When you can measure what you are speaking about, and express it in numbers, you know something about it; when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the stage of science.

Lord Kelvin

In contrast to early economics, much of modern economic analysis today largely ignores geography. Human populations cluster mainly on coasts and rarely on ice-sheets. Yet, modern growth economics generally ignores geographic factors such as climate, proximity to coasts, soils, tropical pests, and permafrost.[1]

[1] The authors are Sterling Professor of Economics at Yale University and Associate in Research at Yale University. Authors can be contacted at william.nordhaus@yale.edu and xi.chen@yale.edu. Version is *graphics_120408.docx*, as of *12/4/2008.*

There are many reasons why geography plays but a limited role in modern economics. One important reason is that most social and economic data are collected on the basis of political boundaries – cities, counties, states, and nations. Such sources yield very rich data for socioeconomic variables, but cannot be easily integrated with most geophysical measures, which are collected on a geophysical basis. The data set examined here is organized around geophysical boundaries. We have developed observations on "gross cell product," which measures output with a resolution of 1° latitude by 1° longitude.

The original data set, published in 2006, was available for a single year (1990). We have updated the observations, corrected several data and methodological issues, and extended it to three periods (1990, 1995, and 2000). The current data set covers approximately 27,500 terrestrial grid cells, with 18,000 minimum-quality observations. The data are available at gecon.yale.edu.

The change in emphasis proposed here has a major effect on our ability to examine the geographic attributes of economic activity. The G-Econ database 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. In the spirit of Lord Kelvin, our new measures of spatial economic activity may give impetus to the reemerging geographic economics.

I. Output on a Gridded Basis

The major statistical contribution of the G-Econ project has been the development of data on "gridded output," *gross cell product,* or GCP. In this work, the "cell" is the surface bounded by 1-degree latitude by 1-degree longitude contours. The globe contains 64,800 such grid cells; we have partial data on 27,442 observations, of which virtually all have reasonably complete data on climate, population, and output. The grid cell is the selected geographic unit because it is the geophysical system for which data are most plentiful, particularly population data. It also is the most convenient for integrating with global environmental data. Additionally, it has the features that the coordinate system is (to a first approximation) statistically independent of economic data (which obviously is not the case for political boundaries) and that the elements are (except at high latitudes) of nearly uniform size. From a practical point of view, there is no alternative to the geophysical measurement system used in the paper.

The conceptual basis of GCP is the same as that of gross domestic product (GDP) 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. Gross cell product 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 businesses. GCP aggregates across all cells in a country to gross domestic product. We measure output in purchasing-power-corrected 2000 U.S. dollars using national aggregates estimated by the World Bank.

We begin by calculating GCP as follows:

(1) GCP by grid cell = (population by grid cell) x (per capita GCP by grid cell)

The approach in (1) is particularly attractive because a team of geographers and demographers has constructed a detailed set of population estimates by grid cell, the first term on the right-hand side of

(1).[2] Estimates of gross cell product therefore primarily require new estimates of per capita output by grid cell.

II. Global Economic Graphics

This paper is primarily devoted to providing several graphical images of the GEcon data set. Using a graphical information system (here ArcGIS), we can display an "economic globe," similar to the ones that are found in the family living room. This is an economic elevation map, where the heights represent the level or "height" of output in a particular location. The maps we show here are transformed so that the height is proportional to the output in each grid cell. The graphics in this paper and on our website are created from the data updated in December, 2008.

Figures 1 through 3 are two-dimensional snapshots of the economic globe from different vantage points. To access the rotating globe, with some patience, you can go online at http://gecon.yale.edu/bigglobe.avi. Figure 4 shows a contour map of China.

---

[2] Tobler, W., Deichman, U., Gottsegen, J. & Malloy, K. (1995) *The Global Demography Project Technical Report 1995-6* (National Center for Geographic Information and Analysis, Santa Barbara, CA), available at http://sedac.ciesin.columbia.edu/gpw/.

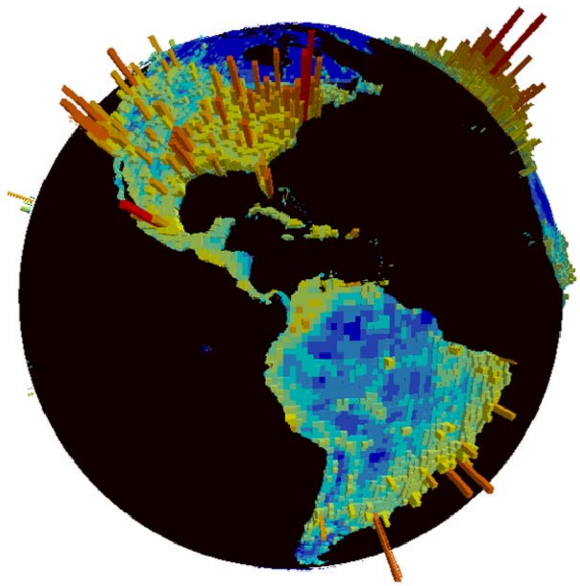

Figure 1. Snapshot of economic map of Western Hemisphere

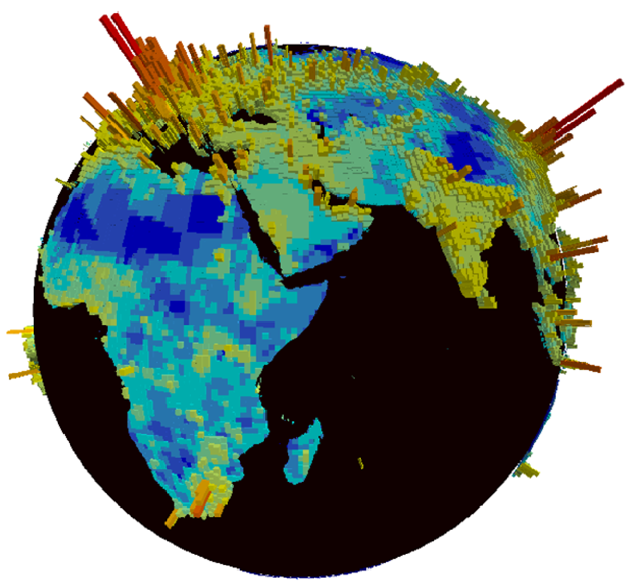

Figure 2. Snapshot of globe over Africa and Europe

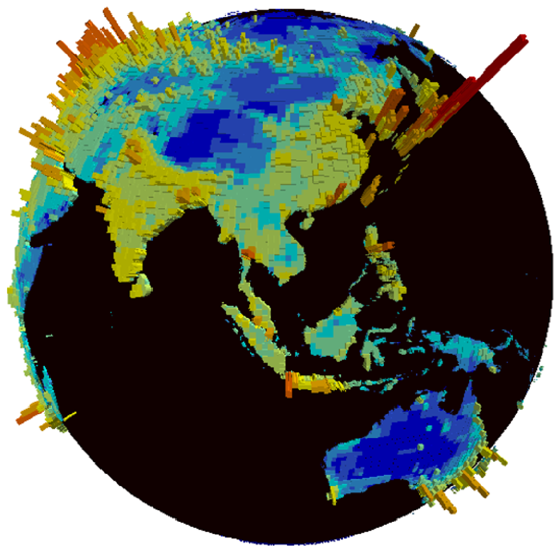

Figure 3. Snapshot of economic globe over Asia

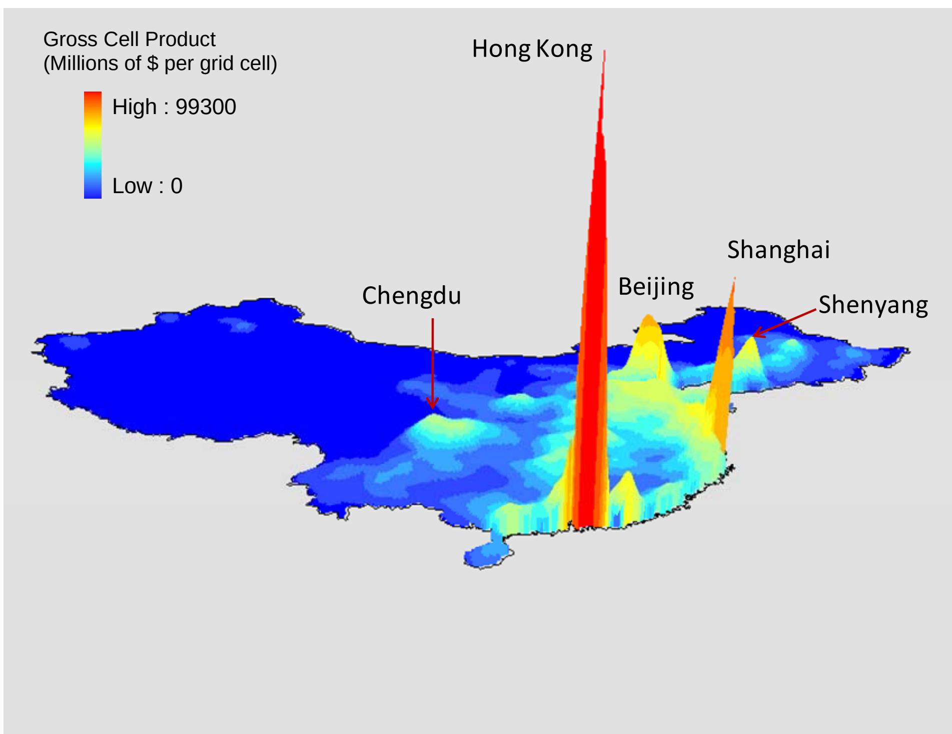

Figure 4. Economic elevation map of China

One feature that is immediately apparent from the globe is how much of economic activity is clustered on coastlines. This is particularly apparent in China, but can be seen in many parts of the world. Two of the areas where economic output is not strongly clustered on coasts are the United States and Europe. The reasons for the differences in the coastal clustering of countries are largely unexplored in economic geography.

III. The Geographical Determinants of Economic Activity

There are many applications of geospatial data. We investigate two that have interesting graphical features: the distribution of economic "deserts" and the economic geography of Africa.

A. Where are the economic "deserts"?

An interesting question to ask is, where are the economic deserts of the world? In other words, what are the most *unproductive* parts of the globe? We address this question by looking at grid cells where the economic activity (as measured by gross cell product) is zero. In reality, there are probably few cells with literally zero output, and zero output is the result of truncation. But those cells with a zero estimated output are very likely to have very low output.

There are 6721 cells which we estimate to have zero economic output, and 8335 cells with near-zero output density. We define a “near-zero” output density as one with output density less than $100 per square km; there are 18,884 cells with significant output density (larger than near-zero), for which the average output density is $404,000 per square km.

Figure 5 shows the basic result of our study of economic deserts. Simply put, the economic deserts of the world are the cold regions. Once the mean temperature falls below about 20 degrees C, the probability of being an economic desert exceeds 80 percent. Virtually all these cells are in Antarctica, Greenland, northern Russia, and northern Canada. Although the first impulse of many people is to think of sand deserts as unproductive regions, in fact it is the cold regions (essentially those regions which are covered in ice) where little or no economic activity takes place.

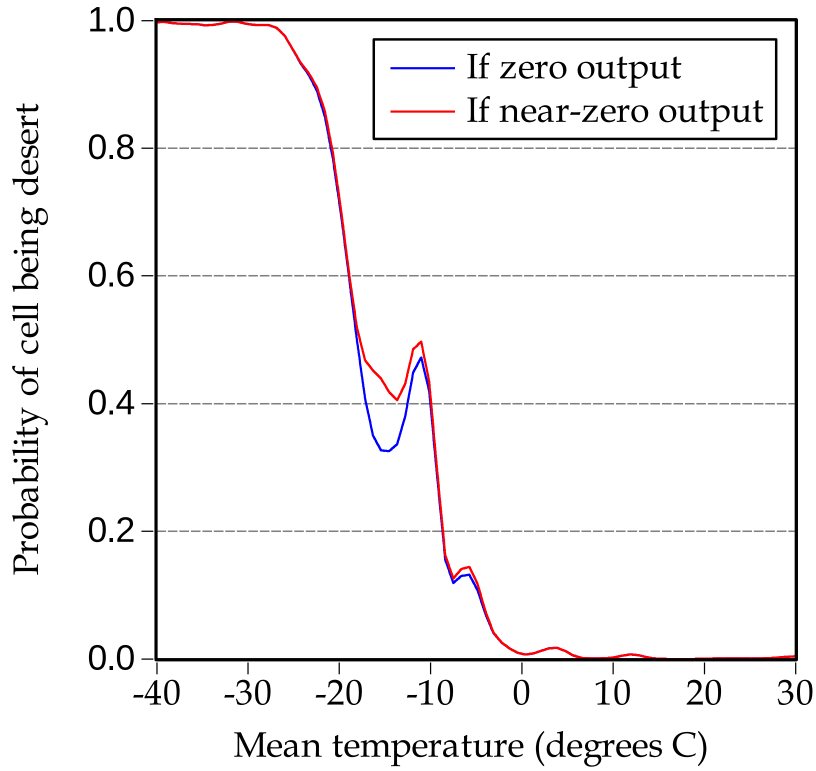

Figure 5. Likelihood of being economic desert as function of temperature

There are very few deserts at the other end of the distribution, with very high temperature and low precipitation. There are only 137 near-zero output hot cells, spread widely across the low latitudes.

B. African geography and economic activity

A second particularly important question is the role of geography in tropical Africa. Africa is widely recognized to be the globe's troubled continent. In terms of economic statistics, while GDP per capita in 2004 was over $30,000 in the high income countries, 10 countries of tropical Africa had estimated output per head less than $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.[3]

What are the sources of the 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 ethno-linguistic 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, Sachs,

[3] See the detailed review in Bloom, D.E., Sachs, J.D. (1998) "Geography, Demography, and Economic Growth in Africa" *Brookings Papers on Economic Activity* **2**, 207-295.

and Collier conclude, “At the root of Africa’s poverty lies its extraordinarily disadvantageous geography…”[4] In their major statistical analysis of Africa, Sachs, Bloom, and Collier use as a dependent variable the growth in output per capita, while their geographic variables are percent land area in tropics, coastal population density, and an Africa dummy. This emphasis on geography has been criticized by scholars who emphasize the importance of institutions, political instability, and policies.

Most studies of African economics cannot capture in a realistic fashion the impact of geography for three major reasons. First, in reality these studies have no interesting measures of geography, and, most important, they omit any climate variables. The major geographic variable in all economic studies is latitude, which is at best a proxy for temperature. Second, as discussed above, the unit of observation is the country. Because countries clearly have different institutional features (for example, North Korea v. 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

---

[4] *Id., p. 211.*

by climate (for example, coastal population density is clearly endogenous).

We estimate the impact of geography on Africa by looking at whether Africa has a lower level of output than other low-latitude regions. For this question, we estimate the following regression:

$$(2) \quad y_{ijk} = \alpha_0 + \sum_{k=1}^{n} \beta_k \, g^k(Geo_{ijk}) + \gamma IFAF_{ijk} + \varepsilon_{ijk}$$

For this regression, y_{ijk} is a measure of economic activity, where *i* stands for latitude, *j* for longitude, and *k* for nation, the variables labeled Geo_{ijk} are the geophysical variables including latitude, $IFAF_{ijk}$ is a dummy variable that equals 1 if a cell is in sub-Saharan Africa and 0 if not, ε_{ijk} is a random error, and γ is the key coefficient which estimates whether being a country in sub-Saharan Africa affects the output variable. For this purpose, we define sub-Saharan Africa as all grid cells of the continent south of latitude 20N. We capture the effect of low-latitude regions by including as a geophysical variable latitude and latitude squared. Note that this study does not focus on per capita output but on economic density as measured by output per square kilometer. The estimates are for the year 1990. We have not corrected for spatial correlation, but that will be done in future work.

To begin with, we note that there are very few economic deserts at low latitudes. Of the 4465 grid cells between 20N and 20S, there are no cells with zero output and only 45 with near-zero output. Not surprisingly, the low-output high-temperature cells really are deserts.

We use for our estimates the 11,803 cells with positive cell output, with land area greater than 1 percent of a grid cell, and with mean temperature above zero. We estimate equation (2) with a dependent variable being the log of GCP per square kilometer. Independent geophysical variables are linear and squared terms in mean precipitation, mean temperature, elevation, and the distance from coastline, lakes, and rivers. The geophysical variables without latitude explain about 45 percent of the cross-sectional variance. If we add latitude and latitude squared as variables, the geophysical variables explain 55 percent of the variance.

As is well known from earlier studies, it is clear that output density is definitely lower at low latitudes. The predicted output density at latitude

40 (New York) is 2.05 logarithmic points (or about eight times) higher than at the equator.

The question is how much additional output penalty Africa suffers relative to other regions or to other low-latitude regions. Begin by estimating the penalty to output density of being a region in Sub-Saharan Africa relative to other countries. For these regressions, we omit cells with zero output and those in cold regions (mean temperature less than 0 °C). Table 1 shows the results of the tests. The Appendix provides the full regression for the last row of Table 1. The first row in Table 1 shows that grid cells in Sub-Saharan Africa have a penalty for output density of -1.78 ($\pm$ -0.06) relative to all non-African grid cells.

Region and specification	Penalty for economic density	
	Coefficient	Standard error
African penalty v. all non-African countries		
No geography	-1.78	0.06
Geography	-0.77	0.05
Geography and latitude	-0.27	0.05
African penalty v. other low-latitude regions		
No geography	-1.95	0.06
Geography	-1.01	0.06
Geography and latitude	-0.24	0.06

Note: The penalty is the difference in the logarithm of output density between African cells and other cells. A logaritmic difference of -0.24 is a fractional difference of exp(-.24)-1 = 0.787 - 1 = -0.213.

Table 1. Effect of geography on African economic density

The second row adds the geophysical variables to the equation. This shows that after correction for the several geophysical variables, Sub-Saharan Africa has an output penalty of only -0.77 ($\pm$ -0.05). Adding latitude and latitude squared to the equation reduces the logarithmic difference to -0.27 ($\pm$ 0.05).

The second set of results compares Sub-Saharan Africa to other low-latitude cells. The output penalty without geography is slightly larger than for all cells. The impact of correcting for geophysical variables explains about one-half of the difference between Sub-Saharan Africa and other low-latitude regions. Again, adding latitude explains a substantial fraction of the remaining difference.

These results indicate that approximately half of the output penalty experienced by Sub-Saharan Africa relative to other regions, as well as other low-latitude regions, can be explained by geophysical variables such as temperature, precipitation, distance from coastlines or major bodies of water, and elevation.

Latitude as a geographical variable

An interesting question involves the interpretation of "latitude" in these equations. Distance from the equator has been widely used in economic studies of geography, and these equations confirm that this variable does indeed explain a substantial fraction of output differences. There is a vigorous debate about the exact significance of latitude.

The interesting point here is that a substantial part of the "latitude effect" does not appear to reflect geophysical variables such as climate, distance from coastline, and similar variables. These results confirm earlier studies that have argued that variables other than pure geography – but which are correlated with latitude – are responsible for a substantial part of the poor economic performance of low-latitude regions. Put differently, geography and latitude explain about 90 percent of the difference in economic performance between Sub-Saharan Africa and other low-latitude regions. However, only about half of the difference is pure geophysical variables alone. The clear implication is that the about one-half the "latitude effect" reflects economic, institutional, and other non-geophysical variables.

The summary of the results here is that we do indeed find that low-latitude regions have lower economic density than mid-latitude regions. The syndrome is found outside of Africa as well, however. We find that, using the high-resolution geophysical data from the G-Econ data set, that about one-half of Africa's economic penalty relative to other countries is associated with its unfavorable physical geography.

Appendix. Regression of Sub-Saharan African grid cells

Dependent Variable: LYDENS
Method: Least Squares
Sample: 1 27395 IF RIG>0.2 AND TEMPAV_F>0
Included observations: 11804

Variable	Coefficient	Std. Error	t-Statistic	Prob.
IFAF-IFLOLAT	-0.220615	0.060611	-3.639853	0.0003
IFLOLAT	-0.116926	0.116464	-1.003966	0.3154
C	-12.67370	0.242240	-52.31880	0.0000
TEMPAV_F	0.316235	0.072209	4.379434	0.0000
TEMPAV_F^2	-0.014932	0.014992	-0.996028	0.3193
TEMPAV_F^3	0.001577	0.001272	1.240067	0.2150
TEMPAV_F^4	-0.000114	4.69E-05	-2.426735	0.0153
TEMPAV_F^5	2.24E-06	6.26E-07	3.587873	0.0003
PRECAV_F	0.016422	0.001083	15.16112	0.0000
PRECAV_F^2	-8.76E-05	2.50E-06	-34.99121	0.0000
D1	-0.001939	0.000175	-11.07077	0.0000
D2	-2.31E-05	0.000209	-0.110962	0.9116
D3	-0.001470	0.000296	-4.964487	0.0000
D4	-0.001779	0.000373	-4.762554	0.0000
ELEV_F	0.000126	7.74E-05	1.629923	0.1031
ELEV_F^2	-4.62E-09	2.08E-08	-0.221860	0.8244
TEMPAV_F*PRECAV_F	0.000531	4.60E-05	11.54448	0.0000
D1^2	5.83E-07	9.02E-08	6.464758	0.0000
D2^2	7.84E-09	1.09E-07	0.072257	0.9424
D3^2	-2.73E-07	2.61E-07	-1.047546	0.2949
D4^2	2.04E-06	5.01E-07	4.064058	0.0000
LAT	0.040283	0.000805	50.05544	0.0000
LAT^2	0.000275	6.14E-05	4.475722	0.0000

R-squared	0.544173	Mean dependent var	-10.51985
Adjusted R-squared	0.543322	S.D. dependent var	2.489206
S.E. of regression	1.682155	Akaike info criterion	3.879975
Sum squared resid	33336.06	Schwarz criterion	3.894348
Log likelihood	-22876.61	Hannan-Quinn criter.	3.884801
F-statistic	639.2872	Durbin-Watson stat	1.568008
Prob(F-statistic)	0.000000		

Variables

LYDENS	log(output per square km)
IFAF	If a cell is in Sub-Saharan Africa
IFLOLAT	If a cell is in a non-African low-latitude cell
C	Constant
TEMPAV_F	Mean cell temperature (C)
PRECAV_F	Mean cell precipitation (mm per month)
D1	Distance from ocean
D2	Distance from major river
D3	Distance from minor river
D4	Distance from lake
ELEV_F	Elevation (meters)
LAT	Latitude (degrees)