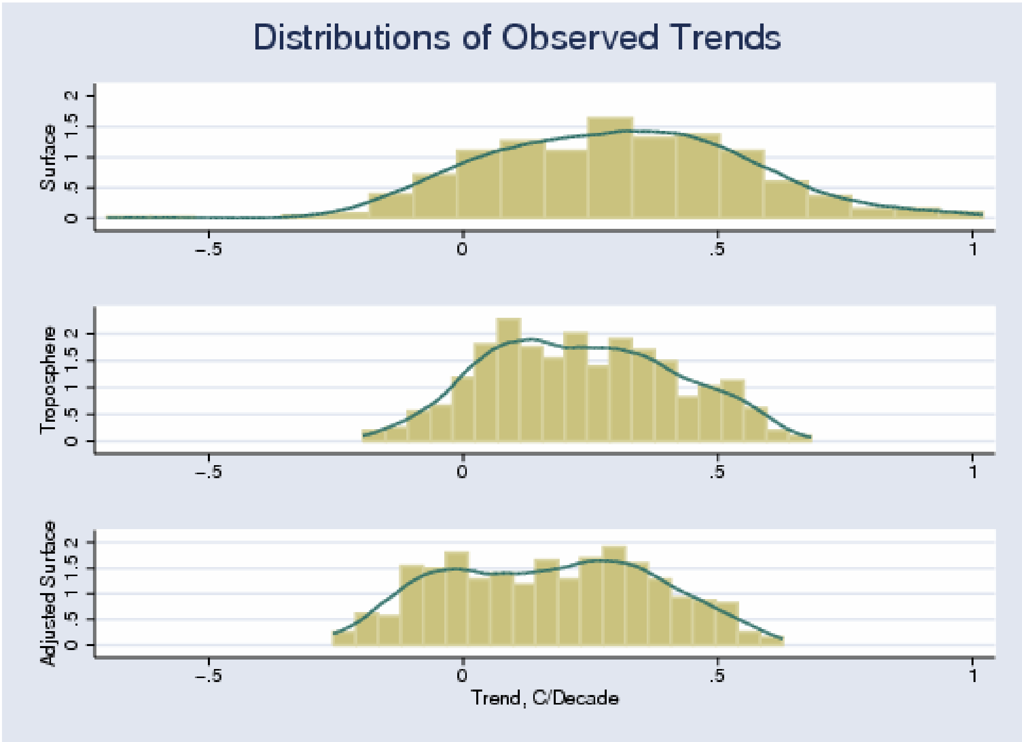

Figure 3: Distributions of temperature trends 1979-2002. Top: IPCC surface data. Middle: satellite (MSU) tropospheric data. Bottom: adjusted surface data. Smoothed kernel density shown.

The positive biases found here are not uniformly distributed around the world. Figure 4 shows the differences ($\theta_i - \hat{\theta}_i^{ADJ}$) on a global map. Note the regions where the adjustments are minimal are North America, Eastern Europe and Australia. Widespread positive biases are observed in Western Europe and Southeast Asia. Africa and South America contain many regions with missing data, though the map overstates this, because at the equator, the raster squares are smaller than the grid cells they represent, due to the global projection used.

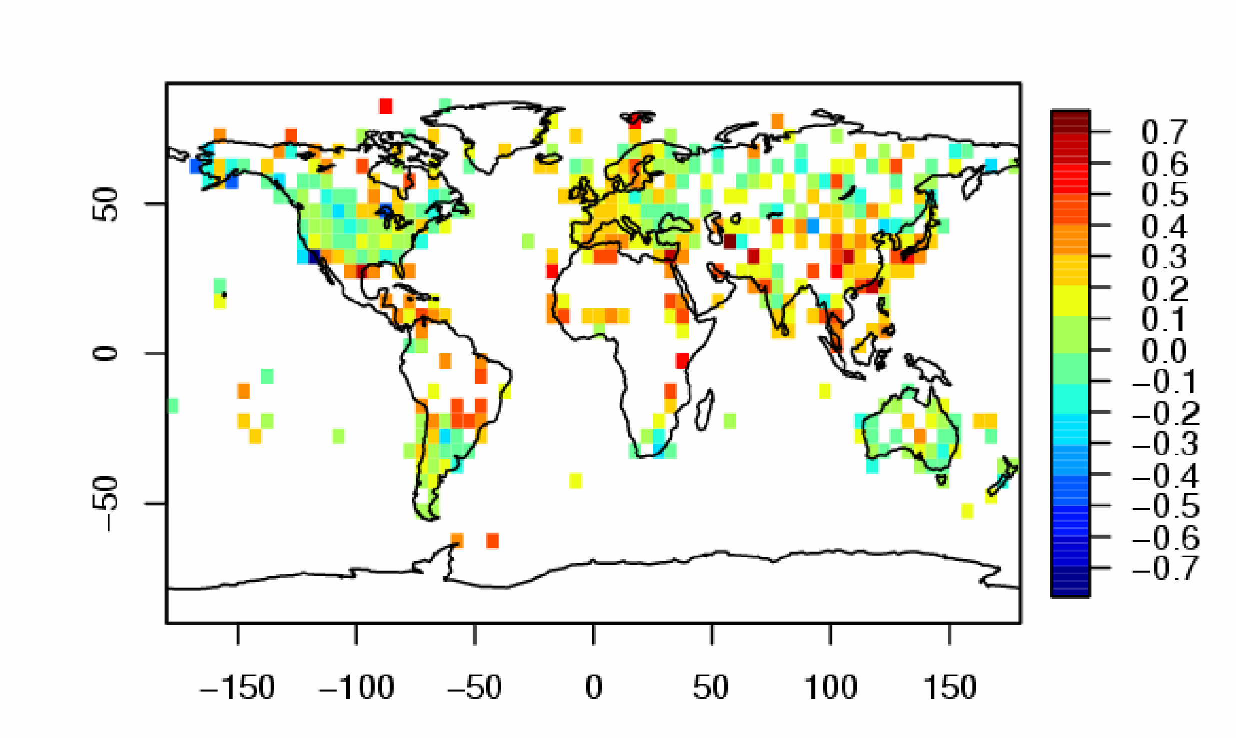

Figure 4. Differences between observed and adjusted trends around the world. Raster squares correspond to center of 5°x5° grid cell, but not to size of grid cell itself. Units are °C/decade. A value of, say, 0.1-0.2 means that the observed trend in that cell was between 0.1 and 0.2 °C/decade higher than the trend as adjusted using equation (8).

8 Conclusions

The standard interpretation of global climate data is that extraneous effects, such as urbanization and other land surface effects, and data quality problems due to inhomogeneities in the temperature series, are removed by adjustment algorithms, and therefore do not bias the large-scale trends. Our empirical model of the post-1980 interval embeds this assumption as a null hypothesis, and it is rejected at very high confidence levels. We show that our results cannot be explained away as outlier effects, model misspecification or reverse causality (endogeneity) bias. Out-of-sample prediction tests consistently perform well, and we show that variables representing changes in economic activity have significant explanatory power on the pattern of trends in published climatic data measured at the Earth's surface, but not in trends measured in the lower part of the atmosphere, thus showing that our results are not likely due to spurious correlation. Taken together, our findings show that trends in gridded climate data are, in part, driven by the varying socioeconomic characteristics of the regions of origin, implying a residual contamination remains even after adjustment algorithms have been applied. Users of gridded climate data products need to interpret their results accordingly.

These results are also consistent with previous findings showing that nonclimatic factors, such as those related to land use change and variations in data quality, likely add up to a net warming bias in climate data, suggesting an overstatement of the rate of global warming over land. They also provide support for attribution of some observed climate changes in recent decades to land surface modifications, rather than greenhouse gas emissions, a factor not typically evaluated in studies that attempt to attribute the causes of recent global warming.

Our data set has a low resolution for strictly local measures of economic density within countries. Since we detect significant effects on temperature trends even with low spatial resolution we conjecture that if future studies are able to examine the issues at the subnational level, even more significant and detailed results will emerge. There is some prospect for future subnational studies, possibly by merging the EDGAR data base (as used in de Laat and Maurellis 2004, 2006) with the new G-Econ data set (http://gecon.yale.edu/) developed by Nordhaus (2006). Additionally, there is always the possibility in cross-sectional regressions that unobservable heterogeneity may explain both climate and economic processes in such a way as to eliminate the significance of results reported herein. That could be formally tested in a panel data set where the time dimension provides additional identification of fixed cross-sectional effects, which is a direction for future research.

Acknowledgments

This research was assisted by financial support from the Social Sciences and Humanities Research Council of Canada. We thank Michael Fleischmann for research assistance and Stephen McIntyre for assistance with the figures.

References

Allen, M.R. and S.F.B. Tett (1999), Checking for Model Consistency in Optimal Fingerprinting, *Climate Dynamics* 15:419—434.

Baker, D.G. (1975), Effect of Observation Time on Mean Temperature Estimation, *Journal of Applied Meteorology* 14(4) 471—476.

Balling, R.C. and W.S. Hughes (1996), Urban influences on South African temperature trends, *International Journal of Climatology* 16:935—940.

Böhm, R. (1998) "Urban bias in temperature time series—A case study for the city of Vienna, Austria." *Climatic Change* 38:113-128.

Chagnon, S. A. (1999), A rare long record of deep soil temperature defines temporal temperature changes and an urban heat island, *Climatic Change* 42: 531—538.

Chung, U., Choi, J. and J.I. Yun (2004), Urbanization effect on the observed change in mean monthly temperatures between 1951—1980 and 1971—2000 in Korea, *Climatic Change* 66:127—136.

Davey, C.A., and R.A. Pielke Sr. (2005), Microclimate exposures of surface-based weather stations - Implications for the assessment of long-term temperature trends, *Bulletin of the American Meteorological Society* 86(4) 497–504

Davidson, R. and J.G. MacKinnon (2004), *Econometric Theory and Methods* Toronto: Oxford.

De Laat, A.T.J., and A.N. Maurellis (2004), Industrial CO_2 emissions as a proxy for anthropogenic influence on lower tropospheric temperature trends, *Geophys. Res. Lett.*Vol. 31, L05204, doi:10.1029/2003GL019024.

De Laat, A.T.J., and A.N. Maurellis (2006), Evidence for influence of anthropogenic surface processes on lower tropospheric and surface temperature trends, *International Journal of Climatology* 26:897—913.

Easterling, D.R., B. Horton, P.D. Jones, T.C. Peterson, T.R. Karl, D.E. Parker, M.J. Salinger, V. Razuvayev, N. Plummer, P. Jamason and C.K. Folland, 1997: Maximum and minimum temperature trends for the globe. Science, **277**, 364-367.

Easterly W. and M. Sewadeh (2003) World Bank Global Development Network Growth Data Base http://www.worldbank.org/research/growth/GDNdata.htm. Accessed fall 2003.

Feddema J.J., K.W. Oleson, G.B. Bonan, L.O. Mearns, L.E. Buja, G.A. Meehl, W.M. Washington (2005a), The importance of land-cover change in simulating future climates *Science* 310 9 December 2005 pp 1674-1678.

Feddema J.J., K.W. Oleson, G.B. Bonan, L. Mearns, W.M. Washington, G. Meehl, D. Nychka (2005b), A comparison of a GCM response to historical anthropogenic land cover change and model sensitivity to uncertainty in present-day land cover representations. *Climate Dynamics* 25 (6) 581—609.

Fujibe, F. (1995). Temperature rising trends at Japanese cities during the last hundred years and their relationships with population, population increasing rates and daily temperature ranges, *Papers in Meteorology and Geophysics* 46(2):35—55.

Gallo, K.P. and T.W. Owen (2002), A sampling strategy for satellite sensor-based assessments of the urban heat-island bias." *International Journal of Remote Sensing*." 23(9) 1935—1939.

Global Hydrology and Climate Center (GHCC) (2005) Web site http://wwwghcc.msfc.nasa.gov/temperature/. Global temperature anomalies. Accessed June 2005. Data are now archived at http://www.nsstc.uah.edu/data/msu/t2lt/.

Hansen, J., R. Ruedy, M. Sato, M. Imhoff, W. Lawrence, D. Easterling, T. Peterson and T. Karl (2001), "A Closer Look at United States and Global Surface Temperature Change." *J. Geophys. Res.*. 106:239-247.

Hingane, L.S. (1996), Is a signature of socioeconomic impact written on the climate? *Climatic Change* 32:91—102.

Intergovernmental Panel on Climate Change. (2001), *Climate Change 2001: The Scientific Basis*. Houghton, J.T., Y. Ding, D.J. Griggs, M. Noguer, P.J. van der Linden, X. Dai, K. Maskell and C.A. Johnson CA., eds. Cambridge: Cambridge University Press.

Intergovernmental Panel on Climate Change (2007), *Summary for Policymakers of the Contribution of Working Group I to the Fourth Assessment Report*. http://www.ipcc.ch/.

Jenne, R.L. (1974), Jenne's northern hemisphere climatology, monthly, 1950-64. National Center for Atmospheric Research Dataset DS205.0, National Center for Atmospheric Research, Boulder CO.

Jones, P.D. and A. Moberg (2003), Hemispheric and large-scale surface air temperature variations: An extensive revision and an update to 2001, *Journal of Climate*. 16:206-223.

Jones, P.D., P. Ya. Groisman, M. Coughlan, N. Plummer, W-C. Wang and T.R. Karl, (1990), Assessment of urbanization effects in time series of surface air temperature over land, *Nature* 347 169—172.

Kalnay E and M. Cai (2003), Impacts of Urbanization and Land Use Change on Climate, *Nature*. 423:528-531.

Karaka, M., M. Tayanc and H. Toros (1995), Effects of urbanization on climate of Istanbul and Ankara, *Atmospheric Environment*, 29(23) 3411—3421.

Karl, T.R. and C.N. Williams Jr. (1987), An approach to adjusting climatological time series for discontinuous inhomogeneities, *Journal of Climate and Applied Meteorology* 26 pp. 1744—1763.

Karl, T.R., H.F. Diaz and G. Kukla (1988), Urbanization: Its detection and effect in the United States climate record, *Journal of Climate*, 1:1099-1123.

Klysik, K. and K. Fortuniak, (1999), Temporal and spatial characteristics of the urban heat island of Lodz , Poland. *Atmospheric Environment* 33:3885—3895.

Kmenta, Jan (1986). *Elements of Econometrics*. New York: MacMillan.

Linacre, E. (1992) *Climate data and resources: A reference and guide*. Routledge, London.

Magee N, J. Curtis and N. Wendler (1999), The urban heat island effect at Fairbanks, Alaska, *Theoretical and Applied Climatology* 64:39—47.

McKendry, I. G. (2003), Progress Report: Applied Climatology, *Progress in Physical Geography* 27(4) pp. 597–606

McKitrick, R.R. and P. J. Michaels (2004), A test of corrections for extraneous signals in gridded surface temperature data, *Climate Research* 26(2) pp. 159-173, Erratum, *Climate Research* 27(3) 265—268.

Mendelsohn, R., W. Morrison, M. E. Schlesinger and N. G. Andronova (2000), Country-specific market impacts of climate change, *Climatic Change* 45(3-4): 553-569.

Michaels, P.J, P.C. Knappenberger, R.C. Balling Jr, and R.E. Davis (2000), Observed warming in cold anticyclones, *Climate Research*. 14:1-6.

Mitchell Jr., J.M. (1953), On the causes of instrumentally-observed secular temperature trends, *Journal of Meteorology* 10, pp. 244-261.

Mitchell, T.D. and P.D. Jones (2005), An improved method of constructing a database of monthly climate observations and associated high resolution grids, *International Journal of Climatology* 25: 693—712.

Nordhaus, W. (2006), Geography and macroeconomics: New data and new findings, *Proceedings of the National Academy of Sciences (US)*, March 7, 2006, vol. 103, no. 10, pp. 3510-3517.

Parker, D.E. (2004), Climate: Large-scale warming is not urban, *Nature* 432, 290 (18 November 2004); doi:10.1038/432290a.

Peterson T.C. and R.S. Vose (1997), An overview of the Global Historical Climatology Network temperature database, *Bulletin of the American Meteorological Society* 78:2837—2849.

Peterson, T.C. (2003), Assessment of urban versus rural *in situ* surface temperatures in the contiguous United States: No difference found, *Journal of Climate* 16(18) 2941—2959.

Pielke R.A. Sr., G. Marland, R.A. Betts, T.N. Chase, J.L. Eastman, J.O. Niles, D.D.S. Niyogi and S.W. Running (2002), The influence of land-use change and landscape dynamics on the climate system: Relevance to climate-change policy beyond the radiative effect of greenhouse gases. *Philosophical Transactions of the Royal Society of London*. A360:1705-1719

Pielke, R.A. Sr., T. N. Chase and B. Herman (2004), Some thoughts on the surface-mid-tropospheric temperature trend differences, Unpublished Technical Comment, University of Colorado, http://climatesci.colorado.edu/publications/pdf/naturecomments.pdf.

Pielke, R.A. Sr. and T. Matsui (2005), Should light wind and windy nights have the same temperature trends at individual levels even if the boundary layer averaged heat content change is the same?" *Geophys. Res. Lett.*(32) L21813, doi:10.1029/2005GL024407.

STATA Release 8.0 (2003), *Users Guide*. College Station: Stata Corporation.

Schaal, L.A. and R.F. Dale (1977), Time of observation bias and climatic change, *Journal of Applied Meteorology* 16 (3), 215—222.

Schelling, T. (1992), Some economics of global warming, *American Economic Review* 82(1): 1-14.

Spencer, R.W. and J.C. Christy (1990), Precise monitoring of global temperature trends from satellites, *Science* 247:1558—1562.

Staley, D.O. and G.M. Jurica (1970), Flux emissivity tables for water vapor, carbon dioxide and ozone, *Journal of Applied Meteorology*, 9:365-372.

Streutker, D.R. (2002), Satellite-measured growth of the urban heat island of Houston, Texas, *Remote Sensing of Environment* 85: 282—289.

Tett, S.F.B., P.A. Stott, M.R. Allen, W.J. Ingram and J.F.B. Mitchell (1999), Causes of twentieth century temperature change near the Earth's surface, *Nature*. 399:569-572.

Tso, C.P. (1995), A survey of urban heat island effects in two tropical cities, *Atmospheric Environment*, 30(3) 507—519.

United Nations Education, Scientific and Cultural Organization (UNESCO) (2003). Web site http://unescostat.unesco.org/en/stats/stats0.htm. World education indicators. Data accessed from http://www.uis.unesco.org/TEMPLATE/html/Exceltables/education/View_Table_dist_pop_edu_attainment.xls

White, Halbert. (1980). "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." Econometrica 48 817–38.

Wilmott, C.J., S.M. Robeson, J.J. Feddema (1991), Influence of spatially variable instrument networks on climatic averages, *Geophys. Res. Lett.*18(2) 2249—2251.

Var	Definition	Obs	Mean	Std. Dev.	Min	Max
Surf	Surface temperature trend (θ_i)	440	0.3015	0.2574	-0.7	1.02
Trop	Tropospheric temperature trend	440	0.2325	0.1838	-0.1969	0.6832
Press	Sea level air pressure	440	1016.2	5.3024	993	1029
Dry	Dummy for dry region	440	0.4614	0.4991	0	1
Dslp	Dry x Press	440	469.40	507.78	0	1029
Water	Grid cell contains a coast line	440	0.6045	0.4895	0	1
Abslat	Absolute latitude	440	40.602	17.953	2.5	82.5
g	1979 Real National GDP per sq km in millions	440	0.2965	0.5999	0.0014	3.0023
e	Literacy +Post-secondary education rates	440	106.52	26.200	11.6	144.2
x	# missing months in grid cell temperature record	440	0.7636	2.5522	0	24
p	% growth in population*	440	0.2792	0.2089	-0.0692	1.2353
m	% growth in real average income*	440	0.3799	0.6142	-0.7901	2.1472
y	% growth in real national GDP**	440	0.7710	0.8391	-0.6686	3.0025
c	% growth in coal consumption*	440	1.0158	4.0557	-1	39.333
Rich	1999 real income > median	440	0.4932	0.5005	0	1
Grow	1999 real income > 1979 real income	440	0.7614	0.4267	0	1

Table 1: Model Variables. Definitions discussed further in text. *over the interval 1979 to 1999. **Over the interval 1980 to 2000. % Changes should be multiplied by 100, e.g. mean population growth is 27.92%.

Variable	*SURF*	G1	G2	G3	G4	G5
trop	**0.8631**	**0.9054**	**0.9195**	**0.8884**	**0.8855**	
	(8.62)	(10.28)	(9.73)	(8.94)	(8.89)	
slp	0.0044	-0.0012	0.0009	0.0041	-0.0006	0.0043
	(1.02)	(-0.22)	(0.16)	(0.92)	(-0.13)	(0.91)
dry	0.5704	-4.3301	-2.6643	1.5847	-4.8544	-9.2581
	(0.10)	(-0.59)	(-0.37)	(0.29)	(-0.70)	(-1.58)
dslp	-0.0005	0.0043	0.0027	-0.0015	0.0048	0.0092
	(-0.09)	(0.60)	(0.38)	(-0.27)	(0.71)	(1.61)
water	-0.0289	-0.0374	-0.0308	-0.0245	-0.0403	-0.0024
	(-1.37)	(-1.63)	(-1.37)	(-1.19)	(-1.73)	(-0.09)
abslat	0.0006	-0.0014	-0.0002	-0.0003	0.0004	**0.0061**
	(0.51)	(-1.63)	(-0.16)	(-0.29)	(0.38)	(3.39)
g	**0.0432**			**0.0480**		**0.0798**
	(3.36)			(3.81)		(3.15)
e	**-0.0027**			**-0.0028**		**-0.0030**
	(-5.14)			(-5.49)		(-4.26)
x	0.0041			0.0029		-0.0057
	(1.66)			(1.10)		(-1.52)
p	**0.3839**		**0.1798**		**0.4143**	**0.5432**
	(2.72)		(2.23)		(3.59)	(2.80)
m	**0.4093**				**0.3374**	**0.6334**
	(2.39)				(2.47)	(2.66)
y	**-0.3047**				**-0.2287**	**-0.4834**
	(-2.22)				(-2.17)	(-2.57)
c	**0.0062**				**0.0036**	**0.0093**
	(3.45)				(2.42)	(3.56)
constant	-4.2081	1.3425	-0.8378	-3.7889	0.6149	-4.1492
	(-0.96)	(0.24)	(-0.15)	(-0.84)	(0.13)	(-0.85)
N	440	440	440	440	440	440.00
R^2	0.53	0.45	0.46	0.51	0.48	0.34
ll	139.22	105.03	109.92	131.53	116.88	63.18
P(I)	**0.0000**			**0.0000**		**0.0000**
P(S)	**0.0004**				**0.0001**	**0.0022**
P(all)	**0.0000**					**0.0000**

TABLE 2: MAIN PARAMETER ESTIMATES. Coefficient t-statistics underneath in parentheses, based on robust standard errors. **Bold** denotes significant at 95%. Variable codes: g: 1979 GDP density; e: educational attainment; x: count of missing months; p: % change in population; m: %income growth; y: % growth in GDP; c: &growth in coal consumption. ll = loglikelihood value. P(I) = prob value of test that all inhomogeneity factors (g—x) are jointly zero; P(S) = prob value of test that all surface process coefficients (p—c) are jointly zero; P(all) = prob value of test that g—c are jointly zero.

Variable	*SURF*	*TROP*
trop	**0.8631**	
	(8.62)	
slp	0.0044	-0.0001
	(1.02)	(-0.03)
dry	0.5704	**-11.3879**
	(0.10)	(-3.01)
dslp	-0.0005	**0.0112**
	(-0.09)	(3.02)
water	-0.0289	0.0307
	(-1.37)	(1.37)
abslat	0.0006	**0.0064**
	(0.51)	(5.39)
g	**0.0432**	0.0424
	(3.36)	(1.81)
e	**-0.0027**	-0.0004
	(-5.14)	(-0.58)
x	0.0041	**-0.0114**
	(1.66)	(-3.22)
p	**0.3839**	0.1845
	(2.72)	(1.42)
m	**0.4093**	0.2596
	(2.39)	(1.55)
y	**-0.3047**	-0.2069
	(-2.22)	(-1.59)
c	**0.0062**	**0.0036**
	(3.45)	(2.11)
_cons	-4.2081	0.0682
	(-0.96)	(0.02)
N	440	440.00
R^2	0.53	0.49
ll	139.22	269.02

Table 3: Comparison of basic model and version with dependent variable surface trends replaced by tropospheric trends. **Bold** denotes significant at 95%. *T* statistics in parentheses.

Variable	Sglobe	Rich	Poor	Growing	Declining
trop	**0.8631**	**1.1224**	**0.6257**	**0.9378**	**0.6085**
	(8.62)	(8.59)	(4.52)	(8.20)	(2.47)
slp	0.0044	0.0084	0.0121	0.0043	-0.0017
	(1.02)	(1.57)	(1.33)	(1.20)	(-0.05)
dry	0.5704	6.4594	5.3143	4.4592	-12.1839
	(0.10)	(0.85)	(0.54)	(0.82)	(-0.33)
dslp	-0.0005	-0.0063	-0.0051	-0.0043	0.0121
	(-0.09)	(-0.84)	(-0.52)	(-0.80)	(0.33)
water	-0.0289	**-0.0350**	-0.0326	-0.0295	-0.0358
	(-1.37)	(-2.19)	(-1.11)	(-1.15)	(-0.96)
abslat	0.0006	0.0009	-0.0021	0.0002	0.0026
	(0.51)	(0.47)	(-1.32)	(0.14)	(0.85)
g	**0.0432**	**0.0517**	0.0614	**0.0385**	0.4325
	(3.36)	(3.27)	(0.47)	(3.03)	(1.91)
e	**-0.0027**	**-0.0047**	**-0.0018**	**-0.0026**	-0.0029
	(-5.14)	(-6.79)	(-2.02)	(-4.54)	(-1.27)
x	0.0041	**-0.0066**	**0.0053**	0.0044	0.0002
	(1.66)	(-2.29)	(2.88)	(0.99)	(0.05)
p	**0.3839**	0.8761	0.2554	**0.8867**	-0.0002
	(2.72)	(1.40)	(1.71)	(4.22)	(-0.01)
m	**0.4093**	0.5398	0.2659	**0.8687**	-0.2361
	(2.39)	(1.04)	(1.48)	(3.87)	(-0.59)
y	**-0.3047**	-0.4365	-0.2003	**-0.6453**	0.2913
	(-2.22)	(-1.02)	(-1.41)	(-3.62)	(1.12)
c	**0.0062**	0.0043	0.0022	**0.0075**	-0.0105
	(3.45)	(1.07)	(0.50)	(3.49)	(-0.42)
_cons	-4.2081	-8.1950	-11.9809	-4.1058	1.9676
	(-0.96)	(-1.49)	(-1.30)	(-1.14)	(0.06)
N	440	217	223	335	105
R^2	0.53	0.63	0.47	0.56	0.44
ll	139.22	75.61	79.80	107.13	40.38
P(I)	**0.0000**	**0.0000**	**0.0078**	**0.0000**	0.2372
P(S)	**0.0005**	0.2775	0.3282	**0.0000**	0.1623
P(all)	**0.0000**	**0.0000**	**0.0009**	**0.0000**	0.1030

Table 4: Rich/poor, Growing/declining subsamples. Column 1 shows Sglobe results from Table 2 for comparison. Variable codes as for Table 2. P(I) = prob value of test that all inhomogeneity factors (g—x) are jointly zero; P(S) = prob value of test that all surface process coefficients (p—c) are jointly zero; P(all) = prob value of test that g—c are jointly zero. **Bold** denotes significant at 95%. *T* statistics in parentheses.