Geography and macroeconomics: New data and new findings

William D. Nordhaus*

Yale University, 28 Hillhouse Avenue, New Haven, CT 06520-8268

This contribution is part of the special series of Inaugural Articles by members of the National Academy of Sciences elected on May 1, 2001.

Contributed by William D. Nordhaus, December 2, 2005

The linkage between economic activity and geography is obvious: Populations cluster mainly on coasts and rarely on ice sheets. Past studies of the relationships between economic activity and geography have been hampered by limited spatial data on economic activity. The present study introduces data on global economic activity, the G-Econ database, which measures economic activity for all large countries, measured at a 1° latitude by 1° longitude scale. The methodologies for the study are described. Three applications of the data are investigated. First, the puzzling ''climate-output reversal'' is detected, whereby the relationship between temperature and output is negative when measured on a per capita basis and strongly positive on a per area basis. Second, the database allows better resolution of the impact of geographic attributes on African poverty, finding geography is an important source of income differences relative to high-income regions. Finally, we use the G-Econ data to provide estimates of the economic impact of greenhouse warming, with larger estimates of warming damages than past studies.

economic growth | development | climate change

The linkage between economic activity and geography is obvious to most people: populations cluster mainly on coasts and rarely on ice sheets. Yet, modern macroeconomics and growth economics generally ignore geographic factors such as climate, proximity to water, soils, tropical pests, and permafrost. This inaugural essay examines this intellectual division, presents data on geographically based economic activity, and examines some of the major relationships between macroeconomic activity and geographic measures. A full description of the data and methods can be found at the project web site (http://gecon.yale.edu).

Why has macroeconomics generally ignored geography? As will be discussed in subsequent sections, three factors have prevented a thorough integration of geographic factors into macroeconomic analysis. First, economic growth theory has emphasized the role of endogenous and policy factors, such as capital formation, education, and technology, rather than exogenous factors such as geography or even population. Although natural resources (particularly land and minerals) have been featured in some studies, climate, soils, tropical diseases, and similar "unchanging" factors have typically been omitted from modern economic growth analysis.

Second, studies of the impact of geography on economic activity have emphasized the level or growth in per capita output. Although this focus is sensible for a discipline like economics, which focuses on national economic policies and living standards, it is difficult to capture time-invariant geographic factors in such studies. To separate out geographic factors, this study examines areal density of output and per capita output. We will see that shifting the measure dramatically changes the estimated effects of geography on economic activity.

Third, most measures of economic activity have been time series or panels measured at the level of the country, which provide ≈100 observations at enormously different geographic scales. The data set presented here (GECON 1.1), which measures output with a resolution of 1° latitude by 1° longitude, covers 25,572 terrestrial grid cells. Such an increase in resolution is analogous to pictures from the Hubble telescope, which provide clear and crisp answers to many previously difficult and fuzzily answered questions.

The change in emphasis proposed here has an enormous effect on the estimated impact of geographic attributes on economic activity. The G-Econ database (described in detail in the second part of this article) can be useful not only for economists interested in spatial economics but equally for environmental scientists looking to link their satellite and other geographically based data with economic data.

I begin with a brief survey of the role of geographic factors in economic analysis and empirical work. In this survey, I will discuss mainly macroeconomics, and it must be emphasized that these remarks present a highly condensed view of studies that relate to global economic processes. The vast and impressive literature in geography and regional economics is largely outside the scope of this study.

It will be useful to state what I mean by "geographical" factors (or, better, geophysical factors studied in "physical geography"). These physical attributes are tied to specific locations. They may be nonstochastic on the relevant time scale (such as latitude, distance from coastlines, or elevation) or they may be stochastic with slowly moving means and variability (such as climate or soils). One of the critical features of the present approach is that geographic factors are statistically exogenous in the sense that they cause, but to a first approximation are not caused by, economic and other social variables. For our purposes, we omit most environmental and endogenous geographic variables, such as pollution, land use, and the natural-resource content of trade or output. Although these factors are of critical importance for many purposes, the focus here is on exogenous and large-scale factors that are largely unaffected by human activities on decadal time scales.

In reflecting on the wealth of nations, early economists assumed that climate was one of the prime determinants of national differences. In societies where most of the population lived on farms, this presumption was probably correct. Earlier civilizations, such as those investigated in Landes's history of economic growth (1) or Diamond's analysis of societal collapses (2), were highly dependent on local resources and climatic conditions and less able to specialize and trade than most economies today.

However, one of the major factors in economic development has been the movement from climatic-sensitive farming and into climate-insensitive manufacturing and services. In 1820, 72% of U.S. employment was on farms, whereas by 2004, the share was down to 1.2%. Many studies suggest that the market economy in the developed world is relatively insensitive to moderate and gradual changes in climate or similar geographic conditions (see below).

Current theories and empirical studies of economic growth today give short shrift to climate as the basis for the differences in the

Conflict of interest statement: No conflicts declared.

Freely available online through the PNAS open access option.

Abbreviation: GCP, gross cell product.

*E-mail: william.nordhaus@yale.edu.

wealth of nations. A review of a handful of textbooks on economic development shows that climate is confined to a few lines in hundreds of pages. (Exceptions are ref. 3 and more recent work discussed later.) The modern view of economic growth presents development as an engine fueled by capital, labor, and technology; sometimes, mineral resources are included, but only with a major stretch of interpretation would we equate resources with geographic attributes. The recent wave of studies investigating international differences in productivity has generally omitted climate as a determining variable.

Over the last decade, economists have begun to introduce geography into studies of economic growth and development. One early set of studies by Hall and Jones (4) investigated the reasons for the enormous diversity of per capita incomes across nations. Their main hypothesis was that average output differences across nations are primarily determined by institutions and government policies. In examining statistically exogenous instruments, they found that geography (measured as distance from the equator) was among the most significant variables behind differences in per capita output by country. They speculated that location affects economic success because of patterns of human settlements, which influence institutions.

The study of economic geography has been revitalized by the work of Sachs and his colleagues (5, 6). The major thrust of this work is to understand the economic problems of tropical Africa. They examine geographic factors such as the percent of the land area in the tropics as influencing growth. Their surprising conclusion is "Our statistical estimates, admittedly imprecise, actually give approximately two-thirds of the weight of Africa's growth shortfall to the 'noneconomic' conditions, and only one-third to economic policy and institutions" (5). Other studies examine the role of "landlockedness," coastal settlements, and tropical diseases on economic activity (6).

The geographically based studies on Africa have come under heavy criticism for both technical and economic reasons. One set of issues concerns the statistical "endogeneity" of the independent variables. A second and more far-reaching criticism concerns the relative importance of institutions. Several studies argue (along the lines of Hall and Jones discussed above) that high incomes today are best understood as determined by historical conditions in which geography led to settlement patterns that were favorable to good institutions (such as British settlement in North America), and then that good institutions led to high incomes (7). Although these studies are not the last word on the subject, a casual look at East and West Germany, North and South Korea, and Baja and Alta California surely suggests the importance of institutions in economic growth.

Existing studies serve many useful purposes, but they have three distinct shortcomings for determining the impact of geographic attributes on economic activity, all of which are remedied by the present data set. First, virtually all studies focus on national data. If institutions are indeed a key ingredient in economic growth, then it would be very difficult to sort out geographic from national influences without disaggregating below the national level. The gridded data used here overcome this obstacle by employing almost 20,000 terrestrial observations (hence, many per nation) as compared to the 100 or so national observations customarily used in the studies just reviewed.

Second, the analysis here is primarily concerned with the geographic intensity of economic activity rather than the personal intensity of economic activity. In other words, it focuses on the intensity of economic activity per unit area rather than per capita or per hour worked. Although geographic intensity may be less interesting for many policy purposes than the determinants of per capita income, the present approach places the emphasis clearly on geography.

Third, virtually all prior studies have focused on proxies for geographic variables rather than those that are intrinsically important. Distance from the equator and percent area in the tropics, for example, have no intrinsic economic significance. One of the advantages of using gridded data rather than national data is that they allow us to use a much richer set of geographic attributes. Most of the important geographic data (including climate, location, distance from markets or seacoasts, and soils) are collated on a geographic basis rather than based on political boundaries. There is also an important interaction between the finer resolution of the economic data and the use of geographic data because, for many countries, averages of most geographic variables (such as temperature or distance from seacoast) cover such a huge area that they are virtually meaningless, whereas for most grid cells the averages cover a reasonably small area.

Methodology for Estimating Gross Cell Product

The Concept of Gross Cell Product (GCP). The major statistical contribution of the present research program has been the development of "gridded output" data, or GCP. In this work, the "cell" is the surface bounded by 1-degree latitude by 1-degree longitude contours. A full description of the data and methods can be found at the project web site (http://gecon.yale.edu).

The globe contains 64,800 such grid cells; we provide output estimates for 25,572 terrestrial cells. Of these terrestrial cells, 19,136 cells are outside Antarctica, 17,433 have complete and minimum-quality data, and 14,859 have complete, minimum-quality data with nonzero population and output.

The grid cell is selected because it is the unit for which data, particularly on population, are most plentiful. It also is the most convenient for integrating with global environmental data. Additionally, this coordinate system is (to a first approximation) statistically independent of economic data (which obviously is not the case for political boundaries), and the elements are of (almost) uniform size except in polar regions. From a practical point of view, there is no alternative to a grid measurement system such as the one used in the paper.

The conceptual basis of GCP is the same as that of gross domestic product and gross regional product as developed in the national income and product accounts of major countries, except that the geographic unit is the latitude-longitude grid cell. GCP is gross value added in a specific geographic region; gross value added is equal to total production of market goods and services in a region less purchases from other businesses. GCP aggregates across all cells in a country to gross domestic product. We measure output in purchasing-power-corrected 1995 U.S. dollars by using national aggregates estimated by the World Bank. We do not generally adjust for purchasing-power differences within individual countries. The exception to this rule is that we make purchasing-power adjustments for oil and mineral production in countries with a high proportion of output coming from these sources.

The general methodology for calculating GCP is the following:

$$\text{GCP by grid cell} = (\text{population by grid cell}) \times (\text{per capita GCP by grid cell}). \quad [1]$$

The approach in Eq. **1** is particularly attractive because a team of geographers and demographers has recently constructed a detailed set of population estimates by grid cell, the first term on the right-hand side of Eq. **1**.[†] Estimates of GCP, therefore, primarily require new estimates of per capita output by grid cell.

Methodologies for Estimating Per Capita GCP. The detail and accuracy of economic and demographic data vary widely among countries, and we have developed alternative methodologies depending on the data availability and quality. The methodologies are de-

[†]The gridded population data are available online at http://sedac.ciesin.columbia.edu/plue/gpw with full documentation in ref. 8 and updated in ref. 9.

scribed (http://gecon.yale.edu; W.N., Q. Azam, D. Corderi, N. M. Victor, M. Mohammed, and A. Miltner, unpublished data), and data for each country are also available upon request.

In developing the data and methods for the project, two different attributes are central: the level of spatial disaggregation and the source data used to construct the estimates of gross cell product. In terms of spatial disaggregation, there are usually three political subdivisions: (*i*) national data, (*ii*) "state data" from the first political subdivision, and (*iii*) "province data" from the second political subdivision. We use the lowest political subdivision for which data are available, although different levels are sometimes combined.

There are four major sources of the economic data: (*i*) gross regional product (such as gross state product for the United States), (*ii*) regional income by industry (such as labor income by industry and counties or provinces for the United States and Canada), (*iii*) regional employment by industry (such as detailed employment by industry and region for Egypt), and (*iv*) regional urban and rural population or employment along with aggregate sectoral data on agricultural and nonagricultural incomes (used for African countries such as Niger). For each country, we combine one or more of the four data sets at one or more regional levels.

Specific Methodologies. Some examples illustrate the variety of methodologies. (*i*) For the United States, government estimates are available for gross state product for 50 states. We use detailed data on labor income by industry for 3,100 counties to develop per capita gross county product. We then apply spatial rescaling described below to convert the county data to the 1,369 terrestrial grid cells for the United States. We would judge these estimates to be highly reliable. A similar approach was used for Canada, the European Union, and Brazil. (*ii*) For most other high-income countries, we use gross regional product by first political subdivision (such as oblasts for the Russian Federation). For small- or medium-sized countries (Argentina), this approach will be relatively reliable, whereas for large countries (Russia) the regions are sometimes very large and the spatial resolution is consequently poor. (*iii*) For many middle-income countries, such as Egypt, we have data from recent censuses, which collect data on employment by region and industry. We then use these data along with national accounts data on national output by industry to estimate output by region and industry and then aggregate these data across industries to obtain estimates of gross regional product. (*iv*) For Nigeria and many of the lowest-income countries, we have no regional economic data. In these cases, we combine population censuses on rural and urban populations with national employment and output data to estimate output per capita by region. For these countries, because of the sparse economic data and limited regional data, estimates of GCP are less accurate than those for high-income countries.

Spatial Rescaling. The data on output and per capita output are estimated by political boundaries. To create gridded data, we need to transform the data to geographic boundaries. I call this process "spatial rescaling," although it goes by many names in quantitative geography such as "the modifiable areal unit problem," "cross-area aggregation," or "areal interpolation" (10–12). Spatial rescaling arises in a number of different contexts and requires inferring the distribution of the data in one set of spatial aggregates based on the distribution in another set of spatial aggregates, where neither is a subset of the other. The scaling problem arises here because all economic data are published by using political boundaries, and these data need to be converted to geographic boundaries.

Having reviewed alternative approaches and done some simulations with economic data, we settled on the "proportional allocation" rule (details available upon request). The first step is to divide each grid cell into "subgrid cells," each of which belongs uniquely to the smallest available political unit (call them "provinces"). The next step is to collect or estimate per capita output for each province. Third, the proportional allocation rule assumes that per capita output is uniformly distributed in each province and that population is uniformly distributed in each grid cell. Based on these assumptions, we can calculate a tentative estimate of output for each subgrid cell as the product of the subgrid cell area times the population density of the grid cell times the per capita output of the province. We next calculate the GCP as the sum of the outputs of each subgrid cell. The final step is to adjust the GCPs to conform to the totals for the province and the country.

This approach is data-intensive and computationally burdensome because it requires estimating the fraction of each grid cell belonging to each province and estimating the economic data for each of the provinces. Calculations indicate that there are significant gains in accuracy from disaggregating. For the United States, using actual county data, we estimate that disaggregating from the national average to counties decreases the root mean squared error of the cell average by a factor of 5.

Impact of Geography, Climate, and Other Geographic Activities on Economic Activity

This final section presents some results of analyzing the patterns of economic activity by using the new G-Econ data set. This study is not meant to be a comprehensive analysis, which must await integrating the data with a further geographic attributes and time series on spatial economic data. Moreover, at this point, we are primarily examining basic patterns and reduced-form estimates; future work should focus on structural estimates of the major variables.

An Economic Map of Europe. Fig. 1 shows an economic contour map of Europe, with some important mountains and lowlands marked. Unlike familiar contour maps, this one has height proportional to the output density (output per square kilometer) in different regions. The economic Mt. Everests are located along a core region from southern England through northern Italy, whereas the peripheral areas, particularly arctic Europe, are the economic lowlands. Maps for other countries are available upon request.

Fig. 2 shows fractile kernel plots of five major geographic variables. A fractile plot first orders the variable from lowest to highest observation. It then estimates a kernel density function or smoothed nonlinear relationship between the fractile and the logarithm ($\log_{10}$) of output density. For example, output density differs by a factor of 10^5 from fractile trough to peak for mean temperature, whereas the difference for mean precipitation is 10^2 from trough to peak. Clearly, all geographic variables have major systematic impacts on economic activity. The relationships are difficult to capture in a simple fashion because the impacts are highly nonlinear.

The Climate-Output Reversal. The first set of tests examines the relationship between economic activity and a limited set of geographic activities, focusing primarily on climate. Many economic studies have examined the relationship between geography and economic activity. One of the major findings is that output per capita rises with distance from the equator. Those studies have used countries as the unit of observation. Are the results confirmed when the unit of observation is refined to grid cells within countries?

Fig. 3 shows a "box plot" of the relationship between mean temperature in each grid cell and the output per capita in that grid cell. A box plot groups the observations in each bin and then estimates several statistics for those observations; the different statistics are explained in the Fig. 3 legend. For this purpose, bins have a 2°C width (every second bin is shown on the bottom axis). Fig. 3 confirms the strong negative but highly nonlinear relationship between temperature and per capita output.

Although per capita output is a key economic variable, output per unit area is the key variable from a geographic and ecological point of view. Fig. 4 shows a box plot of the relationship between mean

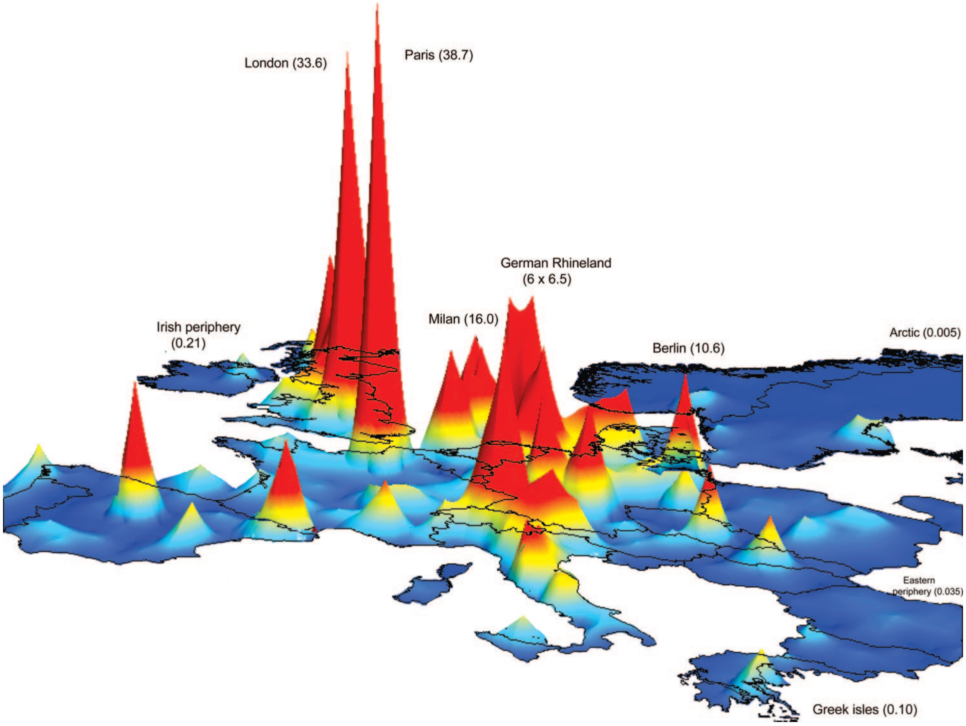

Fig. 1. Economic map of Europe. This figure shows an economic topographical map of Europe with heights proportional to gross domestic product per area. Note how economic activity clusters in the core, whereas the periphery has much lower economic elevations. The observations measure economic activity in millions of 1995 U.S. dollars per km^2 at a 1° latitude by 1° longitude scale.

temperature in each grid cell and the output density in that grid cell. For this purpose, we have assumed that the data are censored and that zero observations have a lower truncation range of 1 dollar per km^2, so $\log_{10}$(truncated observations) = 0. The estimates are relatively unreliable for $\log_{10}$ densities <1.

The striking finding is the very sharp positive gradient between output density and temperature from the lowest observations to ≈5°C; the difference between the peak and the lowest temperature (polar) regions is a factor of at least 10^5. The temperature output varies modestly above 0°C, peaking between 7°C and 14°C. Output density falls by a factor of ≈100 from the peak to the high-temperature regions.

The striking paradox shown in Figs. 3 and 4 can be labeled the climate-output reversal. This reversal indicates opposite relationships between climate and output depending on whether we look at output per person or output per area. (This relationship is similar if the geographic variable is latitude.)

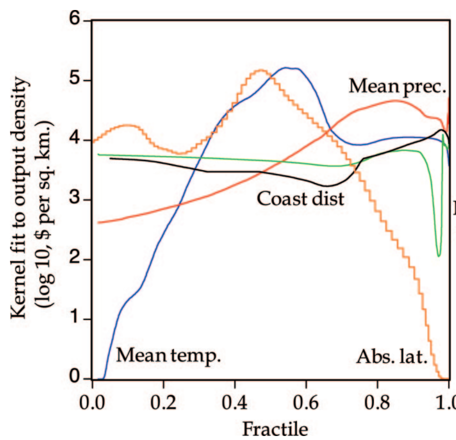

Fig. 2. Fractile plot for key geographic variables. The figure shows the fractile plots for key variables (mean temperature, mean precipitation, mean distance from coast, mean elevation, and absolute value of latitude). Fractiles rank each variable from lowest to highest cell observations. For each variable, we have fitted a kernel density function to the bivariate relationship between the $\log_{10}$ (output density) and the geographic variable. Zero values of output are included as equal 0 (n = 17,796).

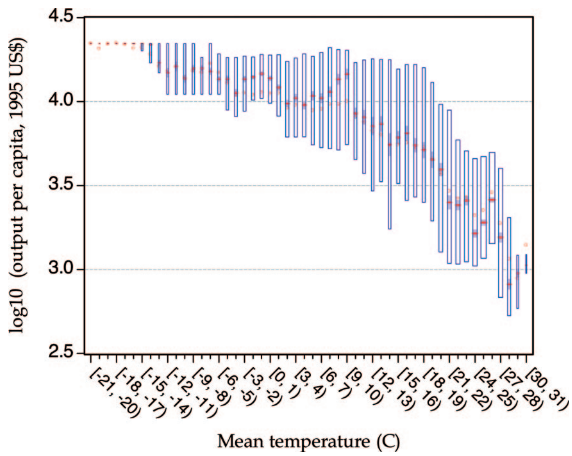

Fig. 3. Boxplot of output per capita and temperature. Earlier studies indicate that high-latitude countries have higher output per capita than those in low latitudes. This relationship is verified by using mean temperature as the geographic variable for grid cells. Coldest regions have an output per capita ≈12 times that of warmest regions. In boxplots, the means are the red circles, the medians are the heavy red horizontal line, the one-sigma ranges of the median are the blue shaded regions, and the interquartile ranges are the boxes. The width of the box is proportional to the square root of the number of observations in each bin. The bin is shown on the horizontal axis, but only every other bin can fit on this graph. The vertical scale is $\log_{10}$, so that each unit is a factor of 10. Zero observations are omitted (n = 15,755).

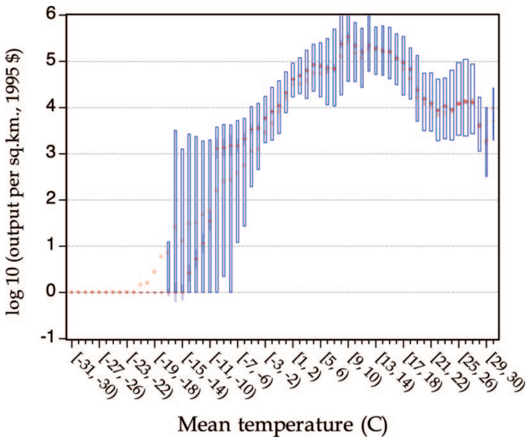

Fig. 4. Boxplot of output density and temperature. This boxplot shows the distribution of output density by temperature. Output density varies by at least five orders of magnitude from cold to temperate region. For the explanation of the boxplots, see Fig. 3. Zero observations are set at $\log_{10}(x) = 0$ ($n = 18{,}995$).

What is the explanation for the climate-output reversal? To a first approximation, the reason is that people are mobile, whereas land is not. Under economic conditions that have existed for recorded history, areal productivity is low in ice-covered and very cold regions. This point is obvious for agriculture, but with few exceptions (such as skiing and glaciology), it is also true for other sectors of the economy. We can use different regions of high-income countries to illustrate. The output density in northern Greenland is \$500 per km^2, the non-oil output density in Alaska averages \$6,000 per km^2, whereas output density in the lower 48 U.S. is \$800,000 per km^2. Unless the global economy becomes devoted substantially to the extraction of ice, it seems likely that the low areal productivity of cold regions will prevail.

The reasons for high per capita productivity of low-temperature regions are not so obvious. To see what can be explained by human behavior, take the case of perfect economic mobility over space. In other words, assume that people migrate until average outputs are equalized in all regions. Under this assumption, the temperature/output-per-capita gradient of Fig. 3 would be horizontal. Although the assumption of perfect mobility does not hold for recent years, particularly across national boundaries, human mobility is surely at the heart of the difference between Figs. 3 and 4.

Three other factors might give the temperature/output-per-capita gradient its negative slope. First, some output in low-temperature regions is highly capital-intensive (such as oil production) and, therefore, tends to have high output per capita. This factor can be ruled out given the relative unimportance of mineral production in high-latitude regions. Second, there may be "compensating differentials," whereby people require higher real wages to live in unpleasant frigid conditions. Here again, although cold regions are unattractive, evidence on compensating differentials cannot explain the large differences. Moreover, tropical regions have their own perils and fail to show similarly large compensating differentials. Third, and most complex, is the poorly understood fact that countries in temperate and colder regions have higher per capita output than most low-latitude and high-temperature regions. As described in the first section, a major debate in the economic-growth literature is whether these differences involve primarily geography or national institutions.

We can use the G-Econ data set to separate out institutional and other national factors from geographical ones. Estimates of a linear equation of the natural logarithm of output per capita on mean temperature can be run with and without country fixed effects for all large and medium countries ($n = 15{,}229$). If national effects involving institutions and political systems are most important, the country effects would capture such impacts. The country effects decrease the sensitivity of log per capita output with respect to mean temperature by approximately one-third, from -0.058 (± 0.00068) to -0.040 (± 0.00092). This result indicates that approximately two-thirds of the gradient shown in Fig. 3 cannot be explained by country-specific factors such as institutional differences, history, major locational advantages, and such national factors. At this point, the remaining sources of the climate-output reversal are still an open question.

Are Output Differences in Output Explained by Pure Geography? One of the central questions in economic geography is how much of the dispersion of output is explained by geographic variables. The G-Econ data provides an ideal laboratory to answer this question. For this purpose, I estimated a multivariate regression with the logarithm of output per km^2 as the dependent variables, with independent variables being temperature, precipitation, and other geographic variables. More precisely, the equation is

$$\ln(y_{ij}) = \beta_{0j}\, Count_j + \sum_{k=1}^{n} \beta_k g^k(Geo_{ijk}) + \varepsilon_{ij}, \qquad [2]$$

where i is the cell, j is the country or region, and k is the geographical variable. The variables are y_{ij} is output per km^2 in 1995 international U.S. prices, $Count_j$ is country effects, and ε_{ij} is the equation residual. Geographic variables, $Geo_{ijk,}$ are mean annual temperature, mean annual precipitation, mean elevation, "roughness" measured as standard deviation of elevation in grid cell, soil categories, and distance from coastline. The g^k represent polynomial functions of geographic variables. The Greek variables β_{0j} are coefficients on regions, whereas the β_k are regression coefficients on geographic variables. It should be noted that we omit all clearly endogenous variables (such as coastal density, proximity to markets, and health status).

This test uses a dense set of exogenous variables to capture all interactions.[‡] The equation explains 91% of the variance of output density for all 17,409 minimum-quality observations. The geographic variables are all highly significant (as is clear for temperature in Figs. 3 and 4).

The equation has some interesting features. It indicates that the "optimal" temperature (which maximizes output density) is $\approx$12°C. Moreover, it suggests that some countries do particularly well or badly given their climates. Countries that are big negative outliers are Australia, Mozambique, Madagascar, and Angola. Those with positive country effects are Denmark, Japan, France, and Italy. The low density of output in Greenland, Canada, Russia, and Alaska are consistent with the economically inclement climates in those regions.

It should be recognized that much of the dispersion of economic activity is unexplained; the standard error of the multivariate regression is 1.97, which is the equivalent of an average error of a factor of $\exp(1.97) = 7.2$. Geographical variables will probably never explain the high densities of economic activity in Madrid, Paris, or Moscow –nor the relatively low levels in temperate South America and South Africa. Geography is important, but much variability remains.

Africa: Geography, Economics, and Destiny. Africa is widely recognized to be the globe's troubled continent. In terms of economic statistics, although gross domestic product per capita in 2004 was

[‡]The precise specification in Eq. **2** contains 72 country effects plus nine polynomial terms in temperature and precipitation, six statistics on extremes and higher moments in temperature and precipitation, the first and second moments of elevation, three variables for distance from coast (<50 km, <100 km, and <200 km), and 27 soil types. The equation has 17,305 degrees of freedom, although that is probably overstated because of spatial correlation. Undertaking further analysis of these data by using the techniques of spatial statistics is an important area of research. All results are described in detail in the background documentation available upon request.

over \$30,000 in the high income countries, 10 countries of tropical Africa had estimated output per person of <\$1,000 in that year. For those living in the peaceful and prosperous north, these abstract numbers can hardly capture the state of living conditions in this region (6, 13).

What are the sources of poverty in tropical Africa? This topic has engaged scholars for at least two centuries, and recent work focuses on a complex interaction of factors: slavery and colonial repression, dependence on primary commodities, poorly designed economic policies, political instability and civil conflict, overpopulation, high levels of ethnolinguistic and religious diversity, and poor health and the recent AIDS epidemic. Throughout the analysis of Africa's development, unfavorable geographic conditions have been emphasized. For example, Bloom and Sachs conclude, "At the root of Africa's poverty lies its extraordinarily disadvantageous geography. . . " (6) In their major statistical analysis of Africa, Bloom and Sachs use as a dependent variable the growth in output per capita, and their geographic variables are percent land area in tropics, coastal population density, and an Africa dummy. Recent work examines structural estimates of the relationship between disease and climate (14).

These studies are extremely useful, but they cannot capture in a realistic fashion the impact of geography for three major reasons. First, in reality, many studies have no interesting measures of geography, and, most important, they omit any climate variables. The major geographic variable in most economic studies is latitude, which is, at best, a proxy for temperature. Second, as discussed above, the unit of observation is generally the country. Because countries clearly have different institutional features (see North Korea vs. South Korea), there are essentially zero degrees of freedom for whatever geographic variables are used. Third, the statistical analysis is plagued by identification problems, with many of the explanatory variables being endogenous and, therefore, in part determined by climate (for example, coastal population density is clearly endogenous).

The G-Econ database can be used to get a more precise estimate of the impact of climate on the economic performance of tropical Africa. For this purpose, the sample is the 22 countries of tropical Africa for which there are economic data in the G-Econ database. We then estimate Eq. 2 and calculate the impact of geography for tropical Africa and six other regions. The other regions are all low-latitude grid cells (latitude < 25) outside Africa, industrial Europe (the industrial regions of Western Europe), Greenland, and three countries, Australia, Russia, and the contiguous United States.

The approach is to estimate the impact of geography by using Eq. **2** above and then apply the coefficients to the geography of different regions. These equations are reduced-form rather than structural econometric estimates. The impact of geography is calculated as the estimated coefficients times the values of the geographical variables for each region. For example, the impact of geography for tropical Africa is estimated to be

$$\sum_{\substack{k=1 \\ i=(j\in Af)}}^{n} \beta_k g^k(Geo_{ijk}), \qquad [3]$$

where $i = (j \in Af)$ indicates that the estimate contains only grid cells for tropical Africa (Af). We then calculate the differential impact of geography between regions p and m, as

$$\Delta_{\mathrm{pm}} = \sum_{\substack{k=1 \\ i=(j\in p)}}^{n} \beta_k g^k(Geo_{ijk}) - \sum_{\substack{k=1 \\ i=(j\in m)}}^{n} \beta_k g^k(Geo_{ijk}). \qquad [4]$$

Table 1 shows the estimates of geographic impacts, Δ_{pm}, as a matrix for the different reference regions. Each entry shows the logarithmic advantage of the region on the top relative to the region on the left-hand side. The analysis shows the impacts on three variables: output density (\$US per km^2), per capita output (\$US per person), and population density (persons per km^2). The sample size is 2,315 for tropical Africa and from 153 observations (industrial Europe) to 3,066 observations (other low latitude) for other regions.

The first block of the table shows the overall impact of geography on output density. This estimate shows that tropical Africa is severely disadvantaged relative to industrial regions, with a geography handicap of -2.25 relative to industrial Europe (equivalent to 89% lower density in Africa). It is also disadvantaged relative to other low-latitude regions. On the other hand, it has advantageous geography relative to Russia and frozen Greenland.

With respect to per capita output, Africa is significantly disadvantaged by geography relative to all regions, although there is only a small geographic handicap relative to other low-latitude regions. Geography lowers ln per capita output by 0.67 relative to industrial countries, while lowering ln per capita output by 0.16 relative to other low-latitude regions. The relative impact of geography on population density is mixed, Africa being disadvantaged relative to mid-latitude regions and advantaged relative to cold regions.

How much of the difference in economic performance is explained by geography? The last column of Table 1 shows the logarithmic difference in actual values between Africa and the region in columns 3–8. The ratio of the coefficient in the first to last columns is an estimate of the fraction of the logarithmic difference explained by geography. Geography explains 20% of the difference in per capita output between tropical Africa and the two industrial regions; and it explains ≈12% of the difference in per capita output between tropical Africa and other low-latitude regions. Hence, geography contributes substantially to Africa's poor economic performance, but other factors appear to contribute more.

Overall, Africa's geography imposes a significant handicap on output density and per capita output relative to industrial regions. On the other hand, tropical Africa's per capita output does not seem to have a major geographic disadvantage relative to other low-latitude regions.

It should be emphasized that this analysis excludes many geographic factors, dynamics, and societal factors. It is clear that tropical geography is currently economically unproductive, but the reasons are beyond the power of the current data to resolve. Nonetheless, the major finding is that tropical geography has a substantial negative impact on output density and output per capita compared to temperate regions.

Impact of Climate Change on Output. In addition to understanding current patterns of global output, the G-Econ data can be used to investigate the implications of environmental changes. One prominent example is the impact of global warming on output, both globally and by region.

Most studies of the economic impacts of global warming have analyzed the impacts on specific sectors (such as agriculture) or on regional ecosystems (15–17). However, impact studies have concentrated on the United States and high-income countries with extrapolations to other regions.

Using the G-Econ database, we can estimate the impact of different warming scenarios on output by using a global database. The assumptions underlying this projection are similar to those using the "Ricardian" technique for estimating economic impacts of climate change in agriculture (18). More specifically, this approach assumes that economies are in long-run equilibrium with respect to climatic and other geographic variables (this relationship is called a "climate-economy equilibrium"). Because climatic variables in recent years have changed slowly relative to the turnover time of most capital stocks and other underlying economic variables, the assumption of climate-economy equilibrium is reasonable except for those areas where the capital or natural stocks change extremely slowly (counterexamples being soils, wetlands, or the location of cities such as New Orleans).

Table 1. Estimates of impact of geography for different regions

Variable country/region	Country or region: Africa	Contiguous U.S.	Industrial Europe	Other low latitude	Russia	Australia	Greenland	Difference between Africa and reference region
Output density								
Africa	0.00							0.00
Contiguous U.S.	−1.35	0.00						−3.60
Industrial Europe	−2.25	−0.90	0.00					−6.44
Other low latitude	−0.66	0.70	1.60	0.00				−1.06
Russia	2.51	3.86	4.76	3.17	0.00			−0.63
Australia	−0.86	0.49	1.39	−0.21	−3.37	0.00		3.25
Greenland	6.43	7.78	8.68	7.08	3.92	6.54	0.00	8.17
Per capital output								
Africa	0.00							0.00
Contiguous U.S.	−0.68	0.00						−3.13
Industrial Europe	−0.66	0.02	0.00					−3.00
Other low latitude	−0.16	0.52	0.50	0.00				−1.41
Russia	−1.77	−1.09	−1.11	−1.61	0.00			−2.53
Australia	−0.26	0.42	0.40	−0.10	1.51	0.00		−3.02
Greenland	−1.52	−0.84	−0.86	−1.36	0.25	−0.50	0.00	−3.05
Population density								
Africa	0.00							0.00
Contiguous U.S.	−0.72	0.00						−0.39
Industrial Europe	−1.56	−0.84	0.00					−3.35
Other low latitude	−0.13	0.59	1.43	0.00				0.12
Russia	2.00	2.72	3.56	2.13	0.00			1.38
Australia	−0.61	0.11	0.95	−0.48	−2.61	0.00		3.50
Greenland	3.17	3.89	4.73	3.30	1.17	2.45	0.00	4.53

This table shows the estimated effect of geography on relative output or population. Figures in columns 2–8 are Δ_{pm}, or the difference in the logarithm of output or population densities between regions. A positive sign indicates that geography is relatively advantageous for the region shown on the top of each row. For example, geography is estimated to lower ln output density in tropical Africa by 2.25 relative to industrial Europe but to raise ln output density 2.51 relative to Russia. Omitted entries in the upper right are the symmetrical entry with sign changed. A difference of 0.69 is a factor of 2. The most disadvantaged region shown in Greenland. The last column shows the logarithm of the ratio of actual African value to value for region in the left-hand column. For example, the difference in the ln of per capita output between tropical Africa and Australia is −3.02, so Africa's level is exp(−3.02) = 0.049 of Australian per capita output.

To estimate the impact of climate change, I compare the economic productivity of the existing climate with that of two climate-change scenarios that reflect an equilibrium impact of doubling of CO_2-equivalent atmospheric concentrations.

CC1. The first scenario is one in which only temperature is assumed to change. We take a standard scenario that corresponds to a doubling of atmospheric concentrations of CO_2-equilvalent greenhouse gases. This scenario assumes a mean surface temperature change of 3.0°C over all terrestrial grid cells in the sample, and the temperature change is latitude-dependent to capture estimates from general-circulation models. The first scenario assumes no change in precipitation.

CC2. The second scenario is one in which there is mid-continental drying as well as the temperature change assumed in CC1. To model the mid-continental drying, it is assumed that precipitation declines by 15% in areas at least 500 km from the coast in mid-latitude regions (between latitudes 20 and 50 north or south), whereas precipitation rises 7% in other areas (19).

The scenarios are drawn from the multimodel assessments in the Intergovernmental Panel on Climate Change Third Assessment Report (19), Chapter 9, figures 9.10 and 9.11. They have been rescaled to correspond to a 3°C global average equilibrium increase. CC1 has been widely used in the impacts literature. Although oversimplified, it captures the results of general-circulation models reasonably well. The assumptions underlying the second scenario are less well established because the extent and location of the mid-continental drying differ significantly across models. One task on the future research agenda in this field will be to couple directly the gridded output data and other economic relationships to climate models.

The projection of the impact of climate change begins with Eq. **2** described above with one further modification. I have modified the specification in Eq. **2** by using a more parsimonious list of variables and adding variables that are country-specific linear temperature effects. The purpose of these modifications is to reduce the possibility of spurious correlations and to ensure that low-quality country data do not contaminate the estimates.[§]

To estimate the impact of the two scenarios involves the following steps: (*i*) First, estimate a regression of cell output by using the historical climate and other variables. (*ii*) Next, change temperature and precipitation by grid cell according to scenarios CC1 or CC2. (*iii*) Then, estimate the change in output as the difference between the projections for scenarios in (*i*) and (*ii*). (*iv*) Next, aggregate the changes by using as weights cell area, output, and population. (*v*) Because the equations and transformations are highly nonlinear, estimate the statistical variability of the estimates and projections by using "bootstrap" techniques with 100 replications.

The basic results are shown in Table 2, where we combine the two scenarios and the bootstraps with different aggregation approaches. The population weights measure the change in average incomes, the output weights estimate the impact on global output, and the area weights ask what happens to the average terrestrial location.

[§]The equation used for the global warming equation is the log of output density as a dependent variable and, as independent variables, mean and squared temperature, mean and squared precipitation, elevation, roughness, roughness squared, the three distance-from-coast variables, country effects, and linear temperature effects by country.

Table 2. Estimated impact of global warming on world output

Variables	Impact on global output: Estimated impact, %	Bootstrap estimated standard error, %
Scenario CC1		
Output weights	−0.93	0.13
Population weights	−1.73	0.24
Area weights	−0.72	0.10
Scenario CC2		
Output weights	−1.05	0.13
Population weights	−2.95	0.25
Area weights	−1.41	0.10

Estimate is for impact of ln output density as determined by geographic variables. Scenario CC1 is warming only, where as scenario CC2 includes mid-continental drying, as explained in text. Different weights take average output change by grid cell weighted by the fraction of global output, population, or area in grid cell. Estimates omit cells with zero output. Bootstrap standard error is for 100 samples.

The basic message is that the CC1 scenario (warming with no precipitation change) shows a negative impact on output by any of the three weighting systems. The projected output change is −0.9% by using cell output weights and −1.7% by using cell population weights. The one-sigma ranges around the estimates indicate that the estimates are very tightly determined.

The CC2 scenario (warming with mid-continental drying) shows more adverse effects than the CC1 scenario. The differences between the two scenarios are progressively greater as the weights move from output to area to population. The intuition here is that the largest impacts occur where population density is highest. Perhaps the most relevant result is the population-weighted CC2 scenario, which indicates an average impact of −3.0% of average output from the doubling scenario.

These results are among the first comprehensive estimates of the global economic impact of greenhouse warming. These global estimates also have a statistical basis and, therefore, can determine the associated statistical errors. The estimated impacts are larger than most existing estimates of market damages. Nordhaus and Boyer estimated impacts of a 2.5°C warming to be −0.2% and −0.4% of global output for output and population weights, respectively (17). Tol's benchmark estimate for a 1°C warming is +2.3% for output weighted-impacts (20). R. Mendelsohn, A. Dinar, and L. Williams (unpublished study) use an approach similar to the present one, focusing on countries, and find a neutral effect of climate change (±0.1% globally by using output weights) in 2100, although low-latitude countries are expected to experience serious negative economic impacts.

At the same time, three reservations should be emphasized. First, the estimates of the impact of geographic variables on output leave a significant fraction of output unexplained. Until output variations are fully and robustly explained, we cannot be confident about a projection based on an incomplete model. Second, these estimates include only market output and do not incorporate any nonmarket impacts or abrupt climate changes. Hence, impacts on ecosystems or amenities, and particularly the potential for abrupt climate change, need to be included in a full impacts analysis (21, 22). Finally, the model underlying the estimates here, particularly the assumption of climate-economy equilibrium, is highly simplified. The dynamic nature of economic growth cannot be adequately captured in cross-sectional estimates. Given the sluggish reactions of population distributions to changing conditions, existing settlement and economic patterns may still be adjusting to economic and climatic conditions. Pursuing each of these issues requires further data and methodological developments.

Next Steps. This article describes but the initial excursion into the use of geographically scaled economic data to understand the location of economic activity on a global scale. Much further work remains. For the database, it is important not only to extend the data spatially, but even more important, to create time series, and eventually to disaggregate the data for the major sectors (especially agriculture, mining, and manufacturing). For the analytical work, it will be necessary to incorporate other geographic variables and variables reflecting historical, technological, and institutional factors. Structural estimates of the pathways from climate to output and living standards are an important next step.

Notwithstanding the limitations of the analysis and data, four points stand out. First, it is clearly possible to measure global economic activity on a finer scale than has been done up to now; approaches such as the G-Econ data allow more uniform measurement, produce greater spatial resolution by a factor of ≈100, and allow better linkage of economic data to geographic data. Second, the data reveal a pattern in which the density of economic activity is very strongly related to geographic conditions, especially temperature, precipitation, and coastal proximity. Third, applying the data to the tropical Africa, we estimate that Africa's geography is indeed a major economic disadvantage relative to temperate countries, but Africa's geography is only marginally disadvantageous relative to other low-latitude regions. Finally, using the G-Econ data to estimate the impact of global warming, we estimate that an equilibrium doubling of CO_2-equivalent greenhouse gas concentrations will have significantly more negative impacts than was found in earlier studies.

I thank the many coworkers who helped develop the G-Econ database and Robert Mendelsohn, Jeffrey Sachs, V. Kerry Smith, Richard Tol, Waldo Tobler, B. L. Turner II, and Christopher Udry. This research was funded by the National Science Foundation and the Glaser Progress Foundation.

1. Landes, D. (1998) *The Wealth and Poverty of Nations: Why Some Are So Rich and Some So Poor?* (Norton, New York).
2. Diamond, J. (2005) *Collapse: How Societies Choose to Fail or Succeed* (Viking, New York).
3. Kamarck, A. N. (1976) *The Tropics and Economic Development: A Provocative Inquiry into the Poverty of Nations* (Johns Hopkins Univ. Press, Baltimore, MD).
4. Hall, R. E. & Jones, C. I. (1999) *Q. J. Econ.* **114**, 83–116.
5. Bloom, D. E. & Sachs, J. D. (1998) *Brookings Pap. Econ. Act.* **2**, 207–295.
6. Sachs, J. D., Mellinger, A. & Gallup, J. (2000) in *Oxford Handbook of Economic Geography*, eds. Gordon L., Clark, G. L., Feldman, M. P. & Gertler, M. S., (Oxford Univ. Press, Oxford).
7. Acemoglu, D., Johnson, S. & Robinson, J. A. (2001) *Am. Econ. Rev.* **91**, 1369–1401.
8. Tobler, W., Deichmann, U., Gottsegen, J. & Malloy, K. (1995) *The Global Demography Project Technical Report 1995-6* (Natl. Center Geogr. Info. Anal., Santa Barbara, CA).
9. Deichmann, U., Balk, D. & Yetman, G. (2001) *Transforming Population Data for Interdisciplinary Usages: From Census to Grid* (Cent. Intl. Earth Sci. Info. Netw., Palisades, NY).
10. Tobler, W.A. (1979) *J. Am. Stat. Assoc.* **74**, 519–530.
11. Flowerdew, R. & Green, M. (1989) in *The Accuracy of Spatial Data bases*, eds. Goodchild, M. & Gopal, S. (Taylor and Francis, London) pp. 239–247.
12. Gotway, C. A. & Young, L. J. (2002) *J. Am. Stat. Assoc.* **97**, 632–648.
13. Sachs, J. D., McArthur, J. W., Schmidt-Traub, G., Kruk, M., Bahadur, C., Faye, M. & McCord, G. (2004) *Brookings Pap. Econ. Act.* **1**, 117–240.
14. Kiszewski, A., Mellinger, A., Spielman, A., Malaney, P., Sachs, S. E. & Sachs, J. D. (2004) *Am. J. Trop. Med. Hyg.* **70**, 486–498.
15. Fankhauser, S. (1995) *Valuing Climate Change: The Economics of the Greenhouse Effect* (Earthscan, London).
16. Mendelsohn, R. & Neumann, J. E., eds. (1999) *The Impact of Climate Change on the United States Economy* (Cambridge Univ. Press, Cambridge, U.K.).
17. Nordhaus, W. D. & Boyer, J. (2000) *Warming the World: Economic Models of Global Warming* (MIT Press, Cambridge, MA).
18. Mendelsohn, R., Nordhaus, W. D., Shaw, D. (1994) *Am. Econ. Rev.* **84**, 753–771.
19. Houghton, J. T., Ding, Y., Griggs, D. J., Noguer, M., van der Linden, P. J. & Xiaosu, D., eds. (2001) *Climate Change 2001: The Scientific Basis: Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge Univ. Press, Cambridge, U.K.).
20. Tol, R. (2002) *Environ. Res. Econ.* **21**, 47–73.
21. National Research Council (2002) *Abrupt Climate Change: Inevitable Surprises* (National Academy Press, Washington, D.C.).
22. Alley, R. B., Marotzke, J., Nordhaus, W. D., Overpeck, J. T., Peteet, D. M., Pielke R. A., Jr., Pierrehumbert, R. T., Rhines, P. B., Stocker, T. F., Talley, L. D. & Wallace, J. M. (2003) *Science* **299**, 2005–2010.