Population Research and Policy Review **22:** 297–331, 2003.

Blowin' down the road: Investigating bilateral causality between dust storms and population in the Great Plains

GLENN DEANE[1] & MYRON P. GUTMANN[2]
[1]*Department of Sociology and Lewis Mumford Center for Comparative Urban and Regional Research, SUNY University at Albany, USA;* [2]*Department of History and Inter-University Consortium for Political and Social Research, Institute for Social Research, The University of Michigan, USA*

Abstract. Recently, the National Academy of Sciences concluded "it is clear that population and the environment are usually interrelated . . . ". This paper directly tests the expected interrelationship using annual county-level population estimates provided by the U.S. Census Bureau and annual counts of dust storms from the 1960s, '70s, and '80s at weather stations situated throughout the U.S. Great Plains. In doing so, it implements a research design that extends methods (far removed from conventional demography) for pure time series analysis with multilevel regression models. The result is a method for causal modeling in panel data that produces, in this application, evidence of bilateral causality between population size and deleterious environmental conditions.

Keywords: climate change, granger causality, population and environment, U.S.Great Plains

Recently, in a reflection on the current (and future) state of demography, Samuel Preston observed that "Several forces are converging to create powerful pressures for conducting research between population growth and environmental quality . . . . The necessary research designs, incorporating both macro and micro-level features, are far removed from those conventional in demography . . . . Although the study of relations between population growth and environmental change isn't demography, it isn't anything else either. We can expect new interdisciplinary research structures to be created in which demographers will play a prominent role" (Preston 1993: 600).

In the same year, the National Academy of Sciences concluded "it is clear that population and the environment are usually interrelated, but the strength and mechanism of action of the relationship varies widely from setting to setting . . . . To date, cross-national studies have been intriguing, but have failed to resolve the magnitude and mechanism of action governing the relationship between population and environment. The next logical step for research is to examine a number of case studies of differing dimensions to see how population change and the environment are interrelated . . . . It is also important to account for the changes over time and to be able to relate population change to the environment meaningfully" (NAS 1993).

The purpose of this paper is to test directly this expected "interrelationship" using annual county-level population estimates provided by the U.S. Census Bureau and annual counts of dust storms from weather stations situated throughout the U.S. Great Plains. We compile and analyze two complementary panel data files, one which exploits a longer time series (1962–1988) with fewer cross-sectional units (twenty-two counties) and one of a shorter time series (1969–1988) but with more cross-sectional units (thirty-nine counties). In doing so, we implement a research design that incorporates macro- and micro-level features in a single application by extending methods (far removed from conventional demography) for pure time series analysis with multilevel regression models. The result is a method for causal modeling of panel data that produces, in this application, evidence of bilateral causality between population change and deleterious environmental conditions.

Structural models of population and environment

The interrelationship between population and the environment alluded to by the National Academy of Sciences in fact implies three possible causal models: (1) unidirectional causality from population to environment; (2) unidirectional causality from environment to population; or (3) feedback, or bilateral causality. History holds many examples of this interrelationship, some lending credence to the unidirectional effect of population on the environment, some tending toward the unidirectional effect of environment on population, and others favoring bilateral causality. To illustrate what we might expect of the causal structure of population and dust storms in the latter part of the twentieth century, we review some of the historical record below.

Causal impact of environment on population

It is probably fair to attribute the origin of our current awareness of the impact of the environment on population to the social historian Karl Wittfogel. Wittfogel began formulating his hypothesis linking the role of irrigation to the development of early civilization, using the designations "hydraulic civilization" and "hydraulic society", in the 1930s with his insight reaching fruition in the classic comparative study of total power he titled Oriental Despotism (Wittfogel 1935, 1939-1940, 1957). Writing during an era when the main currents of anthropological and historical theorizing on cultural evolution were descriptive and particularistic, Wittfogel pointed to cross-cultural regularities and "cultural causality". His "irrigation hypothesis" integrated the development of early civilizations in the East and the West by the managerial

control required to construct and maintain irrigation systems. As water was brought to arid lands, food production and population increased and became the basis for class-structured states. Wittfogel's irrigation hypothesis initiated a research trend that continues to inform social histories (Steward 1978).

Climatic impact on population growth/redistribution is witnessed by two recent, and coincident, U.S. migration streams: the South–North migration reversal, usually referred to as "migration to the Sunbelt" (Biggar 1979) and the metropolitan-nonmetropolitan turnaround (Long 1987) of the 1970s represent another commonly acknowledged population response to environment. In both instances, positive net in-migration (to the South and nonmetropolitan areas) is thought to have been (largely) determined by climate and desires for a lifestyle that emphasizes outdoor recreation. Economists have also documented a lower cost of living, itself a potential determinant of migration, associated with the milder climate found in the American South (Dickinson 1978).

Another place we find the causal impact of environment on population is in the influence of disease in history. Since the mid- 1970s, a sizeable literature has been amassed through the efforts of social and medical historians, anthropologists, and paleopathologists on "Old World" and "New World" disease reservoirs, the susceptibility of hosts and the virulence of infections, vectors of disease transfer, and the direction of transfer. Although most of the debate among the practitioners of this area of inquiry has concerned population contact and the consequences thereof, coexisting with these central themes is speculation over the negative relationship between the ecological conditions favorable to disease and human population settlement. William McNeill provides the most succinct articulation of this relationship.

In *Plagues and Peoples* (1976), McNeill locates the origin of Chinese settlement in the northern semi-arid environment near the Yellow River (Huang He) flood plain. From this toehold, Chinese population extended into the river's flood plain by 600 B.C. but, despite the inhospitable conditions of the geologically unstable Yellow River, it would be another thousand years before Chinese settlement extended south into the valley of the Yangtze (Chang Jiang) River.

The reason for this slow march to the south, according to McNeill, was that "Chinese pioneers were also climbing a rather steep disease gradient .... Put very simply, too many immigrants from the cooler, drier North died to permit a more rapid buildup" (see McNeill 1976: 73–80). Timothy Bratton makes a related argument for the exceptional virulence of smallpox, and other "virgin soil epidemics", among AmerIndians (Bratton 1988). Substantial documentation of "pre-contact" pathogens in the New World can be found in the issues of *The Paleopathology Newsletter*. These afflictions in-

clude tuberculosis, emphysema, pneumonia, atherosclerosis, coccidioidomycosis, trichinosis, melanoma, asthma, typhoid, systematic lupus, rheumatoid arthritis, osteosarcoma, and spina bifida.

And of course there is considerable debate over the origin of syphilis. Indeed a compelling account of the dual forces of geographic and disease environments on the fates of human societies is found in Jared Diamond's recent prize-winning book *Guns, Germs, and Steel* (1997). Nowhere, however, is the causal impact of environment on human population more apparent than in the images of the American Dust Bowl of the 1930s. These images have been powerfully captured by photojournalists such as Russell Lee, Dorothea Lange, and Arthur Rothstein – their photographs, are vivid reminders of the enormity of the "rollers" as they enveloped whole counties and the devastation they left in their wake; by novelists such as Nathaneal West (*The Day of the Locust*) and John Steinbeck (*Grapes of Wrath*); and by folk singers, most notably, Woody Guthrie (Dust Bowl references in Guthrie's song titles include: *Dust Storm Disaster, Dust Can't Kill Me, Dust Bowl Refugee, Talkin' Dust Bowl, Dusty Old Dust*, and the song from which we excerpted the title of this paper *Blowin' Down This Road*).

Although most Americans understand the referent and have general knowledge of the period referred to as the "Dust Bowl Era", the details of the place and period are far less well known. The "Dust Bowl" a term coined by Associated Press reporter Robert Geiger in 1935, became an official Soil Conservation Service region (SCS Region VI) that same year consisting of the western third of Kansas, Southeastern Colorado, the Oklahoma Panhandle, the northern two-thirds of the Texas Panhandle, and Northeastern New Mexico. In contemporary geographic nomenclature, the Dust Bowl is located within the southern Great Plains states.

Although observers in weather stations of the U.S. Weather Service have been recording occurrences of dust, blowing sand, blowing dust, and dust storms as "present weather" conditions since at least the late 1940s, comprehensive records of these "dust events" do not exist for the Dust Bowl region during the 1930s. Still, it is possible to relate the comparative severity of dust storms during that decade by drawing together information from a variety of sources. First, the Soil Conservation Service did compile a frequency chart of all dust storms of *regional extent* during the 1930s. In 1932 there were 14; in 1933, 38; 1934, 22; 1935, 40; 1936, 68; 1937, 72; 1938, 61; 1939, 30; and by 1940 and 1941, the count had again dropped into the teens with 17 regional storms recorded in each of these years. Second, in November of 1932 and again in May of 1934, dust storms originating in the Great Plains affected cities as far east as New York and Savannah, Georgia (Worster 1979: 13–15). Third, records kept at the Panhandle A & M Experimental Station at

Goodwell, Oklahoma reported 70 days of severe dust storms in 1933; 22 in 1934; 1935, 53; 1936, 73; and a staggering 134 dust storms in 1937 (Bonnifield 1979: 65, 70–71). The loss in topsoil due to wind erosion represents another measure of devastation during the 1930s. By 1938, 23.5 million acres in the Dust Bowl had lost at least their upper two and a half inches of topsoil. Throughout the region, 850 million tons of soil a year was being lost due to erosion during the 1930s, or about 408 tons of dirt blown away per acre of cultivated land (Worster 1979: 29).

The terms "Okies", "exodusters", and "Dust Bowl Refugees" have become linguistic icons for the population consequence of these dust storms and Steinbeck's The Grapes of Wrath forever etched in the American collective conscience the images of foreclosure, expulsion, and the long, hard migration west. The statistical evidence provides (qualified) support for these impressions. During the 1930s the only states that had fewer residents at the end of the decade than at the beginning were in Great Plains. In a survey of forty Dust Bowl counties, the Resettlement Administration (later called the Farm Security Administration) found that between 1930–1935 county farm population decreased by less than three percent, but between 1935–1937, after the dust storms began, over 34 percent of the population had left. In all, almost a million plains people left their farms in the first half of the 1930s, and 2.5 million left after 1935. The only difference between the reality of the "dirty thirties" and Steinbeck's archetypical Joads was that most movers didn't move far, most were simply blown into the nearest town or the next county or the next state (Worster 1979: 48–52).

Causal impact of population on the environment

Of all the potential and real human impacts on our environment, land-cover change and species extinction stand out as our most consistent mark. Human population has created a legacy of stripping land cover and animal destruction beginning at the site where Western civilization first emerged, the Fertile Crescent, and accelerating toward global deforestation at the end of the twentieth century (cf. Marsh 1874; Sauer 1963: 145–154; Perlin 1989; Richards & Tucker 1988; Meyer & Turner 1994).

Although the greatest (percentage) rate of deforestation is occurring in Thailand, El Salvador, Honduras, Paraguay, and the Philippines – all lost more than *twenty percent* of their forests just during the1980s and forest cover in the Philippines has decreased by 56 percent in the postwar period – global deforestation is a very real concern stretching from the outright land conversion of tropical rainforests in South America and virgin and second-growth bush in Sahelien Africa to the biotic degradation of existing stock in the former Soviet Union (Richards & Tucker 1988).

Probably the most widely discussed, though hotly contested, human population impact on the environment is global warming, sometimes referred to as the "greenhouse effect", and destruction of the ozone layer. Global warming is caused by the increase in atmospheric concentration of greenhouse gases. These greenhouse gases, primarily carbon dioxide, methane, and nitrous oxide, absorb outgoing terrestrial radiation while permitting incoming solar radiation to pass through the atmosphere. Although it is this basic heat-trapping mechanism that keeps the atmosphere of the earth relatively warm compared to other planets, and is in this sense "natural" and necessary for life on this planet, the enhancement of the greenhouse effect due to the increasing concentration of greenhouse gases is cause for serious concern (Adger & Brown 1994: 3–12). The primary human activity responsible for additional warming of the earth's surface is the burning of fossil fuel for energy. In addition, human land use, including the global deforestation discussed above, also causes global warming (cf. Adger & Brown 1994). Whatever the final scientific and political assessment of the importance of global deforestation and global warming to the well- being of all plant and animal species on earth will be, clearly these phenomena represent outstanding examples of what it means to claim a causal impact of population on the environment.

Feedback, or bilateral causality

If indeed the stripping of land cover and the human contribution to greenhouse gases cause global climate change, sea level rise, ozone destruction, and related impacts, then we have a ready example of a third causal structure, bilateral causality. It is possible, however, to point to less terminal situations for circumstances of causal feedback. For example, it is somewhat ironic that the Sunbelt and nonmetropolitan migration streams we identified above with the causal impact of environment on population have also been cited as the root for destroying the very lifestyle that attracted Snowbelt and metropolitan migrants in the first place. The following passage illustrates this connection painfully well:

> Jack Douglas can't go rabbit hunting near his Florence, Alabama, home the way he used to. The fields are covered with sub-divisions. His favorite bass-fishing spot on nearby Wilson Lake is churned by water skiers all summer, and the fish don't bite anymore. The family-owned restaurant where he used to eat fresh-baked pies and spicy barbeque sold out to the Bonanza steak chain last year. (Biggar 1979: 34)

A fitting ending to this story might continue something like this: "Jack Douglas could no longer tolerate the changes. He packed his truck and moved to Bozeman, Montana".

If we accept the traditional "explanation" of the Dust Bowl – the convergence on single crop agriculture (wheat) in the Dust Bowl region of the Great Plains (Bonnifield 1979; Worster 1979) – we have the necessary ingredients of bilateral causality here too: plains farmers (in part) caused the dust storms that ignited the migrations of the second half of the 1930s. What is more, the human impact on the magnitude and frequency of blowing dust is neither confined to the U.S. Great Plains nor the 1930s. Finnell (1948, 1954) attributed renewed dust storm activity in the southern plains to the plow-up of marginal soils due to the expansion of agribusiness wheat farming and "the 'suitcase farmer', who leases a tract for a short period, mines it to exhaustion, and moves on" and the planting of cotton and feed-crop cultivation in response to increased demand during World War II (Finnel 1948: 10).

Throughout the second half of the twentieth century, geographers and agricultural economists have observed a systematic covariation in the frequency of dust storms with various human activities, including land conservation practices in the (Texas) Southern High Plains (Ervin & Lee 1994; Lee & Tchakerian 1995) and urban and industrial development in Mexico, Saudi Arabia, and Mongolia (Jauregui 1989; Behairy et al. 1985; Gouldie & Middleton 1992). In short, while it is easy to point to a lack of rain as the cause of dust storms, in many cases, according to many researchers, that is not the real explanation. Indeed, it comes as no surprise then that the Canadian soil scientist Chepil and his colleagues developed a wind erosion equation (WEQ) that linked erosion to five factors, three of which are directly controlled by farming practices (Chepil et al. 1963). The blame falls not on the elements, but on human activity.

While issues of causality must ultimately be left to the realm of theory and philosophy, the question of whether one can statistically detect the presence and direction of causality has been given considerable attention in econometrics for time series data. We extend and apply one of these methods, the Wiener-Granger causality test (cf. Granger 1969; Sims 1972, 1980), to panel time series from the

U.S. Great Plains in the 1960s, '70s, and '80s

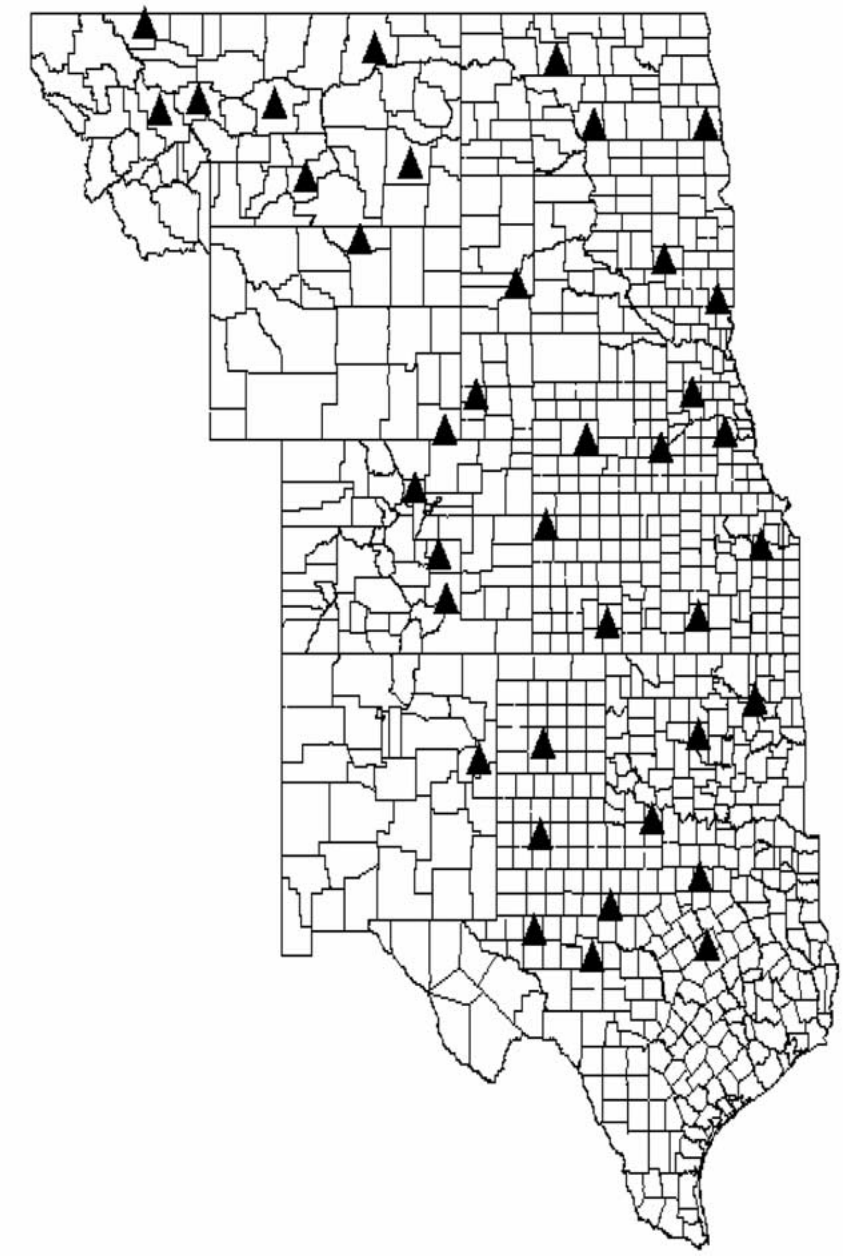
Data: annual time series of dust storms and population in Great Plains counties An empirical test of the interrelationship between population and the environment requires observable measures of these broad concepts and the data requirements for the successful deployment of the Wiener–Granger method are stringent. The method is designed to detect short-term changes in time series, rather than long swings, so the variables we want must, in principle, be able to respond quickly. The method also requires a large number of data points so annual observations (minimally) are desirable. Although

there may be many suitable settings and indicators for such an empirical test, the image of bilateral causality in the Dust Bowl is compelling and the variables involved, population size and number of dust storms, can, as we documented above, meet the conditions necessary for an analysis of this type. Unfortunately annual enumerations, or reliable estimates, of population size and counts of dust storms are not attainable from the 1930s – but extended time series for both of these variables are available.

The U.S. Department of Commerce has published annual county total population estimates since 1969. The Bureau of the Census develops these county estimates with a demographic procedure called the Tax Return method, which is a "component change" procedure.[1] In addition, some states provided their own annual county population estimates, based on the same procedure, before the Census Bureau estimates became available for all counties (U.S. Department of Commerce 1967). The U.S. Weather Service has recorded "present weather" conditions, including occurrences of dust, blowing sand, blowing dust, and dust storms (which together we refer to as "dust storms"), since the 1940s at a large number of weather stations situated throughout the United States (Karl et al. 1990). Thirty-nine of these weather stations are located in (or approximately in) the Great Plains and observations, taken at three hour intervals, of dust events are available from 1961 through 1988.

The analyses to be presented in this paper exploit two overlapping panel data series: one with the full set of cross-sectional units but of shorter duration (the full 39 counties with weather stations have Census Bureau population estimates for 1969–1988) and another of longer duration but with fewer cross-sectional units (twenty-two weather station counties have population estimates available from 1962 through 1988 – nearly exploiting the full time series of dust storms). The resulting study samples have 780 data points [1969–1988 (= 20 years) $\times$ 39 counties] and 594 data points [1962–1988 (= 27 years) $\times$ 22 counties], respectively. These study sample counties are highlighted in the Great Plains map shown below. Weather station counties in Kansas, Montana, New Mexico, Oklahoma, and Texas comprise the long (1962–1988) series.

Both variables show considerable variation across the study space. Figures 1 and 2 plot the annual counts of dust storms and population estimates for each of the weather station counties. The heavy bars in Figures 1 and 2 show the (weighted) averaged time series – the line points are determined by adding all counts (e.g., dust storms) in a given year and dividing by the number of contributing counties. The "dip" in the (log) population line between 1968 and 1969 is not indicative of a sudden decline in population, rather it is due to the shift in the number of contributing counties. Two points are clear from Figure 1. First, there is an outlying time series on dust storms. This time series



Map 1. Weather station locations in Great Plains.

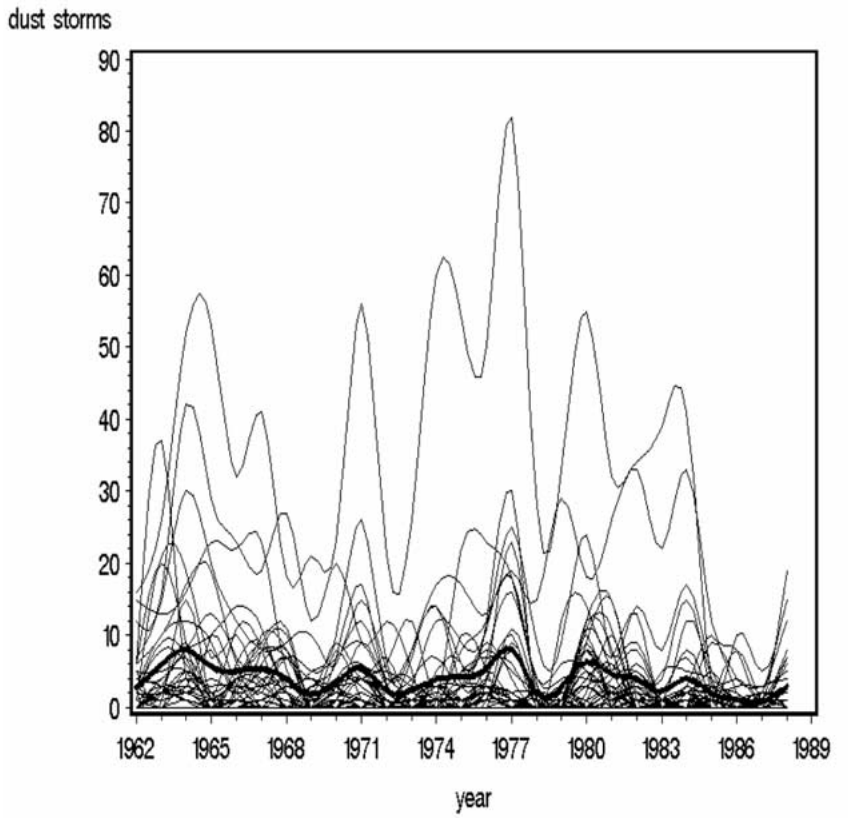


Figure 1. County time series of dust storms, 1962–1988. Heavy bar shows mean of county series.

belongs to Lubbock County, TX. It is far and away the dustiest place in the Great Plains. In 1977, 82 dust storms were recorded, in 1974, 60 dust storms, and a number of years experienced 50 or more storms. Although it is far less obvious than the top end outlying series for Lubbock, there is also a location (Sheridan, Wyoming) that experienced no dust storms over the twenty-year period. Second, if the individual time series are trending over the twenty-two year period, the direction of that trend is not obvious in Figure 1.

Figure 2 displays the population time series. As noted above, twenty-two of the 39 weather station counties have population time series for all twenty-seven years, while seventeen counties have population time series from 1969–1988. We gain visual separation in the time series by placing population size on a natural log scale. In its original metric, Tarrant County, TX (which contains the city of Ft. Worth) is an outlying series. The population

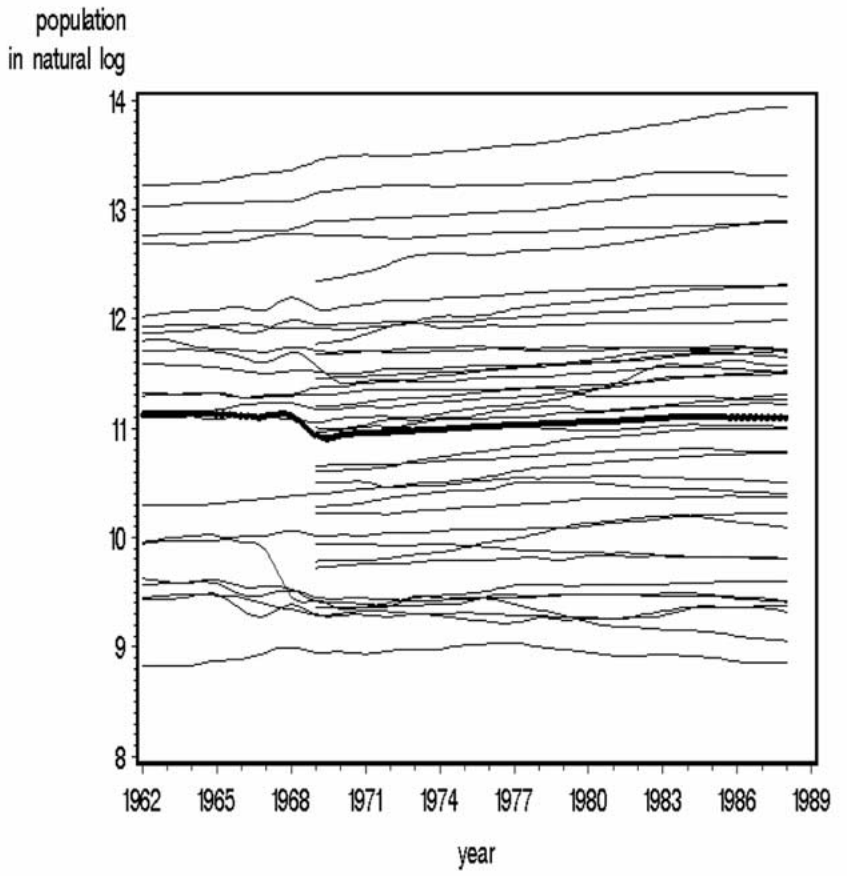


Figure 2. County time series of log population, 1962–1988. Heavy bar shows mean of county series.

of Tarrant County doubled between 1962 and 1989, by which time it is nearly twice as large as the next most populous county.

In the log scale of Figure 2, the geometric growth of Tarrant County is converted to a linear trend (see the time series at the top of Figure 2) but almost all of the other time series move in a nonlinear fashion. Only a handful of counties monotonically trend upwards, what is more remarkable is the absence of a clear pattern of population growth among the rest of these Great Plains counties. The majority of counties show either periods of decline or actual net decline over the twenty to twenty-seven year period.

Methods: time series methods for pooled cross-sections

Our ultimate goal is to diagnose causality in the bivariate time series. To do this we extend and apply a method for determining causal structure in time series (the Wiener-Granger method) to panel data series of population size and counts of dust storms using multilevel models.

The underlying logic of this method is that the future cannot predict the past, so if a variable X causes variable Y, then changes in X should precede changes in Y. The empirical realization of this can be put into operation through a two-step regression procedure. We begin with an auxiliary regression of Y on its own past values. We follow by adding lagged values of X improves the prediction of Y, we say that X causes Y. Typically then one reverses the order to detect whether Y causes X. This results in four possible causal structures: (1) unidirectional causality from X to Y; (2) unidirectional causality from Y to X; (3) feedback, or bilateral causality; and (4) independence.

The Weiner–Granger causality test is perhaps more appropriately considered a test of "feedback" rather than causality. Since "exogeneity" is often equated with the notion of "lack of feedback" in a system of equations, the Granger test may be used to determine whether this condition holds (causal structures 1, 2, or 4 above). In this sense, the method is used to determine that "X does not cause Y" or that "Y does not cause X". The identification of exogeneity is useful because it implies that no information is lost by limiting attention to distributions conditional on the exogenous variable thus considerably simplifying statistical inference (Mills 1990: 289–296). Given the concern over the appropriateness of the use of the term 'causality" in connection with the Granger test, it is often preferred to use the phrase 'X (Granger) causes Y" or "Y (Granger) causes X" if exogeneity is rejected.

Before we can make this application, however, the time series must be either stationary, i.e., not trending, or if trending, they must be trending together (a condition known as "cointegration"). In addition, the distributed-lag structure of the bivariate panels will be investigated so that we do not truncate the lagged effects. The following section will, therefore, include extensions of time series methods for detecting stationarity, cointegration, and lag structure prior to our treatment of Granger causality models. The section concludes with a panel application of a technique for the visual display of causal impact in dynamic vector autoregression systems known as "impulse response functions".

Detrending time series: why do we detrend?

Consider the following population regression function for pure time series, Y_t and X_t, where the subscript t indexes a specific period in the time series:

$$Y_t = B_0 + B_1 X_t + \epsilon_t. \tag{1}$$

Now suppose we have a research hypothesis that posits $B_1 > 0$. If X and Y trend upwards, $\hat{B}_1$ will be positive even if the two variables are unrelated because there is a third variable, time, that is not included in the regression. The consequences of model misspecification are well known (cf. Gujarati 1995: 456–458), thus detrending time series is an essential operation before causal modeling.[2]

Detrending time series: how do we detrend?

A common method of detrending involves controlling for trends by including a time index as an independent variable in the regression analysis, such as:

$$Y_t = B_0 + B_1 T_1 + \epsilon_t. \tag{2}$$

Equation (2) is a simple linear trend model in which Y at time t is a linear function of time and T is generally initialized to take on a value of 0 when $t = 1$.

Unfortunately if trends are stochastic rather than deterministic, as is often the case, this procedure does not detrend the time series (Raffalovich 1994). If a time series is trending, Y_t will be a function of Y at the previous point in time. Thus an alternative method of detrending a time series is through an autoregressive model in which Y is regressed on (a) previous value(s) of Y:

$$Y_t = B_0 + B_1 Y_{t-1} + \epsilon_t. \tag{3}$$

Allison (1990) refers to Equation (3) as the *regressor variable method*. Through a little algebra, Equation (3) can be written as a (first) difference model:

$$\begin{aligned} \Delta_t &= B_0 + (B_1 - 1) Y_{t-1} + \epsilon_t \\ &= B_0 + \delta Y_{t-1} + \epsilon_t, \end{aligned} \tag{4}$$

where $\Delta_t = Y_t - Y_{t-1}$. Indeed, Equation (4) establishes the well-known Dickey–Fuller test of stationarity in time series. Also referred to as the "unit-root test", stationarity (i.e., the time series is not trending) is indicated when $\delta \neq 0$.

The linear growth model

The study of individual change has benefitted greatly from the development of multilevel models. Beginning with Rao's application of random-coefficient models to the analysis of growth curves, the literature on multilevel models for longitudinal data, repeated measures, and individual growth models has rapidly grown large (Rao 1965; see also, Rogosa et al. 1982; Laird & Ware 1982; Rogosa & Willett 1985; Bryk & Raudenbush 1987, 1992; Karney & Bradbury 1995; Willet et al. 1998). In its simplest form, the linear growth model may be written as:

$$Y_{ti} = \pi_{0i} + \pi_{1i} T_{ti} + \epsilon_{ti}, \tag{5}$$

where i indexes the cross-sectional units (e.g., counties) and t indexes the period (as it did in Equation (2)), $\epsilon_{ti} \sim N(0, \sigma^2)$, and T_{ti} is initialized as 0 for the first measurement. Equation (5) is often referred to as the "within-person" or "level-1" individual growth model. The structural part of the level-1 model contains two unknown constants referred to as individual growth parameters whose values determine the trajectory of "true" individual change over time. In Equation (5), growth is hypothesized to be linear, so π_{0i} represents initial status (the "true" level of Y at time 0) and π_{1i} represents the "true" rate of change in Y over time. A "level-2" model expresses variation in parameters from the growth model as random. This level-2 model is commonly referred to as a "between-person" model and may be written as:

$$\begin{aligned} \pi_{0i} &= \beta_{00} + u_{0i}, \\ &\quad , \quad \text{where} \quad \begin{pmatrix} u_{0i} \\ u_{1i} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{pmatrix} \right]. \\ \pi_{1i} &= \beta_{10} + u_{1i} \end{aligned} \tag{6}$$

As Equation (6) reveals, both the intercept, π_{0i}, and slope, π_{1i}, are treated as random effects with no level-2 covariates in this specification. Equations (5) and (6) can be combined and written as:

$$Y_{ti} = [\beta_{00} + \beta_{10} T_{ti}] + [u_{0i} + u_{1i} T_{ti} + \epsilon_{ti}]. \tag{7}$$

Equation (7) is useful because in it, it is easy to see that in Equations (5) and (6) there is a deterministic component containing two "fixed" coefficients (the intercept, β_{00}, and a slope, β_{10}, for the time trend, T_{ti}) and a nonsystematic component containing three "random" coefficients: for the intercepts (u_{0i}), for the slopes (u_{1i}), and for individual observations taken at time t on each person i (ϵ_{ti}). Interpretation of the fixed coefficients is straightforward: $\hat{\beta}_{00}$ is the estimate of the average intercept across persons when T_{ti} (time) is zero,

and if T_{ti} is initialized at zero, then $\hat{\beta}_{00}$ is the estimate of "initial status" in the growth model. Similarly, $\hat{\beta}_{10}$ is the estimate of the average slope across persons.

Multilevel equations for panel data

The similarity between Equations (2) and (5) is obvious. Indeed the linear growth model in Equation (5) is the *panel series expression for the linear trend model* of Equation (2). As with Equation (3), the linear growth model can be replaced by the regressor variable model that detrends on (a) prior value(s) of Y. Thus the level-1 model of Equation (5) becomes:

$$Y_{ti} = \pi_{0i} + \pi_{1i} Y_{t-1i} + \epsilon_{ti}, \tag{8}$$

as with T_{ti} in Equation (5), Y_{t-1i} is a level-1 ("within-county") covariate and the level-2 model is unrelated to any level-2 covariates:

$$\begin{aligned} \pi_{0i} &= \beta_{00} + u_{0i}, \\ &\quad , \quad \text{where} \quad \begin{pmatrix} u_{0i} \\ u_{1i} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{pmatrix} \right]. \\ \pi_{1i} &= \beta_{10} + u_{1i} \end{aligned} \tag{9}$$

but π_{0i} and π_{1i} are allowed to vary between units.

Equations (8) and (9) reveal the potentially complex error covariance structures implied in multilevel models. Each level equation contains its own error term. In the "within-county" equation (Equation (8), ϵ_{ti} captures prediction error in observations taken at time t on each county i. In the level-2 equations (Equation (9), u_{0i} and u_{1i} represent prediction error in each of the i intercepts and i slopes, respectively. Under the classical mixed model, ϵ_{ti} is assumed to be normal with (homoscedastic) variance given by $\sigma^2 I$, where I is the $t \times t$ identity matrix and the variances of u_{0i}, u_{1i} are elements in a diagonal matrix (given as τ_{00} and τ_{11}, respectively, in Equation (9)). However, the classical mixed model is only a special case of the general mixed model that permits arbitrary parameterized covariance structures of the level-1 and level-2 equations (cf. Wolfinger 1993). Indeed, Equation (9) shows an additional *covariance component*, τ_{01} $(= \tau_{10})$, which allows for correlation between intercepts and slopes. In practice, while many alternative covariance structures are plausible, a preferred structure can be empirically determined by fitting competing representations. Although the choice of covariance structures can be an informative aspect of the multilevel analysis, our preference is to allow a nonzero covariance between τ_{00} and τ_{11} (i.e., $\tau_{01} = \tau_{10} \neq 0$)and, because our data are pooled time series for each county i, the residual observations *within*

counties will be modeled as a first-order autoregressive, or AR(1), process, a common error structure for time series.

Another critical feature of multilevel modeling that must be folded into pooled cross-sectional time series analysis concerns the scaling of independent variables. "Centering decisions" have important consequences for interpretation of parameters and for hypothesis testing. The choice of centering options (raw metric vs. grand mean centering vs. group mean centering) must be determined by the research question under investigation. As (panel series) analogs to methods for (pure) time series, the variance in our outcome measures needs to be partitioned into its within and between components. Only group mean centering will provide the appropriate partitioning and allow *separate* structural models for each component of the variance (Hofmann & Gavin 1998). Thus, Y_{t-1i}, a level-1 (within-county) covariate, in Equation (8) should be replaced by Y^*_{t-1i} ($Y_{t-1i} - \overline{Y}_{\cdot i}$, where $\overline{Y}_{\cdot i}$ is the county-specific mean of the lag term Y_{t-1i}). Henceforth all exogenous (right-hand-side) variables will be "starred" as an indication that they are centered on county-specific means and all exogenous variables are entered in the multilevel models as level-1 (within-county) covariates. Equations (8) and (9) can now be collected and written as:

$$Y_{ti} = [\beta_{00} + \beta_{10} Y^*_{t-1i}] + [u_{0i} + u_{1i} Y^*_{t-1i} + \epsilon_{ti}], \tag{10}$$

and $\hat{\beta}_{00}$ and will give the estimate of the average "condition" across counties (e.g., average number of dust storms in the county time series) and, more importantly, $\hat{\beta}_{10}$ is the estimate of the average change in the time series.

Stationarity

In our discussion above we said that H_0L $\delta = 0$ of Equation (4) provides a statistical test of stationarity.[3] In the multilevel model for panel series, the regressor variable model of Equation (8) can be written as:

$$\Delta_{ti} = \pi_{0i} + \delta_i Y^*_{t-1i} + \epsilon_{ti}, \tag{11}$$

where $\epsilon_{ti} \sim N(0, \Sigma)$ and Σ is the within-county error variance-covariance matrix that models correlation in residual observations as an AR(1) process. However, the test of stationarity is now somewhat more complex than Equation (4) in that H_0: $\beta_{10} = 0$ is found in the level-2 model to Equation (11):

$$\begin{aligned}\sigma_{0i} &= \beta_{00} + u_{0i}, \\ &\quad , \quad \text{where} \quad \begin{pmatrix} u_{0i} \\ u_{1i} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{pmatrix}\right]. \quad (12) \\ \sigma_{1i} &= \beta_{10} + u_{1i}\end{aligned}$$

Since δ_i has both a "fixed" component, β_{10}, and a random component, u_{1i}, it is possible that some individual time series in the panel data are nonstationary (i.e., trending) even if we reject H_0: $\beta_{10} = 0$. A further restriction on Equation (12) then would be to treat δ_i as a constant across counties (as in Equation (4)): $\delta = \beta_{10}$. The preference of δ over δ_i can be formally tested as H_0: $\tau_{11} = 0$ in the error variance-covariance matrix from the level-2 equations seen in Equation (12) (preference of δ also implies $\tau_{10} = \tau_{01} = 0$).

If H_0: $\beta_{10} = 0$ If H_0: is accepted (indicating nonstationarity), the standard procedure is to then assess stationarity in the *velocity* of the time series of Y (measured as the first difference in Y_t), then stationarity in the *acceleration* of the time series of Y (measured as the difference of the difference in Y_t), and so on.[4]

Cointegration between time series

Stationarity in each of the time series is enough to proceed with the causal analysis but it is possible to causally model time series that trend as long as they trend on precisely the same "wave length". In time series that satisfy this condition, the trends cancel out and are said to be *cointegrated*. In short, if the residuals from the regression of Y_t on X_t are stationary, even when the time series of Y_t on X_t are not, then the traditional time series methodology is still applicable.

A simple test of cointegration was proposed by Engle and Granger (1987) as a two-stage extension of the Dickey–Fuller/unit-root test for stationarity. In the first stage regression Y_t is regressed on X_t. In the second stage regression the residuals from the first stage regression, ϵ_t, are differenced and regressed on ϵ_{t-1}, as in the Dickey–Fuller test:

$$\Delta_t = \delta\epsilon_{t-1} + \nu_t, \quad (13)$$

where $\Delta_t = \epsilon_t - \epsilon_{t-1}$. Equation (13) does not include an intercept because $E(\epsilon_t \mid \epsilon_{t-1} = e_{t-1}) = 0$ and this assumption is met by OLS estimation. As in Equation (4), H_0: $\delta = 0$ and rejection indicates stationarity in the residuals which in turn indicates cointegration in the time series.

The multilevel Engle–Granger test of cointegration for panel series begins with the following first stage level-1 model:

$$Y_{ti} = \pi_{0i} + \pi_{1i} X^*_{ti} + \epsilon_{ti}, \tag{14}$$

where $\epsilon_{ti} \sim N(0, \sigma^2)$ and X^*_{ti} is treated as a level-1 (within-county) variable. In the second stage level-1 regression, the residuals from the first stage regression are (first) differenced and regressed on a lagged residual term:

$$\Delta_{ti} + \alpha_{0i} + \delta_I \epsilon^*_{t-1i} + \nu_{ti}, \tag{15}$$

where $\nu_{ti} \sim N(0, \Sigma)$ and Σ is the within-county AR(1) error variance-covariance matrix. Although the expectation of α_{0i} is zero, it is a random variable in Equation (15) and is therefore retained in second stage regression (contrary to the pure time series equivalent in Equation (13). The level-2 model is therefore:

$$\begin{aligned} \alpha_{0i} &= \beta_{00} + u_{0i}, \\ &\quad , \quad \text{where} \quad \begin{pmatrix} e_{0i} \\ e_{1i} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{pmatrix} \right]. \\ \delta_i &= \beta_{10} + e_{1i} \end{aligned} \tag{16}$$

and the constraint to Equation (12) can again be applied as $\delta = \beta_{10}$ in Equation (16). If H_0: $\beta_{10} = 0$ is rejected the time series Y_t and X_t are cointegrated.

Distributed lags in the effect of X_t on Y_t

The final step before specifying the causal models is to investigate the bivariate lag structure. The number of lagged terms to be included in Granger causality tests is an important practical question, with the determination of causality often depending critically on this specification (Gujarati 1995: 622–623). In pure time series analysis the standard strategy is to choose a very high lag order (of X_{t-1}, X_{t-2}, …, X_{t-k}) and assess whether the fit of the model (Y_t on X_{t-1}, X_{t-2}, …, X_{t-k}) deteriorates significantly when this lag length is reduced (cf. Davidson & MacKinnon 1993: 675–676). A formal test of lag length in distributed-lag models was developed by Schwarz (1978). The *Schwarz criterion* (SC) is given as:

$$\text{SC} = \ln \hat{\sigma}^2 + m \ln n, \tag{17}$$

where $\hat{\sigma}^2$ is the estimated variance of the residuals, n is sample size, and m is the lag length (i.e., number of lagged terms in the regression), is minimized with respect to m. SC is essentially the BIC (Bayesian Information Criterion) later popularized by Raftery (1995). What generally recommends SC (and BIC) as a method of model selection is that it minimizes the effect of sample

size while capitalizing on parsimony in model parameters. This works well in pure time series as n is reduced by one with each additional lag term. In pooled cross-sectional time series, however, n is reduced by I (the number of cross-sectional units). In practice we have found this effect on model selection to be severe, leading always to a preferred model with smaller n regardless of the impact of the additional lag term on model fit. We propose an alternative approach.

We regress Y_{ti} on a sequential series of lagged X_{ti} terms using pairs of restricted and unrestricted models with constant n. In this way we assess the impact of an additional lag term on model fit while capitalizing on maximum sample size at each step in the sequence of models. For example, as a restricted level-1 model we may fit:

$$Y_{ti} = \pi_{0i} + \pi_{1i} X^*_{t-1i} + \epsilon_{ti}, \tag{18}$$

we would then fit:

$$Y_{ti} = \pi_{0i} + \pi_{1i} X^*_{t-1i} + \pi_{2i} X^*_{t-2i} + \epsilon_{ti}, \tag{19}$$

as the unrestricted model. Both models have their sample size fixed to the smaller n of Equation (19) and the unrestricted model of Equation (19) has only one additional parameter than the restricted model. More generally, if:

$$Y_{ti} = \pi_{0i} + \sum_{k=1}^{m} \pi_{ki} X^*_{t-ki} + \epsilon_{ti}, \tag{20}$$

is the restricted model, then:

$$Y_{ti} = \gamma_{0i} + \sum_{k=1}^{m+1} \gamma_{ki} X^*_{t-ki} + \epsilon_{ti}, \tag{21}$$

gives the unrestricted model to Equation (20). The distributed lag terms are treated as level-1 covariates (that is, they are "group mean" centered and are not used as predictors of π_{0i} and π_{ki} and γ_{0i} and γ_{ki} in the level-2 models to Equations (20) and (21)) because the causal process of interest is internal to each (pooled) time series. The form of the level-2 equations depend on empirical conclusions from the previous sections. If the restriction on β_{10} in Equation (12) is applicable, the π_{ki} and γ_{ki} in Equations (20) and (21) are "fixed" coefficients and the level-2 equations involve only the estimation of π_{0i} and γ_{0i}. If, on the other hand, δ_i is random, the level-2 equations that follow, although simple in form (i.e., without covariates), will be numerous at higher orders of k. To simplify, the covariance (off-diagonal) components in the variance-covariance matrices may be set to zero. Improvement in model

fit, indicating a preference for the lag structure of the unrestricted model, is given by the difference in the −2LL between the paired models. Since the unrestricted model has one additional parameter (degree of freedom), statistically significant improvement (at the 5% probability level) is indicated by $\Delta_{-2\mathrm{LL}} > 3.84$. It should also be noted that model selection criteria of any kind (e.g., BIC, AIC, or the likelihood ratio tests we employ) for multilevel models with different number of fixed coefficients must be based on maximum likelihood (ML) estimation of the −2LL rather than the restricted maximum likelihood (REML) method that is commonly the "default" method in software for multilevel analysis.

Granger causality tests

As in the time series methods described above, we extend the Wiener–Granger method to panel data through the use of multilevel models. Although the specific form of the autoregressive and distributed lag models will be determined by the empirical conclusions of the previous section, the multilevel models that apply the spirit of Granger causality can be seen in the following level-1 equations. The first stage (level-1) regression is a pooled cross-sectional autoregressive model of the form:

$$Y_{ti} = \pi_{0i} + \sum_{k=1}^{m} \pi_{ki} Y^*_{t-ki} + \epsilon_{ti}. \tag{22}$$

A second stage (level-1) regression then introduces distributed lag covariates of the same order from the cointegrated time series:

$$Y_{ti} = \gamma_{0i} + \sum_{k=1}^{m} \gamma_{ki} Y^*_{t-ki} + \sum_{k=1}^{m} \lambda_{ki} X^*_{t-ki} + \epsilon_{ti}, \tag{23}$$

Both the autoregressive (Y^*_{t-ki}) and distributed lag (Y^*_{t-ki}) terms are treated as level-1 covariates and, as in previous procedures, we have no level-2 covariates (so the level-2 equations continue to look like those of Equation (9)). If π_{ki}, γ_{ki}, and λ_{ki} are random, the (level-1) Equation (22) implies k level-2 equations (and a $k \times k$ variance-covariance error matrix) and Equation (23) has $2k$ level-2 equations (with its $2k \times 2k$ variance-covariance error matrix). Following our earlier observation, the covariance (off-diagonal) components in the variance-covariance matrices may be set to zero. It should also be noted that the extended lag structure of Y^* (indicated by $\sum_{k=1}^{m} \gamma_{ki} Y^*_{t-ki}$) may be used to satisfy the assumptions that $\mathrm{cov}(\epsilon_{ti}, \epsilon_{t+si}) = 0$ and $\epsilon_{ti} \sim N(0, \sigma^2)$ as is done in "augmented" Dickey–Fuller and Engle–Granger tests in (pure) time

series (Gujarati 1995: 720–721, 726–727). According to the logic of Granger causality, evidence that X (Granger) causes Y is provided by preference for the unrestricted model (Equation (23)) over the restricted model (Equation (22)). As in the previous section, this evidence is given by the difference in the goodness-of-fit ($\Delta_{-2\mathrm{LL}}$) of the multilevel models (relative to the differences in their degrees of freedom). Equations (22) and (23) are then repeated where X and Y are reversed in the structural equations in order to determine whether Y (Granger) causes X.

Impulse response functions

It is difficult, if not impossible, to interpret the magnitude and lag structure of the effect of X_t on Y_t in the structural equations of Granger causality models. We can, however, visualize the causal impact through a forecasting method known as *impulse response functions*. Impulse response functions (IRF) operate through shocks to the error term of one of the equations in the dynamic system. In practice the IRF can be explained as follows: suppose one has two reciprocally related structural equations, say, equation A and equation B. In equations A and B the empirically estimated regression slopes (usually retained as standardized coefficients) are numerical constants while their variables, say, X_{tk} and Y_{t-k}, are initialized to 0. A one standard deviation "shock" to the error term of equation A, usually at the maximum time lag in the dynamic system, initiates the forecasting. This shock to equation A is immediately felt in equation B as the predicted value of equation A is substituted into equation B (replacing the initialized value of 0 for X_t if a contemporaneous effect of X is part of structural equation B). The IRF is generated as we bounce back and forth between equations A and B until the initial shock dissipates and the system returns to its equilibrium state (this will happen only if the time series are stationary however).

In general, metric regression slopes in multilevel models may be standardized by multiplying the "fixed" regression slopes by $S_{xti-\bar{x}_t\bullet}/S_{yti-\bar{y}ti\bullet}$, the ratio of the average county sample standard deviation of X at time t to the average sample standard deviation of Y at time t. Our standardized slopes involved an additional transformation of the metric slopes to *percentage change* in X or Y (depending on the position of (log) population on the RHS or LHS of the regression equation) by the appropriate multiplication or division of $\hat{\beta}$ by 100 (see, for instance, Stolzenberg (1980) on standardized coefficients for logarithmic scaled variables). The reader is also reminded that in standardized regressions, the regression line/plane passes through the origin of Y so the intercept, β_{00}, is zero.

In panel data we can either generate IRFs for each cross-sectional unit (using, for example, $\hat{\gamma}_{0i}$, $\hat{\gamma}_{ki}$, and $\hat{\lambda}_{ki}$ from the level-2 equations to Equation

(23)) or we may use the "fixed" coefficients (e.g., $\hat{\beta}_{00}$, $\hat{\beta}_{10}$, ..., $\hat{\beta}_{k0}$ from the level-2 equations to Equation (23)) that produce a single IRF produced from effects averaged across the pooled county time series. For the sake of parsimony, we will simulate the effects from the latter – the fixed regression slopes.

Results: the interrelationship between population and dust storms

The methods for panel data analysis described above are implemented using the county time series of dust storms and (log) population size shown Figures 1 and 2. We begin by assessing stationarity and cointegration in the univariate and bivariate (panel) time series, respectively, we then determine the appropriate distributed-lag structure, and conclude by fitting Granger causality models to the preferred lag structures. Impulse response functions are used to interpret the reciprocating effects of change in population on change in dust storms and change in dust storms on change in population in the long (1962-1988) and short (1969–1988) county time series. All model estimation is by SAS' PROC MIXED – except the IRF simulations which employ PROC MODEL (for documentation on PROC MIXED, see SAS Institute, Inc. 1997: 571–702; for PROC MODEL, see SAS Institute, Inc. 1988: 315–398).

We continue to employ the natural log transformation of population size because it reduces heterogeneity between the county time series and distributional skew within the time series. Another advantage of the log form can be seen in when the regressor variable model of Equation (3) is written as the (first) difference model of Equation (4). The first difference of the untransformed population, $Y_t - Y_{t-1}$, gives the absolute change in Y, the first difference of the log transformed population, $\ln Y_t - \ln Y_{t-1} = \frac{Y_t - Y_{t-1}}{Y_{t-1}}$, gives the relative or proportional change. In the context of substantial differences in the magnitudes of county populations, the relative change measure standardizes for population size while absolute change does not.

Stationarity

The result of fitting the level-1 Equation (11) with and without the restriction on Equation (12) is shown in Table 1. Panel A assesses stationarity in the county dust storm panels for the long (1962–1988) and short (1969–1988) series. Stationarity is supported by *rejection* of the null that $\beta_{10} = 0$. The test statistic is reported as $\hat{\tau}_{\mu}$. Evidence in favor of a "fixed" β_{10} is given in the *failure to reject* the null that $\tau_{11} = 0$. The test statistic is shown as $Z_{\hat{\tau}_{11}/s.e.(\hat{\tau}_{11})}$.

Stationarity in the dust storm time series is clearly supported by the estimates provided in Panel A – although evidence in favor of a fixed coefficient

Table 1. Multilevel unit root test of stationarity for pooled cross-sectional time series

Panel A. Dust storm time series		
	1962–1988 time series ($N = 572$)	1969–1988 time series ($N = 741$)
$\delta_i = \beta_{10} + u_{1i}$:		
	$\hat{\beta}_{10} = -0.64$ $\hat{\tau}_\mu = -6.00^*$	$\hat{\beta}_{10} = -0.73$ $\hat{\tau}_\mu = -13.03^*$
	$Z_{\hat{\tau}_{11}/s.e.(\hat{\tau}_{11})} = 3.98^*$	$Z_{\hat{\tau}_{11}/s.e.(\hat{\tau}_{11})} = 1.51^*$
$\delta = \beta_{10}$:		
	$\hat{\beta}_{10} = -0.60$ $\hat{\tau}_\mu = -15.62^*$	$\hat{\beta}_{10} = -0.64$ $\hat{\tau}_\mu = -18.52^*$
Panel B. Population time series (natural log transformation of population)		
	1962–1988 time series ($N = 572$)	1969–1988 time series ($N = 741$)
$\delta_i = \beta_{10} + u_{1i}$:		
	$\hat{\beta}_{10} = -0.06$ $\hat{\tau}_\mu = -3.39^*$	$\hat{\beta}_{10} = -0.25$ $\hat{\tau}_\mu = -6.20^*$
	$Z_{\hat{\tau}_{11}/s.e.(\hat{\tau}_{11})} = 1.27^*$	$Z_{\hat{\tau}_{11}/s.e.(\hat{\tau}_{11})} = 3.32^*$
$\delta = \beta_{10}$:		
	$\hat{\beta}_{10} = -0.05$ $\hat{\tau}_\mu = -4.08^*$	$\hat{\beta}_{10} = -0.08$ $\hat{\tau}_\mu = -6.81^*$
Panel C. Population change time series (natural log transformation of population)		
	1962–1988 time series ($N = 550$)	1969–1988 time series ($N = 702$)
$\delta_i = \beta_{10} + u_{1i}$:		
	$\hat{\beta}_{10} = -0.76$ $\hat{\tau}_\mu = -6.66^*$	$\hat{\beta}_{10} = -0.57$ $\hat{\tau}_\mu = -11.50^*$
	$Z_{\hat{\tau}_{11}/s.e.(\hat{\tau}_{11})} = 3.71^*$	$Z_{\hat{\tau}_{11}/s.e.(\hat{\tau}_{11})} = 2.52^*$
$\delta = \beta_{10}$:		
	$\hat{\beta}_{10} = -0.76$ $\hat{\tau}_\mu = -18.17^*$	$\hat{\beta}_{10} = -0.54$ $\hat{\tau}_\mu = -15.03^*$

The 1%, 5%, and 10% critical values (τ_μ) for the null hypothesis that δ (or δ_i) $= 0$, at $n = \infty$, are -3.42, -2.86, and -2.57 (Fuller 1996, Table 10.A.2, second panel). $Z_{\hat{\tau}_{11}/s.e.(\hat{\tau}_{11})}$ is the ratio of the variance component estimate for π_{1i} ($\hat{\tau}_{11}$) to its estimated standard error (this ratio is standard normal distributed). Stars (*) indicate rejection of H_0: $\beta_{10} = 0$ and H_0: $\tau_{11} = 0$ at 1% probability. Absolute values of the critical values for τ_μ are applied in Panel C.

is not supported in the long (but less cross-sectionally rich) 1962–1988 time series.

Evidence in favor of stationarity in the (log) population county time series is slightly more ambiguous. Panel B does support stationarity in the 1969–1988 time series but we must reject the null that $\tau_{11} = 0$ so we must continue to entertain the possibility that individual county time series may not be stationary. When we consider the 1962–1988 time series we observe stationarity only *when the constant effect restriction is in place* – although our failure to reject the null on τ_{11} does support that restriction. Given the qualified support in favor of stationarity in Panel B, we investigate stationarity in the first difference of (log) population. If the change in population size is not constant, perhaps the change in the change in population size is constant (i.e.,

Table 2. Multilevel Engle–Granger test of cointegration for pooled cross-sectional time series

Second stage regression of $\Delta_t (= \epsilon_t - \epsilon_{t-1})$ on ϵ^*_{t-1i}	
1962–1988 time series ($N = 572$)	1969–1988 time series ($N = 741$)
$\delta_i = \beta_{10} + u_{1i}$:	$\delta_i = \beta_{10} + u_{1i}$:
$\hat{\beta}_{10} = -0.065 \quad \hat{\tau}_\mu = -3.89^*$	$\hat{\beta}_{10} = -0.101 \quad \hat{\tau}_\mu = -6.26^*$
$Z_{\hat{\tau}_{11}/s.e.(\hat{\tau}_{11})} = 1.55^*$	$Z_{\hat{\tau}_{11}/s.e.(\hat{\tau}_{11})} = 1.85^*$
$\delta = \beta_{10}$:	$\delta = \beta_{10}$:
$\hat{\beta}_{10} = -0.057 \quad \hat{\tau}_\mu = -4.50^*$	$\hat{\beta}_{10} = -0.067 \quad \hat{\tau}_\mu = -6.45^*$

First stage regression is (log) population size on dust storm county time series. Since β_{00} is to be estimated (see Equation (16)) we retain the same 1%, 5%, and 10% critical values (τ_μ) for the null hypothesis that δ (or δ_i) $= 0$ given in Table 1. $Z_{\hat{\tau}_{11}/s.e.(\hat{\tau}_{11})}$ is also defined as in Table 1. Stars (*) indicate rejection of H_0: $\beta_{10} = 0$ and H_0: $\tau_{11} = 0$ at 1% probability.

the "velocity" of the time series is constant). Indeed, Panel C does evidence stationarity in the long and the short differenced time series – but we cannot accept the null that $\tau_{11} = 0$ in either series.

Cointegration

However, before we leave the population size form in favor of the differenced time series, we can (and should) investigate cointegration between the dust storm and (log) population size time series. As we noted earlier, it is possible that time series that trend may trend together, in which case methods for assessing causality will still be appropriate. Table 2 presents results from our multilevel test of cointegration.

The evidence in Table 2 is not at all equivocal. The multi-level Engle–Granger tests support stationarity and a "fixed" constraint on β_{10} in the 1962–1988 and 1969–1988 time series. As a consequence, we apply the fixed coefficient restriction on slopes of all exogenous terms in our subsequent tests of distributed lag structure and Granger causality and we investigate the causal relationship between dust storms and (log) population size.

Distributed-lag structure

Figures 3 and 4 display the improvement in model fit ($\Delta_{-2\mathrm{LL}}$) in the unrestricted model of Equation (21) over the restricted model of Equation (20) for the 1962–1988 and 1969–1988 time series. A difference in the -2LL above the critical value of 3.84 (shown as the horizontal line in the figures) indicates a preference in the model fit of the lag structure of the unrestricted model.

Figure 3, constructed from the $\Delta_{-2\mathrm{LL}}$ between unrestricted and restricted regressions of dust storms at time t on (log) population at $t - 1$, $t - 2$, ...,

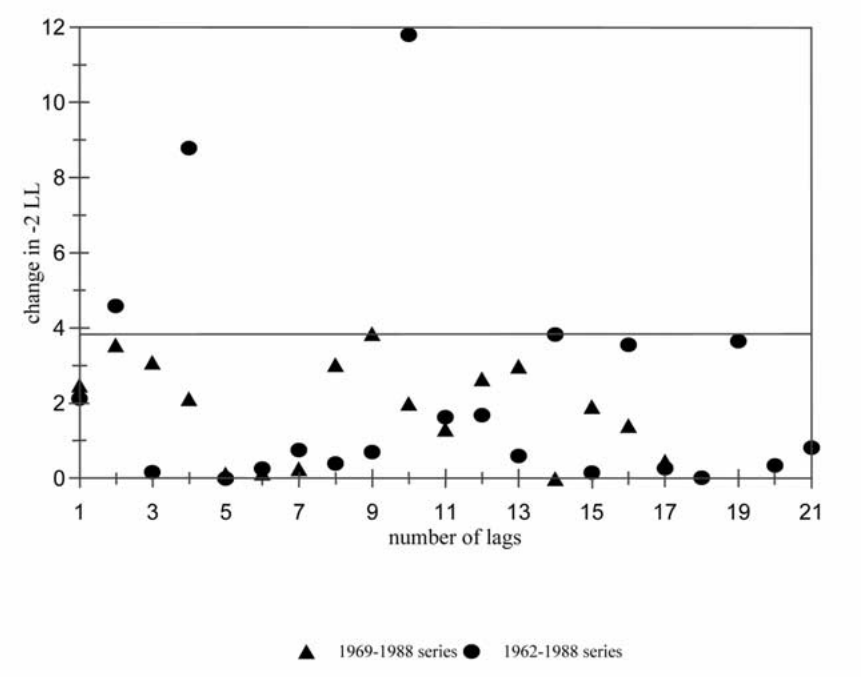


Figure 3. Distributed lag structure dust storms = (log) population.

$t - k$, evidences four distributed lag regressions with statistically significant improvement in model fit in the 1962–1988 time series. Only one short (1969–1988) series model attains statistical significance (at a 5% probability). Figure 4 repeats this exercise for regressions of (log) population on dust storms. Here, the pattern is reversed as nine 1969–1988 time series attain statistical significance in improvement in model fit, yet only three long (1962–1988) series models show a change in -2LL above the critical value.

Based on this information our preferred lag structures to be used in the Granger causality models will maximize lag length in the 1962–1988 time series while building a shorter lag model around the single statistically significant model improvement in the 1969–1988 series regression of dust storms on (log) population. This strategy results in a 14-year lag structure for the 1962-1988 series regression of dust storms on (log) population, a 15-year lag structure for the 1962–1988 series regression of (log) population on dust storms, a 9-year lag structure for the 1969–1988 series regression of dust storms on (log) population, and an 11-year lag structure for the short series regression of (log) population on dust storms. The first three choices employ maximum (statistically significant) lag structures as shown in the figures. The final choice of lag length (the 11-year lag for the 1969–1988 (log) population

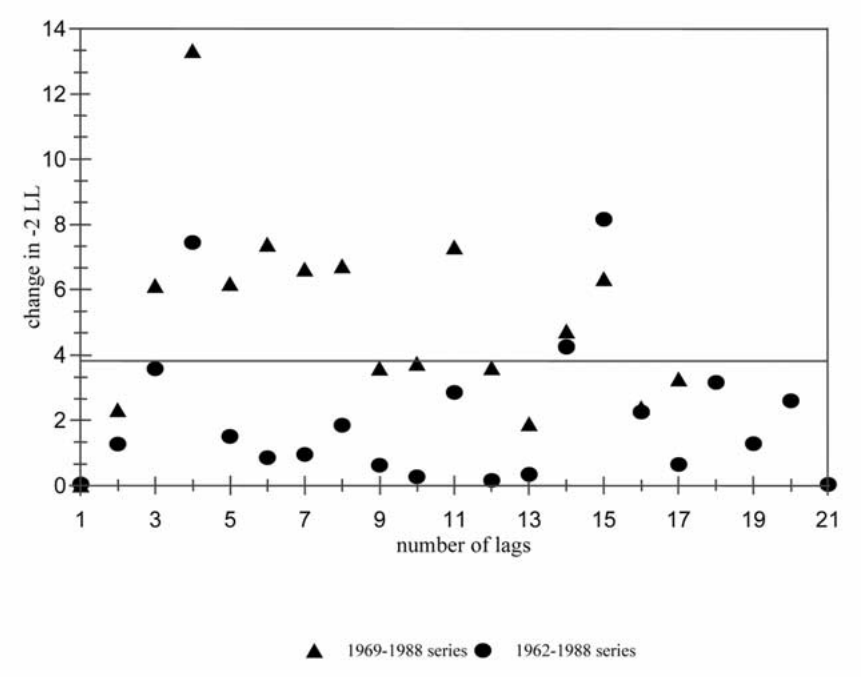


Figure 4. Distributed lag structure (log) population = dust storms.

regression) is motivated by fitting a shorter lag structure to the shorter time series - which is consistent with the single choice of lag length available for the dust storm regression.

Although our choice of lag structures is empirically derived, the advantage of this outcome is that we can use the differences in lag structures (that is, high order lag length in the long time series and more moderate lag length in the short time series) to assess robustness in our determination of Granger causality. In addition, it is important to recognize that the distributed-lag structure determined here is intended to ensure that the lagged effect of population on dust storms and dust storms on population is not prematurely truncated in the Granger causality tests. It is quite probable that the choice of the long lag structure derives more from the number of cross-sectional units and less from the scant information left in the time series at these long lag lengths. This is not particularly troublesome for two reasons. First, the longer lags with simply enter the Granger models as "irrelevant" included variables, while the shorter lags will capture the deterministic lag structure – in which case we are proceeding under the procedure advocated by Davidson and MacKinnon (1993). Second, if statistical evidence of a feedback system is determined, the lag structure of this dynamic system will not necessarily correspond to

the recursive distributed-lag structure found here. This remains to be shown by the impulse response functions.

It is, however, worth noting at this point that previous research by economic historians on short-term and long-term fluctuations in climate and population is consistent with the extended lag structures we find here (Lee 1981; Galloway 1985, 1986). Although this research used much shorter distributed-lags (up to four years) than we find here, the time series were detrended around an 11-year moving average, the net result of which is that the temporal reach of climatic conditions on their population outcomes can be quite long.

Granger causality models

Each of the stages in the analysis thus far have been simply diagnostics to our Granger causality models for panel series. Our ultimate interest has always been in determining the causal relationship between population and the environment. Table 3 presents the empirical evidence of this relationship.

Using the preferred distributed lag structures from above we assess Granger causality in the 14- and 15-year lag models for the 1962–1988 time series and in the 9- and 11-year lag models for the 1969–1988 time series. Panel A reports results from the regressions of storm storms on (log) population. We then reverse the causal order by regressing (log) population on dust storms and display the results in Panel B. The summary evidence is given in the $\Delta_{-2\mathrm{LL}}$ (the improvement in model fit of the second stage regression over the first stage regression) relative to Δ_{fc} (the difference in degrees of freedom in the regressions). In each case, the improvement in model fit is statistically significant at (at least) a 5% probability. In other words, the addition of information from the bivariate panels improves prediction of the event – either number of dust storms (Panel A) or size of population (Panel B) – above and beyond what is expected from prior occurrences of that event. Table 3 clearly points to bilateral causality between the time series – (log) population size (Granger) causes dust storms and dust storms (Granger) cause (log) population size.

Impulse Response Functions

Because of the difficulty in interpreting the magnitude, the lag structure, and even the direction of effects in dynamically interrelated autoregressive and distributed lag models, we omitted the structural coefficients from Table 3. We can, however, visualize the causal impact through *impulse response functions*. Figures 5 and 6 illustrate these reciprocal impacts in the 1962–1988 and 1969–1988 time series. In both figures, the solid line forecasts the

Table 3. First and second stage granger models for pooled cross-sectional time series. Dust storms = Y_{ti}; (log) population = X_{ti}

1962–1988 time series ($N = 286$)	1969–1988 time series ($N = 429$)
Panel A. Dust storms	
$Y_{ti} = \pi_{0i} + \sum_{k=1}^{14} \pi_{ki} Y^*_{t-ki} + \epsilon_{ti}$	$Y_{ti} = \pi_{0i} + \sum_{k=1}^{9} \pi_{ki} Y^*_{t-ki} + \epsilon_{ti}$
$-2\text{LL} = 1828.4$ #f.c. = 15	$-2\text{LL} = 2519.2$ #f.c. = 10
$Y_{ti} = \gamma_{0i} + \sum_{k=1}^{14} \gamma_{ki} Y^*_{t-ki} + \sum_{k=1}^{14} \lambda_{ki} X^*_{t-ki} + \epsilon_{ti}$	$Y_{ti} = \gamma_{0i} + \sum_{k=1}^{9} \gamma_{ki} Y^*_{t-ki} + \sum_{k=1}^{9} \lambda_{ki} X^*_{t-ki} + \epsilon_{ti}$
$-2\text{LL} = 1795.7$ #f.c. = 29	$-2\text{LL} = 2500.7$ #f.c. = 19
$\Delta_{-2\text{LL}} = 32.7^{**}$ $\Delta_{fc} = 14$	$\Delta_{-2\text{LL}} = 18.5^{*}$ $\Delta_{fc} = 9$
Panel B. (log) population	
1962–1988 time series ($N = 264$)	1969–1988 time series ($N = 351$)
$X_{ti} = \pi_{0i} + \sum_{k=1}^{15} \pi_{ki} Y^*_{t-ki} + \epsilon_{ti}$	$X_{ti} = \pi_{0i} + \sum_{k=1}^{11} \pi_{ki} Y^*_{t-ki} + \epsilon_{ti}$
$-2\text{LL} = 1254.4$ #f.c. = 16	$-2\text{LL} = 1672.0$ #f.c. = 12
$X_{ti} = \gamma_{0i} + \sum_{k=1}^{15} \gamma_{ki} Y^*_{t-ki} + \sum_{k=1}^{15} \lambda_{ki} X^*_{t-ki} + \epsilon_{ti}$	$X_{ti} = \gamma_{0i} + \sum_{k=1}^{11} \gamma_{ki} Y^*_{t-ki} + \sum_{k=1}^{11} \lambda_{ki} X^*_{t-ki} + \epsilon_{ti}$
$-2\text{LL} = 1279.7$ #f.c. = 31	$-2\text{LL} = 1698.3$ #f.c. = 23
$\Delta_{-2\text{LL}} = 25.3^{*}$ $\Delta_{fc} = 15$	$\Delta_{-2\text{LL}} = 28.3^{**}$ $\Delta_{fc} = 11$

Order of lag structure is identified over the summations ($\sum$) in the first and second stage regressions. #f.c. is the number of "fixed" coefficients in the regression model. Star (*) indicates rejection of H_0: $\Delta_{-2\text{LL}} = 0$ at 5% probability in χ^2 sampling distribution for d.f. $= \Delta_{fc}$; two stars (**) indicate rejection of null at 1% probability.

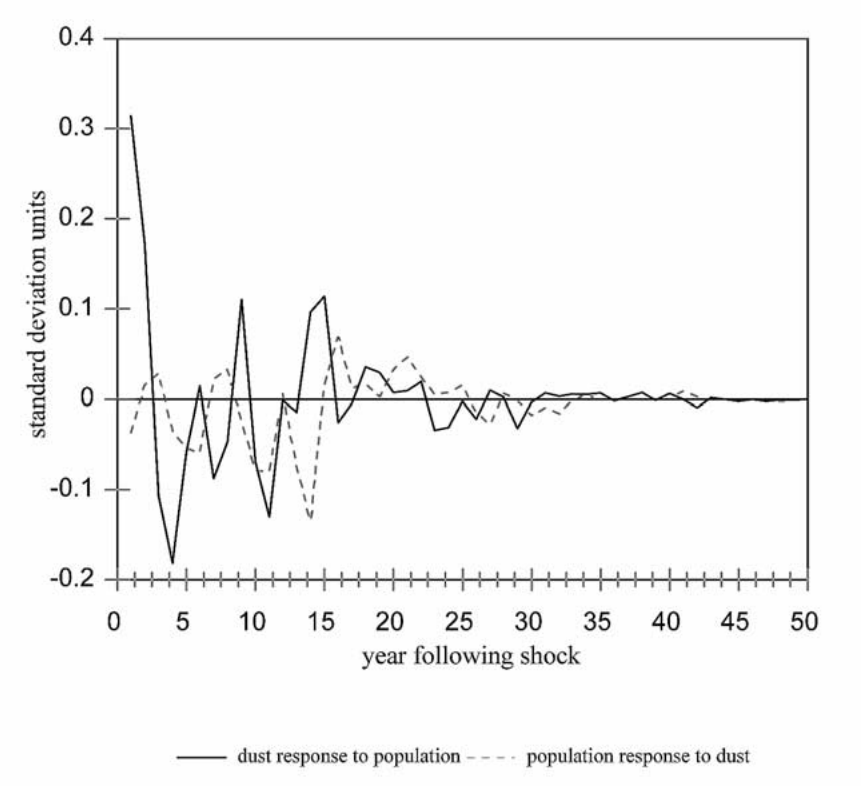


Figure 5. Impulse response functions. 1 s.d. shock to 1962–1988 time series.

response pattern of dust storms to a (one standard deviation) shock in the (log) population size series and the dashed line maps the population response to a shock in the dust storm series.

The response patterns in the longer (1962–1988) time series of Figure 5 are particularly impressive. Focusing first on the population response to dust storms, if a county witnesses a large increase in dust storms, its population falls into an approximately 15-year trough before it recovers slightly and then the shock leaves the (county time series) system. This response pattern is of nearly identical duration to that given by the search for a distributed-lag structure. On the other hand, if a county's population were to suddenly increase, Figure 5 shows that that increase would be followed by an almost immediate and dramatic increase in dust storms of several years duration before bouncing around between net decreases and net increases and leaving the system. This response pattern shows that when treated as a dynamic system the lagged impact of (log) population on dust storms is of much shorter duration than that suggested by the recursive (14-year) distributed-lag model.

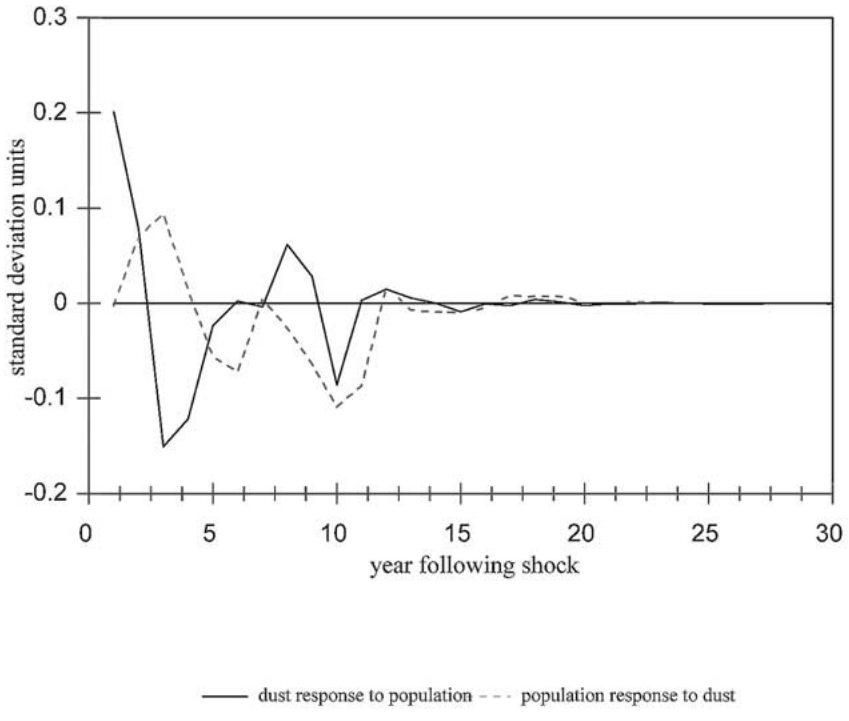


Figure 6. Impulse response functions. 1 s.d. shock to 1969–1988 time series.

Figure 6 gives the response patterns in the shorter (1969–1988) times series with more cross-sectional (county) units. The response of dust storms to a shock in (log) population is nearly identical to that seen in Figure 5. Although less definitive of the anticipated direction and strength of response, we still see in Figure 6 the extended trough in the population response to a shock in dust storms – though the extended decline in population does not begin until fourth year following the intervention. Both systems return to equilibrium in about twelve years.

Discussion

In the introduction to this paper we promised two things: (1) to test directly the interrelationship between population and environmental change, and (2) to do so through the development of a hybrid method that combines a sequence of well-known techniques for (pure) time series analysis with a new, but widely accessible, approach known as multilevel modeling. Both of these achievements deserve further comment.

We have already acknowledged that the Wiener–Granger method is a necessary but not a sufficient demonstration of causality in bivariate time series. Indeed, regardless of how literal Woody Guthrie may have intended the phrase "Blowin' Down the Road" to be, we use it as allegory, rather than fact. Dust storms do not blow people out of their counties, but they do ruin their crops, precipitate foreclosures on their farms, weaken their local economies, and make opportunities elsewhere seem more attractive. Likewise, dust storms are not the result of increasing numbers of people stamping their feet in concert until the ground loosens and swirls away in a Spring or Autumn gust, but there is a well-documented human antecedent to be found in land use practices and urban and industrial development.

We are aware that in directly linking dust storms to population size we are perhaps asserting a broader definition of causality than some of our readers will be willing to accept. On the other hand, dust storms are a concrete and direct measure of an (undesirable) environmental quality and surely there is something elegant in measuring the concept of population as population size. Our preference has been to start simply, and on this score, it is clear that population and environmental change is reciprocal - with the specifics being told in the tables and figures above.

We have also shown how a well-developed branch of hierarchical linear modeling, the linear growth model, converges with the methodology of (pure) time series analysis. We are well aware of the attention econometricians have given to panel data analysis and that efforts have been made previously to bring this literature together with the (pure) time series techniques cited in this paper (cf. Pakes & Griliches 1984; Holtz-Eakin et al. 1988). We believe that what recommends our approach is, similar to Empirical Bayes (EB) estimation, multilevel modeling is a way of "borrowing strength" in order to obtain improved estimates of individual effects. The concept of "borrowing strength" is useful in situations where sparse data are a problem for the prediction of the behavior (Kreft & de Leeuw 1998). In panel data analysis, data is generally sparse in either the time or cross-sectional dimension as one often has relatively few time points but many cross-sectional units or relatively few cross-sectional units but many time points. The solution for this problem is solved by EB/ML estimation, using all information over cross-sectional units and time together. In our application, the prediction for each county separately will be a mixture of the relationship in its own data and the relationship in the total or pooled data. In EB/ML estimation all estimates are more or less shrunken to the mean, thus by estimating all county time series together, the time series are "borrowing strength" from the pooled data.

Given the wide availability of software for multilevel modeling and the rapidly growing expertise in a number of disciplines for applying that soft-

ware, our approach opens a path along which researchers well-versed in growth curve analysis can move seamlessly into a powerful set of methods largely under the purview of econometricians.

Notes

1. The method is described in a hypertext file available on the Census Bureau webpage at http://www.census.gov/population/methods/stco99.txt. The population estimates we use were downloaded from the Bureau of Economic Analysis – Regional Economic Information System's "Personal Income, Total Income, & Per Capital Personal Income by County and Metropolitan Area: 1969–96" file available at http://www.bea.doc.gov/bea/regional/reis/. The annual population estimates in this file were derived by the Bureau of the Census.
2. Most of our discussion of time series methods in this section was informed by Gujarati (1995, esp. chapters 17, 21, and 22) and Mills (1990). In general specific citations will not be provided as this material is widely available in standard econometric texts.
3. In a regression of the form given by Equation (4) (and by extension, Equations (11)–(12), critical values of the test statistic, ($\hat{\tau}_\mu$), for a null hypothesis that $\beta_{10} = 0$ are given in Fuller (1996, Table 10.A.2, second panel). The 1%, 5%, and 10% critical τ_μ values, at $n = \infty$, are -3.42, -2.86, and -2.57, respectively. Obviously these are considerably larger (in absolute value) than the t-scores used in hypothesis testing in ordinary linear regression. For discussion, see Fuller (1996, chapter 10).
4. Stationarity in the velocity of a time series is indicated by rejection of H_0: $\beta_{10} = 0$ in a level-2 equation of the same form as given in Equation (12) (i.e., $\delta_i = \beta_{10} + u_{1i}$) but in which the level-1 equation is:

$$\Delta D_{ti} = \pi_{0i} + \delta_i D^*_{t-1i} + \epsilon_{ti},$$

where D_{ti}, for convenience, denotes $\Delta Y_{ti} = (Y_{ti} - Y_{t-1i})$.

References

Adger, W.N. & Brown, K. (1994), Land use and the causes of global warming, Chichester: John Wiley.

Allison, P. (1993), Change scores as dependent variables in regression analysis, pp. 93–114 in P.V. Marsden (ed.), *Sociological Methodology*, Washington D.C.: American Sociological Association.

Behairy, A.K.A., El-Sayed, M.K., & Rao, N.V.D. (1985), Eolian dust in the coastal area north of Jeddah, Saudi Arabia, *Journal of Arid Environments* 8: 89–98.

Biggar, J.C. (1979), The sunning of America: Migration to the sunbelt, *Population Bulletin* 34(1): 1–42.

Bonnifield, P. (1979), *The Dust Bowl: Men, Dirt, and Depression*, Albuquerque: University of New Mexico Press.

Bratton, T.L. (1988), The identity of the New England Indian epidemic of 1616–19, *Bulletin of the History of Medicine* 62: 351–383.

Bryk, A.S. & Raudenbush, S.W. (1987), Application of hierarchical linear models to assessing change, *Psychological Bulletin* 101: 147–158.

Bryk, A.S. & Raudenbush, S.W. (1992), *Hierarchical Linear Models: Applications and Data Analysis Methods*, Newbury Park: Sage.

Chepil, W.S., Siddoway, F.H., & Armbrust, D.V. (1963), Climatic index of wind erosion conditions in the Great Plains, *Soil Science Society of America Proceedings* 27: 449–452.

Davidson, R. & MacKinnon, J.G. (1993), *Estimation and Inference in Econometrics*, New York: Oxford University Press.

Diamond, J. (1997), *Guns, Germs, and Steel*, New York: W.W. Norton & Company.

Dickinson, P.A. (1978), Retirement in the sunbelt, *Modern Maturity* (October and November): 11–15.

Engle, R.F. & Granger, C.W.J. (1987), Co-integration and error correction: Representation, estimation, and testing, *Econometrica* 55: 251–276.

Ervin, R.T. & Lee, J.A. (1994), Impact of conservation practices on airborne dust in the southern high plains of Texas, *Journal of Soil and Water Conservation* 49: 430–437.

Finnell, H.H. (1948), The dust storms of 1948, *Scientific American* 179(2): 7–11.

Finnell, H.H. (1954), The dust storms of 1954, *Scientific American* 191(1): 25–29.

Fuller, W.A. (1996), *Introduction to Statistical Time Series*, 2nd edn, New York: John Wiley.

Galloway, P.R. (1985), Annual variations in deaths by age, deaths by cause, prices, and weather in London 1670 to 1830, *Population Studies* 39: 487–505.

Galloway, P.R. (1986), Long-term fluctuations in climate and population in the preindustrial era, *Population and Development Review* 12: 1–24.

Goudie, A.S. & Middleton, H.J. (1992), The changing frequency of dust storms through time, *Climatic Change* 20: 197–225.

Granger, C.W.J. (1969), Investigating causal relations by econometric models and cross-spectral methods, *Econometrica* 37: 424–438.

Gujarati, D.N. (1995), *Basic econometrics*, 3rd edn, New York: McGraw-Hill.

Hofmann, D.A. & Gavin, M.B. (1998), Centering decisions in hierarchical linear models: implications for research in organizations, *Journal of Management* 24: 623–641.

Holtz-Eakin, D., Newey, W., & Rosen, H.S. (1988), Estimating vector autoregressions with panel data, *Econometrica* 56: 1371–1395.

Jauregui, E. (1989), The dust storms of Mexico City, *International Journal of Climatology* 9: 169–180.

Karl, T.R., Williams, Jr., C.N., Quinlan, F.T., & Boden, T.A. (1990), United States historical climatology network (HCN) serial temperature and precipitation data, Environmental Science Division, Publication No. 3404, Carbon Dioxide Information and Analysis Center, Oak Ridge National Laboratory, Oak Ridge, TN.

Karney, B.R. & Bradbury, T.N. (1995), Assessing longitudinal change in marriage: An introduction to the analysis of growth curves, *Journal of Marriage and the Family* 57: 1091–1108.

Kreft, I.G.G. & de Leeuw, J. (1998), *Introducing Multilevel Modeling*, Thousand Oaks, CA.: Sage.

Laird, N.M. & Ware, J.H. (1982), Random effects models for longitudinal data, *Biometrics* 38: 963–974.

Lee, J.A. & Tchakerian, V.P. (1995), Magnitude and frequency of blowing dust on the southern high plains of the United States, 1947–1989, *Annals of the Association of American Geographers* 85: 684–693.

Lee, R.D. (1981), Short-run fluctuations of vital rates, prices and weather in England, 1539 to 1840, pp. 356–401 in E.A. Wrigley & R. Schofield (eds.), *The Population History of England, 1541–1871: A Reconstruction*, Cambridge, MA: Harvard University Press.

Long, L.E. (1987), *Migration and Residential Mobility in the United States*, New York: Russell Sage Foundation.
Marsh, G.P. (1874), *The Earth as Modified by Human Action*, New York: Scribner, Armstrong & Company.
McNeill, W.H. (1976), *Plagues and Peoples*, New York: Doubleday.
Meyer, W.B. & Turner II, B.L. (1994), *Changes in Land Use and Land Cover: A Global Perspective*, Cambridge: Cambridge University Press.
Mills, T.C. (1990), *Time Series Techniques for Economists*, Cambridge: Cambridge University Press.
National Academy of Sciences (1993), *Population and Land Use in Developing Countries*, Washington, D.C.: National Academy Press.
Pakes, A. & Griliches, Z. (1984), Estimating distributed lags in short panels with an application to the specification of depreciation patterns and capital constructs, *Review of Economic Studies* 243–262.
Perlin, J. (1989), *A Forest Journey: The Role of Wood in the Development of Civilization*, New York: W.W. Norton.
Preston, S.H. (1993), The contours of demography: Estimates and projections, *Demography* 30: 593–606.
Raffalovich, L.E. (1994), Detrending time series: A cautionary note, *Sociological Methods and Research* 22: 492–519.
Raftery, A.E. (1995), Bayesian model selection in social research, pp. 111–163 in P.V. Marsden (ed.), *Sociological Methodology*, Cambridge, MA: Blackwell.
Rao, C.R. (1965), The theory of least squares when parameters are stochastic and its application to the analysis of growth curves, *Biometrika* 52: 447–458.
Richards, J.F. & Tucker, R.P., eds. (1988), *World Deforestation in the Twentieth Century*, Durham, N.C.: Duke University Press.
Rogosa, D.R., Brandt, D., & Zimowski, M. (1982), A growth curve approach to the measurement of change, *Psychological Bulletin* 90: 726–748.
Rogosa, D.R. & Willett, J.B. (1985), Understanding correlates of change by modeling individual differences in growth, *Psychometrika* 50: 203–228.
SAS Institute, Inc. (1988), *SAS/ETS User's Guide, Version 6*, 1st edn, Cary, N.C.: SAS Institute, Inc.
SAS Institute, Inc. (1997), *SAS/STAT Software: Changes and Enhancements through Release 6.12*, Cary, N.C.: SAS Institute, Inc.
Sauer, C.O. (1963), *Land and Life: A Selection from the Writings of Carl Ortwin Sauer*, Berkeley: University of California Press.
Schwarz, G. (1978), Estimating the Dimension of a Model, *Annals of Statistics* 6: 461–464.
Sims, C.A. (1972), Money, income, and causality, *American Economic Review* 62: 540–552.
Sims, C.A. (1980), Macroeconomics and reality, *Econometrica* 48: 1–48.
Steward, J.H. (1978), Initiation of a research trend: Wittfogel's irrigation Hypothesis, pp. 3–14 in G.L. Ulmen (ed.), *Society and History: Essays in Honor of Karl August Wittfogel*, The Hague: Mouton Publishers.
Stolzenberg, R.M. (1980), The measurement and decomposition of causal effects in nonlinear and nonadditive models, pp. 459–488 in K.F. Schuessler (ed.), *Sociological Methodology*, San Francisco: Jossey-Bass.
U.S. Dept. of Commerce [Bureau of the Census] (1967), *County and City Data Book*, Washington: Supt. of Docs., U.S. Govt. Print.

Willett, J.B., Singer, J.D., & Martin, N.C. (1998), The design and analysis of longitudinal studies of development and psychopathology in context: Statistical models and methodological recommendations, *Development and Psychology* 10: 395–426.

Wittfogel, K.A. (1935), The foundations and stages of Chinese economic history, *Zeitschrift für Sozialforschung* 4: 26–60.

Wittfogel, K.A. (1939–1940), The society of prehistoric China, *Studies in Philosophy and Social Science* 8: 138–186.

Wittfogel, K.A. (1957), *Oriental Despotism; A Comparative Study of Total Power*, New Haven: Yale University Press.

Wolfinger, R. (1993), Covariance structure selection in general mixed models, *Communications in Statistics – Simulations* 22: 1079–1106.

Worster, D. (1979), *Dust Bowl: The Southern Plains in the 1930s*, New York: Oxford University Press.

Address for correspondence: Glenn D. Deane. Department of Sociology, University at Albany, Albany, NY 12222, USA
Phone: 518-442-4587; Fax: 518-442-4936;
E-mail: gdd@albany.edu