ENCOMPASSING TESTS OF SOCIOECONOMIC SIGNALS IN SURFACE CLIMATE DATA

Ross McKitrick
Department of Economics
University of Guelph

In Press: Climatic Change

Checked Against Proofs: June 3, 2013

Abstract: The debate over whether urbanization and related socioeconomic developments affect large-scale surface climate trends is stalemated with incommensurable arguments. Each side can appeal to supporting evidence based on statistical models that do not overlap, yielding inferences that merely conflict but do not refute one another. I argue that such debates are only be resolved in an *encompassing* framework, in which both types of results can be demonstrated as restricted forms of the same statistical model, and the restrictions can be tested. The issues under debate make such data sets challenging to construct, but I give two illustrative examples. First, insignificant differences in warming trends in urban temperature data during windy and calm conditions are shown in a restricted model whose general form shows temperature data to be strongly affected by local population growth. Second, an apparent equivalence between trends in a data set stratified by a static measure of urbanization is shown to be a restricted finding in a model whose general form indicates significant influence of local socioeconomic development on temperatures.

Keywords: Urban heat islands, Temperature data quality, Wind, Encompassing tests.

ENCOMPASSING TESTS OF SOCIOECONOMIC SIGNALS IN SURFACE CLIMATE DATA

1 INTRODUCTION

The debate over whether urbanization and land use changes affect large-scale surface climate trends has involved a number of contrasting methodologies. Jones et al. (1990), Peterson et al. (1998) and Hansen et al. (1999), among others, stratified surface data sets according to measures of urbanization levels and used the fact that the trend differences were insignificant to conclude that non-climatic bias was likewise insignificant. McKitrick and Michaels (2004a,b), McKitrick (2010) and McKitrick and Nierenberg (2010), by contrast, found significant correlations between the spatial pattern of warming trends and the spatial pattern of socioeconomic development, despite such a pattern not being a prediction of climate models (contrast Schmidt 2009 with McKitrick and Nierenberg 2010 on this point). de Laat and Maurellis (2004, 2006) showed that measured trends were higher in regions with high industrialization trends (as measured by local carbon dioxide emissions), irrespective of where they set the high/low industrialization threshold, with the largest gap appearing when the threshold was set fairly high. By contrast Parker (2004, 2006) examined a sample of urban locations and found no difference in trends between subsets partitioned according to nighttime wind speed, concluding on this basis that urban warming could not be a significant factor in global averages. More recently, Wickham et al. (2013) divided the "BEST" data set (Berkeley Earth Surface Temperature) using satellite-based measures of rural and urban locations and found no significant difference in average trends, likewise concluding that land surface disruptions were not a factor in global average trends.

These rival statistical arguments are inconclusive because, while each side is using the same (or comparable) temperature data, they use different tests and different statistical models. One result merely contrasts with another, but does not disprove it. Such debates are interminable unless an *encompassing* framework is created, in which both types of results can be demonstrated on a single data set, in such a way that apparent support for one is shown only to occur as a restricted case of a more general specification that supports the other.

The basic concept of the encompassing framework is explained in the next section. In sections 3 and 4 I apply the framework on two examples. In section 3, a panel of Canadian urban and rural post-1979 weather station data are developed which includes both local wind and population figures. A Parker-type result is reproduced, in which near-identical warming trends are demonstrated in urban data during windy and calm conditions. This is shown in a restricted model, the general form of which shows that regional population growth has a strong apparent warming effect in urban areas but not rural areas, notwithstanding the absence of a wind effect. In section 4 the global sample of 1979-2002 surface trends from McKitrick and Nierenberg (2010) is compared to a suite of regional socioeconomic indicators. A BEST-type equivalence is demonstrated between trends in locations defined as urban based on a static measure (in other words, stratification of surface characteristics observed at one point in time). But it is also shown that this result is, in principle, consistent with the absence *or* the presence of urbanization bias, and consequently is uninformative. A general model is estimated showing that significant socioeconomic patterns can be detected even if a restricted form of the model, in which the sample is split based on a static characteristic, fails to show it.

The encompassing approach is especially useful in cases in which certain results depend on the failure to observe an effect, since this is not proof that the effect does not exist. In the case of Parker (2004, 2006), it is argued that when data are unaffected by urban heat islands (UHI), night time minimum trends will be the same under calm and windy conditions. But this does not demonstrate the reverse, namely that an insignificant difference in trends between calm and windy conditions implies the data are unaffected by land surface disruptions. Likewise Wickham et al. argue that in data unaffected by urbanization, trends will be equivalent between urban and rural locations. But again, it is incorrect logic to argue the reverse, namely that the observation of trend equivalence implies the data are unaffected by urbanization. In each case, estimation of a general encompassing form allows for a more decisive treatment of the underlying dispute.

2 The Encompassing Framework

Consider two rival regression models:

$$y = \mathbf{X}\beta + e_1 \tag{1}$$

$$y = \mathbf{Z}\gamma + e_2 \tag{2}$$

where y is the dependent variable of interest (for example, the spatial pattern of temperature trends), $\mathbf{X}$ is a matrix of explanatory variables including those unique to one theory and $\mathbf{Z}$ is a matrix containing explanatory variables unique to the rival theory. For example, suppose the "Model 1" school includes only local windspeed in $\mathbf{X}$, interpreting this as a measure of urbanization. If $\hat{\beta} = 0$ this would be adduced as evidence that urbanization does not affect y. The "Model 2" school instead uses local population growth in $\mathbf{Z}$ and finds $\hat{\gamma} > 0$. At this point, since the models are non-overlapping, the findings merely conflict, but one does not disprove the other.

The principle of encompassing (Mizon 1984) states that if equation (1) is the true model, it ought not only to explain y in terms of $\mathbf{X}$, but any correlation between $\mathbf{X}$ and $\mathbf{Z}$ should account for the apparent explanatory power of $\mathbf{Z}$ in (2). Ordinary least squares (OLS) estimation of equation (2) yields $\hat{\gamma} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'y$. But if equation (1) is true then in the limit y can be replaced by $\mathbf{X}\beta$, so the limiting value of $\hat{\gamma}$ is $(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X}\beta$. Substituting in the OLS estimator of $\hat{\beta}$ we obtain $\tilde{\gamma}$, the estimate of γ from (2) under the assumption that (1) is correct:

$$\tilde{\gamma} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'y = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{P}_{\mathrm{X}}y$$

where $\mathbf{P}_{\mathrm{X}}$ denotes the OLS projection matrix, hence $\mathbf{P}_{\mathrm{X}}y$ is the vector of fitted values from a regression of y on $\mathbf{X}$. If the differences between the estimators $(\hat{\gamma} - \tilde{\gamma})$ are statistically insignificant, then model (1) can account for the explanatory power of model (2) and is therefore said to *encompass* it.

In linear models this test turns out to be equivalent to the F test on $\gamma = 0$ from the regression

$$y = \mathbf{X}\beta + \mathbf{Z}\gamma + e \qquad (3)$$

(Davidson and MacKinnon 2004, p. 672). Both models (1) and (2) may include other explanatory variables in common, which we denote **G**. In this case if we estimate

$$y = \mathbf{X}\beta + \mathbf{Z}\gamma + \mathbf{G}\mu + e \qquad (4)$$

then the F test on $\gamma = 0$ from equation (4) tests whether **Z** contributes significant explanatory power to the model, and equivalently whether the variables in **Z** are *encompassed* by those in **X** and **G**.

To make the setup more concrete, suppose one school argues that temperatures y are explained by greenhouse gases **G** and a measure of urbanization **Z**, whereas another school argues for **G** alone, after arguing that **X** is a sufficient measure of urbanization and finding it is insignificant. Estimation of (4) allows us to resolve this debate by asking whether **Z** contributes unique explanatory power ($\hat{\gamma} \neq 0$) in a model also containing **G** and **X**, which would imply it is not encompassed by them, and equivalently whether a model like (4) has significantly more explanatory power than one without **Z**. Only by creating an encompassing framework can the stalemate between equations (1) and (2) be resolved.

Developing an encompassing model tends to require a larger data set since more coefficients need to be estimated. However, one should not assume that a richer explanatory model necessarily implies less significant estimates. By adding, say, **G** and **Z** to a model like (1), the coefficient on **X** may turn out to become larger and more significant than before. Multiple regression fits each coefficient on the assumption that effects due to other variables are controlled. Introducing controls into a model where they belonged, but were previously missing, can reveal effects hidden by omitted variable bias.

3 Wind and the Urban Heat Island Effect

3.1 Data

For this example it was necessary to find suitable measures of temperature, wind and economic growth for a group of locations at a high sampling frequency. Building a global sample was not possible, but a pan-Canadian sample was developed that covers a large geographical span.

All temperature, wind speed and precipitation data were obtained from Environment Canada. The source URLs are shown in the caption to Table 1. Sampling locations for the urban data were international airports at the following cities: Victoria, Vancouver, Edmonton, Calgary, Regina, Saskatoon, Winnipeg, Toronto, Ottawa and Montreal. The non-homogeneity adjusted data are only subject to minimal checks for quality control but are otherwise unadjusted (Environment Canada, pers. comm.). Consequently these data can reasonably be expected to be affected by urbanization and regional land surface disruption. For each city I obtained the monthly minimum temperatures. Environment Canada also provides a homogeneity-adjusted data set with corrections applied for discontinuities due to time of observation change, equipment change, station moves and changes in exposure, but not for artificial trends due to regional land surface changes (Vincent et al. 2002). Monthly mean wind speed observations were taken from this archive, although wind speeds for nearby cities had to be used in place of Toronto, Montreal and Vancouver (respectively St. Catherines, McTavish and Abbotsford, see Table 1) due to the absence of those cities in the homogeneity-adjusted archive. Since the current debates have focused on the post-1979 interval this was the period studied in this example as well. The sample extended up to 2006, the last year of the wind speed record.

To build a data set that could reasonably be described as free of UHI problems, remote stations from the same province were selected to replace each city in turn. In some cases the stations were far away from the city being replaced. Each site was examined visually using Google Satellite View to ensure the location was not in an urban or semi-urban location. The sites were selected based on length of the record and suitability of the site and the data were obtained from the homogeneity adjusted archive (Vincent et al. 2002). Table 1 shows the decadal least-squares temperature trends for each location. Standard errors are not shown since they will be calculated in the full regression model with a correction for panel-specific autocorrelation (see Section 3.2). The data set was assembled a year prior to conducting the statistical analysis and no resampling, screening or replacement of locations occurred after the analysis began. Summary statistics are shown in Table 2.

Statistics Canada does not have monthly population data for cities, but a suitable replacement was found in the form of quarterly provincial population data from the online CANSIM system. In the post-1979 interval, population growth across Canada was almost entirely concentrated in major urban centres, so changes in population in each province are good indicators of the growth of the major cities in this sample. The data were filled out to match the monthly frequency of the temperature data. That is, the same value was used for each of the three months in a quarter. The population data were centered on a zero mean city by city and converted to units of one million. Each climatic series was converted to anomalies by removing the monthly means, city-by-city.

3.2 METHODS

I estimated the following regression equation:

$$T_{it}^{j} = a_0 + a_1 D_{it} + a_2 t + a_3 t \times D_{it} + a_4 P_{it} + e_{it} \qquad (5)$$

where T_{it}^{j} denotes the mean temperature anomaly in city i (=1,…,N) in month t (=1,…,T) in either an urban (j=U) or a rural (j=R) location, D_{it} is an indicator (or dummy) variable taking the value 1 in month t if average wind speed was one standard deviation above that city's average and 0 otherwise, P_{it}^{j} is population in city i at time t, and e_{it} is a regression residual. (Similar results to those reported herein are obtained if the cutoff is set to 1.5x or 2x the standard deviation, but in the latter case there are only 1 or 2 percent of cases classified as windy, making the comparisons unreliable.)

Equation (5) implies that the estimated trend through data points during relatively calm periods (D_{it} =0) is $\hat{a}_2$ while that through windy periods (D_{it} =1) is $\hat{a}_2 + \hat{a}_3$. Thus a t-test on $\hat{a}_3$ tests the null hypothesis that the trend under windy conditions is less than or equal to that under calm conditions. A Parker-type result would arise in a restricted version of (5) in which the coefficient on population (a_4) is set to zero and the wind differential term ($\hat{a}_3$) is statistically insignificant. An encompassing model would allow a_4 to be unrestricted. Were the estimate $\hat{a}_4$ to be significant while $\hat{a}_3$ is insignificant whether or not population is included in the equation, that would imply that windiness does not reliably measure whether population growth affects the trends. A consistency test is provided by redoing the analysis on the rural sample and checking that the population effects are not significant, irrespective of the wind effects.

Equation (5) was estimated using Driscoll and Kraay (1998) standard errors, which are robust to heteroskedasticity, serial correlation and spatial dependence of unknown forms. This is a non-parametric approach in the class of estimators known as HAC (heteroskedasticity and autocorrelation consistent) in the econometrics literature, and is implemented in Stata 12 using the "xtscc" command. The estimation requires specifying the maximum autocorrelation lag length, which was assumed to be four periods.

3.3 RESULTS

The results are shown in Table 3. Looking first at the restricted model, the urban results appear to show a Parker-like finding, namely no significant difference in trends between the calm and the windy conditions. This is illustrated in the top panel of Figure 1, where the fitted temperatures (respectively $\hat{a}_0 + \hat{a}_2 t$ on calm nights and $\hat{a}_0 + \hat{a}_1 D_{it} + \hat{a}_2 t + \hat{a}_3 t \times D_{it}$ on windy nights) follow a positive trend and are nearly parallel. The t-test on α_3 (H_0: no difference between the trends) has a p value of 0.922, indicating that we cannot reject the hypothesis of equivalence between the trends. The trend on calm nights equates to 0.216 degrees C per decade, but is not significant (p=0.151). The trend on calm nights in rural locations is 0.240, a bit higher, though likewise insignificant.

Rural areas also exhibit no difference in trends based on night time wind speed, as shown in the bottom panel of Figure 1. The coefficient on the trend difference has a p value of 0.827. The Parker

(2004, 2006) model implies that since neither the urban nor the rural locations exhibit a difference in trends based on wind speed, urbanization has no effect on the overall temperature trend. But the unrestricted model, in which local population is introduced, tells a different story. These results are shown in columns 5 and 6 of Table 3 and in Figure 2. The population variable is positive and highly significant (p=0.013) in the urban data but not in the rural data (p=0.599). This is consistent with the prior expectation that the urban data are contaminated and the rural data are not. It also indicates that an insignificant difference in trends on windy and calm nights does not imply the absence of non-climatic bias in the data, or more precisely that a model with only windspeed and a time trend does not encompass one that also includes population.

It is not the case that these results imply population growth explains all the warming across Canada. In the rural sample where population has no significant effect, the trend over months with low wind speed is 0.336 degrees per decade, though it is not significant (p=0.157). In the urban data the zero trend indicates that the amount of climatic warming simply cannot be identified separately from the population effect.

The example in this section shows that that one must be careful when using one variable to measure something else. A failure to observe a difference of trends in a sample stratified by wind speed cannot reliably be interpreted as the absence of an effect due to land use changes. The encompassing model demonstrates that an insignificant difference in trends in data stratified by calm and windy conditions can still be consistent with a significant urban warming effect due to regional population growth.

4 URBAN-RURAL SAMPLE SPLITS

4.1 CONCEPTUAL ISSUES

Another way of looking for evidence of bias in the surface temperature record is by dividing the data into rural and urban regions based on a static threshold, then comparing temperature trends. While some studies have reported differences based on this method (some results in Jones et al. 1990, as well as de Laat and Maurellis 2004 and 2006), others do not (Peterson et al. 1998, Hansen et al. 1999). Most recently Wickham et al. (2013) reported that in a global sample, the rural trends were slightly higher than those in an urban sample, though the difference was insignificant.

A key problem with this approach is that it relates a change term (temperature trend) to a level variable (land classification) rather than to a corresponding change variable (such as the change in surface conditions). As a result the findings are inherently ambiguous. In principle the urbanization process could result in a faster measured warming rate in a rural area than in a city, or no difference at all. Referring to Case 1 in Figure 3, suppose there are only two weather stations in the world, one rural and one urban. Suppose also that there is zero climatic warming over some interval, but there is a false warming due to local population growth. Suppose also that the effect of urbanization on temperature is logarithmic and that the economic growth over a given time period is indicated by the horizontal run of the respective arrows. A sample split according to the rural/urban distinction would apparently show that the rural station has a faster warming trend than the urban one. Far from proving that there is no urbanization bias in the overall average, this difference emerges easily in an example that assumes there is nothing but such a bias. And the contrast would be larger, the wider is the difference between "urban" and "very rural".

Consequently finding a slightly larger warming rate in rural areas compared to urban areas (as in Wickham et al. 2013) does not imply that there is no urbanization bias. Nor does it prove there is—it is consistent with either hypothesis.

Case 2 refers to a situation in which there is no urbanization bias, so if there is no climatic warming there is no temperature change; and alternatively if there is a temperature trend it is not due to local land use change. Which is the correct view? If Case 2 is correct, then land use change will have no explanatory power in a model of temperature trends. Hence we need to test if a change term is encompassed by urbanization level, or if it has unique explanatory power.

4.2 Data and Methods

A suitable data set is from McKitrick and Nierenberg (2010, herein MN10). This is based on the data set of McKitrick and Michaels (2007), augmented with more recent data products and spatial weights matrices based on maximum likelihood estimation of the distance decay parameter. Wickham et al. (2013) critique the use of national population levels in some remote MN10 grid cells, but the estimation in this section only makes use of the population change measures, not the underlying levels.

MN10 estimated the regression equation

$$CRU3v_i = \beta_0 + \beta_1 UAH4_i + \beta_2 PRESS_i + \beta_3 DRY_i + \beta_4 DSLP_i + \beta_5 WATER_i + \beta_6 ABSLAT_i + \beta_7 p_i + \beta_8 m_i + \beta_9 y_i + \beta_{10} c_i + \beta_{11} e_i + \beta_{12} g_i + \beta_{13} x_i + u_i$$

(6)

where $CRU3v_i$ is the 1979-2002 trend in the CRUTEM3v gridded surface climate data (Brohan et al. 2006) in grid cell i, $UAH4_i$ is the time trend of Spencer-Christy Microwave Sounding Unit (MSU)-derived temperatures in the lower troposphere in the same grid cell as $CRU3v_i$ over the same time interval (Spencer and Christy 1990), $PRESS_i$ is the mean sea level air pressure, 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 ^{o}C), $DSLP_i$ is $DRY_i \times PRESS_i$, $WATER_i$ is a dummy variable indicating the grid cell contains a major coastline, $ABSLAT_i$ denotes the absolute latitude of the grid cell, p_i is local population change from 1979 to 2002, m_i is per capita income change from 1979 to 2002, y_i is total Gross Domestic Product (GDP) change from 1979 to 2002, c_i is coal consumption change from 1979 to 2002, g_i is GDP density (national Gross Domestic Product per square kilometer) as of 1979, e_i is the average level of educational attainment, x_i is the number of missing months in the observed temperature series and u_i is the regression residual. There are 428 observations in this data set. MN10 also used the MSU product from Mears et al. (2003) and found it implied slightly stronger nonclimatic effects once some outliers were removed, but the data set has less spatial coverage than the Spencer-Christy record, so the latter is used here.

Equation (6) explains the spatial pattern of temperature trends in terms of three groups of explanatory variables: temperature trends in the lower troposphere, fixed geographical factors, and socioeconomic variables. The standard interpretation of climate data is that the socioeconomic effects have been filtered out of climatic data products like CRUTEM3v.

Summary statistics are in Table 4. The tropospheric data are at a 2.5x2.5 degree level, one-fourth of the 5x5 CRU surface grid size. To reconcile the spatial scales between surface and tropospheric gridcells MN10 develop matched 5x5 grid cells.

The surface temperature field is spatially autocorrelated, which can, in principle, bias the inferences from regressions on the spatial trend field. We test for spatial dependence in the residuals as follows. The regression model (6) can be rewritten in matrix notation as

$$\mathbf{T} = \mathbf{Xb} + u \tag{7}$$

where $\mathbf{T}$ is a 428x1 vector of temperature trends in each of 428 surface grid cells, $\mathbf{X}$ is a 428xk matrix of climatic and socioeconomic covariates, $\mathbf{b}$ is a kx1 vector of least-squares slope coefficients and u is a 428x1 residual vector.

Spatial autocorrelation in the residual vector can be modeled using

$$u = \lambda \mathbf{W} u + e \tag{8}$$

where λ is the autocorrelation coefficient, $\mathbf{W}$ is a symmetric $n \times n$ matrix of weights that measure the influence of each location on the other, and e is a vector of homoskedastic Gaussian disturbances, (Pisati 2001). The rows of $\mathbf{W}$ are standardized to sum to one. A test of $H_0 : \lambda = 0$ measures whether the error term in (8) is spatially independent. Anselin et al. (1996) point out that if the alternative model allows for possible spatial dependence of $\mathbf{T}$, i.e.

$$\mathbf{T} = \phi \mathbf{ZT} + \mathbf{Xb} + e \qquad (9),$$

where **Z** is a matrix of spatial weights for **T** and may not be identical to **W**, then conventional tests of $\lambda = 0$ assuming an alternative model of the form $\mathbf{T} = \mathbf{Xb} + e$ will be biased towards over-rejection of the null. They derive a $\chi^2(1)$ Lagrange Multiplier (LM) test of $\lambda = 0$ robust to possibly nonzero ϕ in (9). Hypothesis tests and parameter estimations using **W** are conditional on the assumed spatial weights. Denote the great circle distance between the grid cell centers from which observation *i* and observation *j* are drawn as d_{ij}. The weighting function is $d_{ij}^{-\mu}$ where μ determines the rate at which the relative influence of one cell on adjacent cells declines and is estimated in MN10 by a maximum likelihood grid search routine.

On the MN10 data set the spatial lag term in λ in (9) is significant (p=0.002) but the residual lag term ϕ is not (p=0.160) indicating that (6) is a well-specified model of the surface temperature trends. To be conservative the spatial lag term was included in the regression models.

An alternative form of equation (6) was estimated in which the only explanatory variable was g_i, the static (1979) measure of (national) GDP per square km in each grid cell. A version was also tried in which g was replaced with a binary variable indicating if g_i was at least one standard deviation above the mean, but the same results were obtained, so this outcome is not reported.

4.3 Results

Table 5 shows the results from estimating the restricted case of (6) in which all coefficients are set equal to zero except the one on g_i, and the unrestricted case. The restricted case resembles the kinds of tests undertaken in Wickham et al. (2013), Hansen et al. (1999), Peterson et al. (1998) and others, where the sample is conditioned only on a static measure of the level of surface disruption at one point in time. The restricted model appears to show that there is no significant difference in trends based on the level of economic development, thus apparently confirming the conclusions of these studies.

But the unrestricted model tells a different story. GDP now correlates with temperature trends (illustrating the point made at the end of Section 2), as does educational attainment. The rate of population change and coal consumption now do as well. Rather than the restricted model proving that surface changes do not contaminate the temperature record, it appears from the unrestricted model to be more likely the case that a model of that form is simply not capable of measuring the effect. The restrictions necessary to turn equation (6) into a model with only g on the right hand side yield a χ^2 statistic with a value of 111.8, which is significant at <0.0001%. Hence the data reject the hypothesis that the other variables are encompassed by the static measure of changes to the local land surface.

5 CONCLUSIONS

The examples shown herein demonstrate the potential value of using an encompassing framework in order to settle debates between incommensurable statistical models. The examples herein reproduce apparently conflicting findings on a single data set, and then test whether the restrictions necessary to yield one set of results can be rejected or not. In both cases, a model that implies an absence of effects due to socioeconomic development is a restricted version of another model that implies the presence of such effects, and the restrictions can be rejected, indicating that the effects are not encompassed by the variables in the simpler model. In order to move matters closer to overall resolution, it would be useful to develop a global data base, pooling time series and cross sectional information at the international level to permit development of a genuinely comprehensive testing framework. Future work in this direction is planned.

6 REFERENCES

Anselin, L, Anil K. Bera, Raymond Florax and Mann J. Yoon (1996). Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics* 26: 77—104.

Brohan, P., J.J. Kennedy, I. Harris, S.F.B. Tett and P.D. Jones, 2006: Uncertainty estimates in regional and global observed temperature changes: a new dataset from 1850. *J. Geophys. Res.* **111**, D12106, doi:10.1029/2005JD006548

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, *Int. J. Climatol.* 26:897—913.

Driscoll, John C. and Art Kraay (1998) . “Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data.” *Review of Economics and Statistics* November 1998, Vol. 80, No. 4, Pages 549-560.

Hansen, J., R. Ruedy, J. Glascoe, and Mki. Sato (1999), “GISS analysis of surface temperature change,” *Journal of Geophysical Research*, 104, 30997-31022, doi:10.1029/1999JD900835.

Jones P.D, Groisman PY, Coughlan M, Plummer N, Wang WC, Karl TR (1990) “Assessment of urbanization effects in time series of surface air temperature over land.” *Nature* 347:169—172.

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

McKitrick, Ross R. and P. J. Michaels (2004b), Erratum, *Clim. Res.* 27(3) 265—268.

McKitrick, Ross R. and P.J. Michaels (2007), Quantifying the influence of anthropogenic surface processes and inhomogeneities on gridded global climate data, *J. Geophys. Res.*, 112, D24S09, doi:10.1029/2007JD008465.

McKitrick, Ross R. (2010), Atmospheric Oscillations do not Explain the Temperature-Industrialization Correlation. *Statistics, Politics and Policy* Vol 1, July 2010.

McKitrick, Ross R. and Nicolas Nierenberg (2010) Socioeconomic Patterns in Climate Data. *Journal of Economic and Social Measurement*, 35(3,4) pp. 149-175. DOI 10.3233/JEM-2010-0336.

Mears, Carl A., Matthias C. Schabel, and Frank J. Wentz. 2003. A Reanalysis of the MSU Channel 2 Tropospheric Temperature Record. *Journal of Climate* 16, no. 22 (November 1): 3650-3664 .

Mizon, G. E. (1984). 'The encompassing approach in econometrics', in Hendry D. F. and Wallis K. F. (eds), *Econometrics and Quantitative Economics*, Basil Blackwell, Oxford, pp. 135–172.

Parker, D.E. (2006) "A Demonstration that Large-Scale Warming is not Urban." *Journal of Climate* 19:2882—2895.

Parker, D.E. (2004) "Climate: Large-Scale Warming is not Urban." *Nature* 432, 290 (18 November 2004); doi:10.1038/432290a.

Peterson, Thomas C., Russell Vose, Richard Schmoyer and Vyachesvslav Razuavev (1998) "Global Historical Climatology Network (GHCN) Quality Control of Monthly Temperature Data." *International Journal of Climatology* 18: 1169—1179.

Pisati, Maurizio (2001) Tools for spatial data analysis. *Stata Technical Bulletin* STB-60, March 2001, 21—37.

Schmidt, Gavin (2009) "Spurious correlations between recent warming and indices of local economic activity." *International Journal of Climatology* DOI:10.1002/joc.1831.

Spencer, R.W. and J.C. Christy (1990) "Precise Monitoring of Global Temperature Trends from Satellites." Science 247: 1558—1562.

Vincent, Lucie, X. Zhang, B.R. Bonsal and W.D. Hogg (2002) "Homogenization of Daily Temperatures of Canada." *Journal of Climate* 15, pp. 1322—1334.

Wickham C, Rohde R, Muller RA, Wurtele J, Curry J, et al. (2013) Influence of Urban Heating on the Global Temperature Land Average using Rural Sites Identified from MODIS Classifications. *Geoinformatics and Geostatistics: An Overview* 1:2.

7 FIGURES

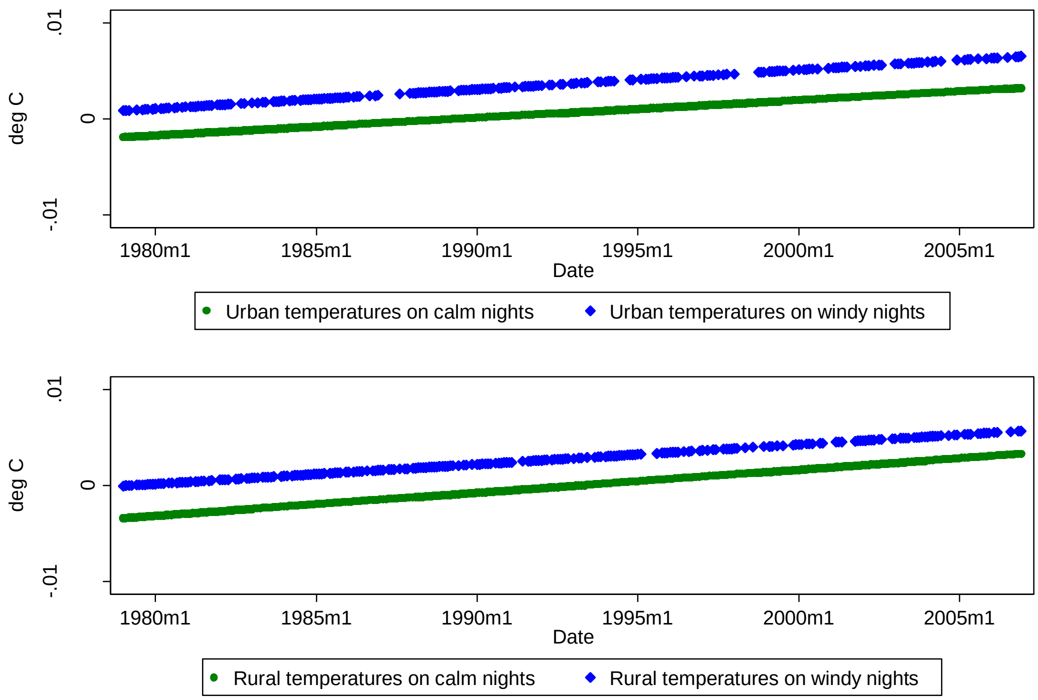

Figure 1. Fitted temperatures in urban (top) and rural (bottom) samples stratified by whether monthly average wind speed is one standard deviation above the local average (blue, "windy") or not (green, "calm"). Trends are estimated without controlling for any effects due to local population growth. Series are offset by constant terms to aid visual comparison. Gaps arise where no data are available for that case.

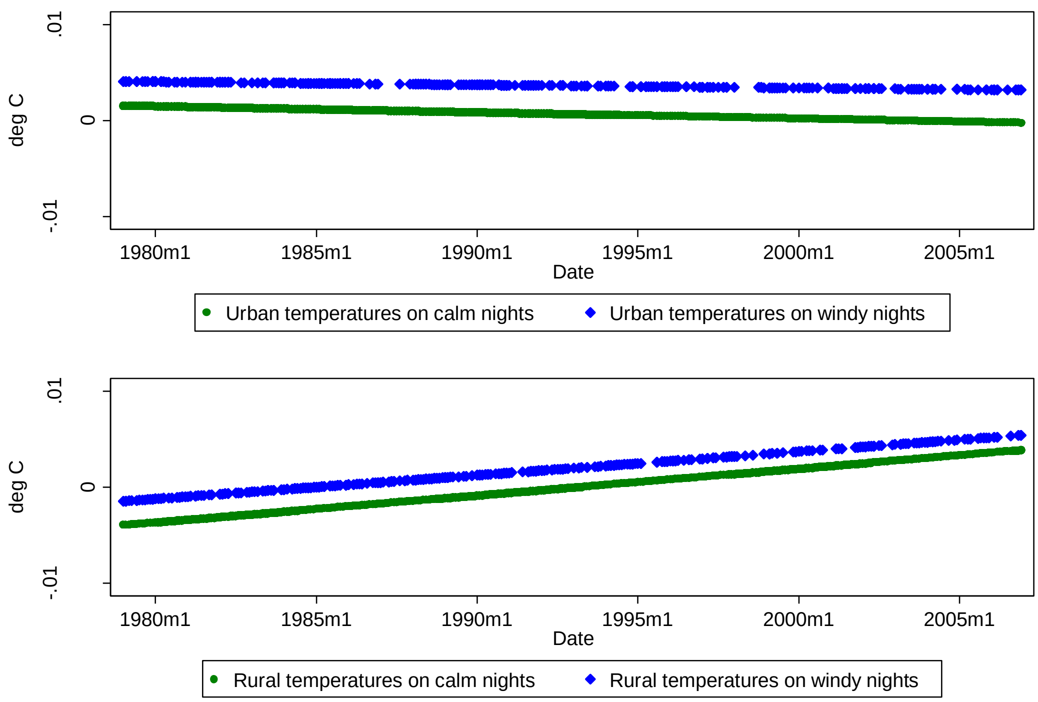

Figure 2. Fitted temperatures in urban (top) and rural (bottom) samples stratified by whether monthly average wind speed is one standard deviation above the local average (blue, "windy") or not (green, "calm"). Trends are estimated after controlling for effects due to local population growth. Series are offset by constant terms to aid visual comparison. Gaps arise where no data are available for that case.

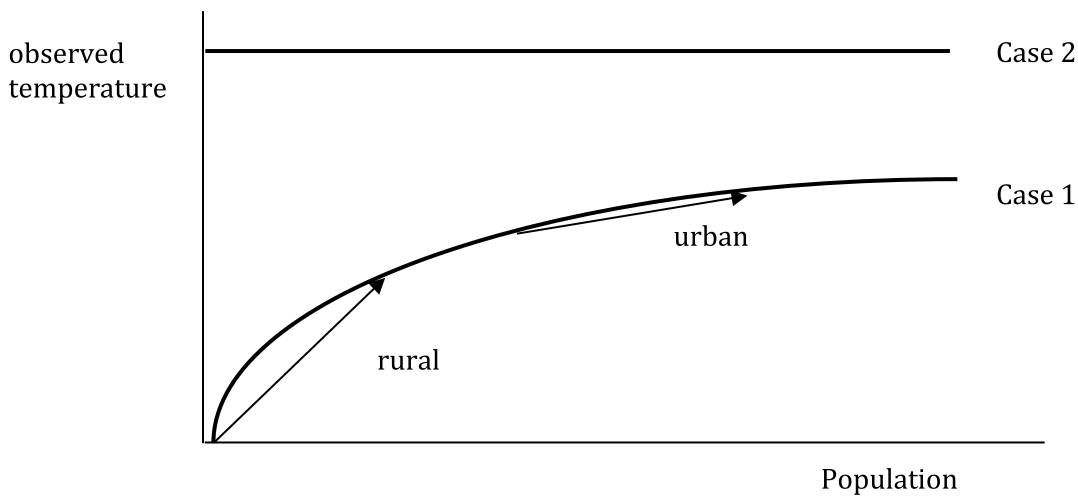

Figure 3. Conceptual representation of different warming bias rates at rural and urban locations.

8 TABLES

	Non-Homogeneity Adjusted Urban Temperature (a)			Homogeneity-Adjusted Remote Temperature (b)		
Province	Location	ID	Trend	Location	ID	Trend
AB	Calgary	3031093	0.097	Carway	3031400	0.405
AB	Edmonton	3012205	-0.229	Campsie	3061200	-0.060
QU	Montreal	7025250	0.614	Gaspe	7052605	0.744
ON	Ottawa ON	6106000	0.318	Peterborough	6166418	0.360
SK	Regina SK	4016560	-0.176	Swift Current	4028040	0.412
SK	Saskatoon	4057120	-0.208	Estevan	4012400	0.137
ON	Toronto ON	6158733	1.028	Wiarton	6119500	0.371
BC	Vancouver	1108447	0.401	Stewart	1067742	0.089
BC	Victoria	1018620	0.395	Quatsino	1036570	-0.213
MA	Winnipeg	5023222	0.122	Sprague	5022759	0.554

TABLE 1: Temperature Data for Section 2.
(a) http://www.climate.weatheroffice.gc.ca/climateData/ canada_e.html.
(b) http://www.ec.gc.ca/dccha-ahccd/default.asp?lang=En&n=B1F8423A-1.
"ID" refers to Environment Canada identifier, not WMO identifier. Trend: 1979-2006 linear temperature trend in degrees C per decade. Provinces: AB Alberta, ON Ontario, QU Quebec, SK Saskatchewan, BC British Columbia, MA Manitoba.

Variables	Obs	Mean	Std.Dev	Min	Max
URBAN					
Monthly min temp	3360	0	2.420	-12.143	10.580
Monthly mean wind speed	3360	0	1.656	-6.282	8.975
RURAL					
Monthly min temp	3360	0	2.623	-12.793	10.854
Monthly mean wind speed	3360	0	1.932	-10.046	12.954
Population (millions)	3360	0	0.554	-1.880	1.850
D (urban)	3360	0.141	0.314	0	1
D (rural)	3360	0.150	0.318	0	1

TABLE 2: Summary statistics of the data set for Section 2. Note that all data are centered.

		Restricted		Unrestricted	
Variables	Coefficient	Urban	Rural	Urban	Rural
Trend (calm)	a_2	0.0018 (1.57)	**0.0024** (1.92)	-0.0006 (0.36)	0.0028 (1.54)
Trend (windy)	$a_2 + a_3$	0.0020 (0.094)	0.0021 (1.14)	-0.0003 (0.12)	0.0025 (1.11)
Trend difference $(\boldsymbol{t} \times \boldsymbol{D_{it}})$	a_3	0.0002 (0.10)	-0.0003 (0.22)	0.0003 (0.16)	-0.0003 (0.22)
Windy Conditions Indicator $(\boldsymbol{D_{it}})$	a_1	-0.6154 (0.70)	0.1803 (0.32)	-0.6722 (0.76)	0.1740 (0.57)
Population $(\boldsymbol{P_{it}})$	a_4			**0.5555**** (3.10)	-0.0917 (0.55)
Constant	a_0	-0.6471 (1.36)	**-0.9541** (1.82)	0.3323 (0.46)	-1.11316 (1.49)
R^2		0.0123	0.0075	0.0188	0.0077

Table 3. Results from estimation of equation (5) . Restricted: population effect set equal to zero. Coefficient on trend terms (first 3 rows) are degrees per month. Terms in parentheses are absolute *t* statistics. Bold: significant at 10%, * significant at 5%, ** significant at 1%. Sample size = 3360 for all regressions.

Variable	Definition	Mean	Std.Dev.	Min	Max
CRU3v	1979-2002 Surface Temp Trend (C/decade)	0.2761	0.2443	-0.717	1.042
UAH4	1979-2002 Tropospheric Temp Trend (C/decade)	0.2206	0.1732	-0.1390	.7414
PRESS	Sea Level Pressure	1016.29	4.987	993	1029
DRY	Predominantly dry region	0.376	0.4835	0.0	1.0
WATER	Grid cell contains coast	0.6060	0.4892	0.0	1.0
ABSLAT	Absolute latitude	35.97	16.79	2.5	82.5
g	GDP per square km	0.3010	0.6029	0.0014	3.002
e	Educational level	103.58	28.10	11.6	144.2
x	Months w/o surface temperature data	0.5812	1.938	0.0	24
p	% Population growth*	0.3110	0.218	-0.0691	1.2353
m	% Income growth*	0.4172	0.6339	-0.7901	2.147
y	% GDP growth**	0.8536	0.8597	-0.6686	3.002
c	% Coal usage growth*	1.2869	4.759	-1.0	39.33

Table 4: Data used for estimating equation (6). Weighted by cosine latitude to control for grid cell size. uah4: 4 2.5x2.5 degree gridcells combined to match area of surface 5x5 gridcell. *over the interval 1979 to 1999. **Over the interval 1980 to 2000. % Changes should be multiplied by 100, e.g. mean population growth is 31.1%. Sample size = 428.

Variable	**Definition**	**Restricted**	**Unrestricted**
UAH4	Trop. Temp Trend (C/decade)		**0.727**** (9.045)
PRESS	Sea Level Pressure		**0.006*** (2.157)
DRY	Predominantly dry region		4.738 (1.367)
DSLP	$dry \times slp$		-0.005 (1.346)
WATER	Grid cell contains coast		-0.029 (1.564)
ABSLAT	Absolute latitude		0.000 (0.450)
g	GDP per square km	0.0116 (1.14)	**0.030*** (2.438)
e	Educational level		**-0.002**** (3.581)
x	Months w/o surface temp. data		0.000 (0.070)
p	% Population growth		**0.240*** (2.069)
m	% Income growth		0.178 (1.361)
y	% GDP growth		-0.139 (1.351)
c	% Coal usage growth		**0.004*** (2.19)
constant		**0.0433** (3.15)	**-6.025*** (-2.108)
R^2		0.521	0.581

Table 5: Coefficient estimates for Equation (6). Dependent variable is gridcell trend *cru3v*. Sample size = 428. Second column: regression on *g* and constant only. Third column, allowing all other variables to enter model. R^2 is the squared correlation between the observed and predicted dependent variable. Bold: significant at 10%, * significant at 5%, ** significant at 1%.