NBER WORKING PAPER SERIES

THE VALUE OF LUMINOSITY DATA AS A PROXY FOR ECONOMIC STATISTICS

Xi Chen
William D. Nordhaus

Working Paper 16317
http://www.nber.org/papers/w16317

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2010

The Value of Luminosity Data as a Proxy for Economic Statistics
Xi Chen and William D. Nordhaus
NBER Working Paper No. 16317
August 2010
JEL No. E01,O47,Q4

ABSTRACT

One of the pervasive issues in social and environmental research has been to improve the quality of socioeconomic data in developing countries. Because of the shortcoming of standard data sources, the present study examines luminosity (measures of nighttime lights) as a proxy for standard measures of output. The paper compares output and luminosity at the country levels and at the 1° x 1° grid-cell levels for the period 1992-2008. The results are that luminosity has very little value added for countries with high-quality statistical systems. However, it may be useful for countries with the lowest statistical grades, particularly for war-torn countries with no recent population or economic censuses. The results also indicate that luminosity has more value added for economic density estimates than for time-series growth rates.

Xi Chen
Yale University
28 Hillhouse Ave
New Haven CT 06511
xi.chen@yale.edu

William D. Nordhaus
Yale University, Department of Economics
28 Hillhouse Avenue
Box 208264
New Haven, CT 06520-8264
and NBER
william.nordhaus@yale.edu

Introduction

One of the pervasive issues in social and environmental research has been to improve the quality of socioeconomic data in developing countries. Many countries of tropical Africa, particularly those in war-torn countries, have no reliable censuses of population and only rudimentary economic statistics. The poor quality of the data has hindered attempts to understand economic growth, poverty, health status, and environmental quality. The lack of regional data has been even more daunting for those working at the sub-national level.

This difficulty has been a major difficulty for the GEcon project at Yale, which has undertaken the research project of constructing geophysically based economic data on gross cell output by 1° latitude by 1° longitude for all terrestrial grid cells for the world. The project has been hampered by very poor quality data in many low-income countries with poorly funded statistical systems and particularly in those regions plagued by civil wars. One example is Somalia. While international agencies routinely produce population and output estimates, the basis for these is very tenuous. The last population census in Somalia was 1975. There is no functional statistical office, and indeed there have been only sporadic central governments for two decades. A search for statistics from a web site stating it is the government web site finds a message, "UNDER RECONSTRUCTION PLEASE CHECK BACK SOON." Similar issues arise for many other African countries. The last functional population census for the Democratic Republic of Congo was 1984. Economic data are generally even sparser with the exception data on international trade.

Because of the shortcoming of standard data sources, the authors undertook to examine measures that could supplement or substitute for our measures of regional output that come from standard sources on income, output, and other demographic information. The obvious place to look was nighttime luminosity. There is now a large literature on both luminosity and the use of luminosity as a proxy for population, output, and poverty (see the references in the next section). The intuitive notion is that luminosity might serve as a useful proxy for output because it is "objectively" measured, is highly correlated with output, and is universally available for the world except for the high latitudes. Figure 1 shows a striking image of nighttime lights for India, and it is immediately obvious that high-income regions have higher luminosity. The question is, can these data be usefully exploited for measuring output?

The tentative bottom line of this project is that luminosity shows considerable promise for regions with very poor data but little or no value added for countries with good to excellent statistical systems. The noise in the light signal is

too high to be useful for improving economic accounts for countries like the United States or Brazil, but appears to have sufficient signal for the poor countries of tropical Africa as well as countries with minimal governing structures.

Use of Luminosity as Socioeconomic Indicators

Nighttime luminosity has been gathered from satellites for almost two decades and has been carefully filtered into a series of data at very disaggregated level, with observations beginning in 1992. The results are visually striking (see Figure 1) and are seemingly related to the level of economic activity in different regions.

Luminosity has been increasingly used in studies of areas with weak data systems. Previous studies have used nighttime image data as a proxy for socioeconomic development of particular geographic areas (Doll, Muller, Elvidge 2000; Ebener, Tandom, and Elvidge 2005; Elvidge et al. 2007; Elvidge et al. 1997; Sutton, Elvidge, and Ghosh 2007; Sutton and Costanza 2002; Henderson, Storeygard, and Weil 2009). Elvidge et al. (2007, p.51) conclude "Nighttime lights provide a useful proxy for development and have great potential for recording humanity's presence on the earth's surface and for measuring important variables such as annual growth for development."

In the past decade, researchers have tried a series of tests to support this conclusion. During the initial phase of their research, Elvidge and his colleague focused on correlation between luminosity and GDP values at the country level. They found a strong log-log correlation (R^2 = 0.97) between "lit area" in square kilometers, derived from luminosity series and defined as all areas that have persistent detectable lights, and GDP for 21 countries (Elvidge, et al. 1997). Similarly, Doll, Muller, and Elvidge (2000) found a strong log-log relationship between total lit area (measured as km^2) of a country and its GDP for 46 countries that have at least 10 identified urban polygons (R^2 = 0.85). Doll, Muller, and Elvidge (2000) also extrapolated 1° latitude by 1° longitude grid cell GDP by applying the coefficient of log-log relationship obtained at the country level, and concluded that lit-area-derived PPP-GDP grid map modeled global economic activity "very well."Sutton and Costanza (2002) also derived GDP per square kilometer by applying the correlation coefficient of luminosity and GDP at the national level and generated fine resolution map of economic activity.

More recently, studies have used nighttime image to predict income per capita at the sub-national level (Ebener et al. 2005). Their results show that lit area and percent frequency of lighting can predict GDP per capita at the national and sub-national level. They write that, when climate and agriculture are considered, the model yields better results in predicting GDP (Ebener et al. 2005). Later Sutton et al.

(2007) improved Ebener's model by adding estimated urban population at the state level to solve the problem of saturation in traditional luminosity image and applied the new model to four countries – China, India, Turkey, and the U.S. The variable "urban population" of each state of a country was estimated by using a log-log linear relationship between the aerial extent of urban areas and population. Then they used the estimated "urban population" as a predictor for GDP values for sub-national administrative units, and concluded that "spatial disaggregation of estimates dramatically improves aggregate national estimates" of GDP. Nighttime luminosity has reached a level of accuracy that it is currently under consideration for use in target acquisition by the U.S. military.

Few studies have undertaken a formal analysis comparing luminosity with traditional output measures. The present study was developed in parallel with a pathbreaking study by Henderson, Storeygard, and Weil (2009) using much the same statistical approach. We are indebted to that study for suggesting a modification of our original modeling as well as for pointing to an approach for measuring errors. That study suggests that the growth rate of true income estimated from luminosity data is more accurate than the World Bank estimates, especially for countries having poor quality data (the "D countries" defined by the Penn World Tables (PWT)). Among these countries, they argued, "general growth is underestimated in the WDI for countries with low measured income growth rates, and overestimated in the WDI for some countries showing very high growth rates…. For D countries at the tails of high or low recorded growth (Myanmar, Liberia, and Congo), lights strongly amend recorded growth rate."

In light of the apparent successes of past studies in applying luminosity measures, the authors undertook a detailed study of the relationship between luminosity and economic activity. We first describe the construction of the data and then present our results.

Description of the Luminosity Data

The data on nighttime lights is one of the most striking images of the presence of human settlements (see Figure 1). One look at the map makes it immediately clear that lighting is closely related to the density of economic activity. We begin with a discussion of how the nighttime lights data are constructed.

The primary nighttime image data were gathered by U.S. Department of Defense satellites starting in the mid-1960s to determine the extent of worldwide cloud cover. It was later declassified and made available as the Defense Meteorological Satellite Program Operational Linescan System (DMSP-OLS). The DMSP satellites collect a complete set of earth images twice a day at a nominal resolution of 0.56 km, smoothed to blocks of 2.8 km (30 arc-seconds). The data, in 30 arc-second resolution, covers 180° West to 180° East longitude and 65° North to 65° South latitude.

Creation of a light-time lights data is an undertaking of monumental difficulty. The raw data are heavily processed on an empirical basis to correct for various optical and atmospheric distortions.[2] Among the important steps are the following: gathering the raw data; identifying clouds; removing glare; identifying the visible–near infrared emissions (lights); gridding the data; calculating the number of cloud-free detections and the number of light detections in cloud-free areas; and manually editing the results to conform to shorelines and other geographical phenomena. Examination of the data indicates the presence of atmospheric disturbances (such as water vapor), scanning errors, errors in mapping of the earth's topography, and blooming or overglow across pixels.

There are different versions of the data. The two major sources are the raw data and the stable lights data. The stable lights version removes ephemeral events such as fires and background noise, while the raw version does not. Because of all the adjustments, the processed data are substantially different from the raw data in ways that are not immediately apparent. Currently DMSP-OLS Nighttime Lights Time Series provides cloud-free digital values of nighttime light image generated from persistent lighting of human activities, and the annual data are available from 1992 to 2008.

We decided to use the raw data because of concerns that low-density cells would be set to zero in the stable lights data set and because some of the processing was related to discrepancies between the lights data and demographic data. A case could be made for using the stable lights version, but from a statistical point of view errors in measurement of luminosity pose less of a difficulty than inappropriate filtering at the low end.

Figure 2 shows a detailed map of Los Angeles that illustrates some of the issues with the use of luminosity as a proxy. One issue is that road networks show distinct luminosity. Generally, road traffic will not be captured in an accurate spatial way by conventional accounts because it will only be recorded by equipment or fuel purchases, not by actual location of traffic. The figure also shows the bleeding that occurs outside of the areas of human settlement and is significant as much as 50 km offshore. These are relatively unimportant at the scale we employ but clearly introduce errors in measurement of the spatial luminosity data. Another feature shown in the figure is the saturation at the high end, with no detectable distinctions among most cells near the city center.

The annual data are presented as digital numbers (DN) from 0 to 63. These can be converted to radiance by the formula Radiance = (Digital Number)$^{3/2}$ x 10^{-10}

---

[2] For a comprehensive discussion, see the description by Christopher Doll (Doll, 2008).

Watts/cm²/sr/um (Elvidge, et al. 1999). For our purposes, we aggregated the data up to 1° x 1° grid cells to match our finest resolution economic data. For this aggregation, we took each pixel's DN, raised it to the power of 1.5, and then summed these radiances over all pixels in the grid cell. (Aggregation errors would be introduced unless this transformation is used.) For a complete grid cell, this would include 120 x 120 pixels. This procedure yielded 40,570 grid cells with data for 17 years, with overlaps of 12 satellite-years.

A preliminary examination of the data indicates that there are substantial measurement errors in both the time-series and the cross-section of the data. For example, we examined the correlation of logarithm of gridded luminosities for satellites F14 and F15 for the year 2000 yielded $R^2 = 0.965$. Regressions of the logarithm of luminosity for the same year across satellites have standard errors in the range of 0.20. Slightly smaller numbers are found for year-to-year variations of individual satellites. The interpretation of these numbers is that the measurement error for individual grid cells is in the order of 20 logarithmic percent. There has been relatively little analysis of the error patterns of nighttime lights. For example, satellite F14 is consistently dimmer than the earlier satellites. This will not pose a major difficulty in our statistical analysis below, but it does raise questions about the overall reliability of the lights data. The size of the errors indicates that luminosity is at best a noisy indicator.

We then merged our grid cell luminosity with new Gecon 3.3 data (available at gecon.yale.edu). Since GEcon data includes country boundaries, and cells that fall between countries are divided into multiple subcells, we divided luminosity values for those subcells based on the share of population of subcells. Generally, to reduce noise caused by remote areas and extremely small cells, we limited our analysis to cells located between latitude 60° N and 60° S, with grid-cell output value of 1990 larger than 10 million U.S. dollars, and with land area at least 10 percent of the area of the grid cell. At the country level, we used GDP PPP values (at constant 2005 international dollar) from World Bank from year 1992 to 2008 as our country GDP and aggregated grid cell luminosity by country to obtain county luminosity for corresponding year.

A First Look at the Data

It is always useful to do a preliminary scan of the data to see what emerges. For this purpose, we have taken a benchmark year for the GEcon data, 2000, and compared the gross cell product (GCP) from the GEcon data with the output-luminosity proxy. Appendix Tables A-3 and A-4 show raw data on luminosity and output.

We show in Figure 3, 4, and 5 the bivariate double-logarithmic plot for the world, for the United States, for tropical Africa. For these plots, we have chosen only grid cells with RIG > .1 (i.e., only cells where the grid cell observation covers at least 10 percent of the grid cell). This avoids border effects as well as difficulties with luminosity that arises near coastlines.

Four points are immediately clear. First, there is a strong positive relationship between luminosity and output. The global R is 0.61 for the global curve. This observation confirms that there is a strong positive relationship between our luminosity proxy and measured GDP at a disaggregated level.

Second, the relationship works very poorly at the low end of the output/luminosity spectrum. This can be seen most clearly for Africa. For ln output density less than -9, the relationship is essentially flat. To understand the numbers, log output density of -9 represents an output density about $100,000 per square km. To put this in perspective, this is approximately the output density of one-thirtieth of a retail gasoline station in the U.S. The meaning of this is that the luminosity at the low end is dominated by noise and errors of the kind discussed above rather than economic activity. Since there are 2270 African grid cells and only 189 observations above the threshold, we may have difficulty predicting output with luminosity unless output is extremely concentrated.

Third, the relationship is strongly positive for high-income regions, but it is extremely noisy. For example, the regression of log of luminosity output on output for grid cells with log output density greater than -8 has an elasticity of 0.74 ($\pm$ 0.013), with a R^2 of 0.573. However, the standard error of the equation is 0.661, indicating that the error is in the order of magnitude of $\pm$ 100 percent. Of course, this might be because of error in the GEcon data, but evidence suggests that it arises from imprecision in the luminosity data.

Fourth, the results for the United States are useful because of the presumption that the regional data for the US are relatively accurate. The U.S. data show the phenomenon of noise at the lower end and has a standard error that is actually larger than the global standard error (being 0.764). Additionally, if we look at the very high end, there is a suggestion of a new problem – saturation. The problem at the high end is that the satellite sensors become saturated and are therefore top coded. This is a problem particularly for the largest U.S. cities, New York, Los Angeles, and Chicago and explains why the curve flattens out at the high end.

The major points that arise from this preliminary look at the data are two. First, based on our examination of our constructed luminosity-output measure at the cell level, it appears to be a very noisy indicator of output using the GEcon cell data as a benchmark. This is true even in countries like the United States where we have the most reliable disaggregated data. Second, the luminosity data are likely to be

least useful in areas where it is most needed – in countries with low output density. Only 6 percent of African grid cells are ones with output beyond the threshold where the luminosity data is highly correlated with the output data.

The Analytical Background

The analytical approach used here is similar to that proposed in Henderson, Storeygard, and Weil (2009). However, we derive different formulas, use different data sets, have different priors for resolving the errors-in-measurement issue, and have dramatically different results.

Constructing "true" GDP

Our ultimate purpose is to determine whether luminosity contains useful information for constructing economic data at either the national or regional level. The most natural place to hope for value added is for countries with poor data or those with little or no data at a subnational scale.

For this purpose, we define the different measures as follows:

Y = measured output (GDP in international U.S. $)
Y^* = true output (GDP in international U.S. $)
X = synthetic measure of output (GDP in international U.S. $)
M = measured luminosity (index value)
Z = luminosity-based measure of output (GDP in international U.S. $)
i = grid cell (here 1° latitude by 1° longitude)
j = country
k= country grade (A, B, C, D, E)
t = year
y = ln (Y)
x = ln (X)
m = ln (M)
z = ln (Z)
ε_i = measurement error in GDP
ξ_i = measurement error in luminosity
u_i = random error in output-luminosity relationship

For notational purposes, we define $x_{i,j}(t)$ as the value of variable x in grid cell i and country j averaged over year t. We assume that there is a "true" level of output for each country and grid cell, which is measured with error:

(1) $y_i = y_i^* + \varepsilon_i$

Measured output is assumed to have a unitary coefficient on true output. This assumption is not completely innocuous as there may be systematic growth mismeasurement due, say, to incomplete source data or infrequent observations.

Luminosity is subject to measurement error (satellite, calibration, and other):

$$(2) \qquad m_i = m_i^* + \xi_i$$

There is a structural relationship between luminosity and true output as follows:

$$(3) \qquad m_i = \alpha + \beta y_i^* + u_i$$

The error in (3) arises from several sources. One important reason why the relationship is noisy is that luminosity is sampled at night, whereas economic activity is generally concentrated in the daytime. More important, the light intensity differs greatly across sectors. The use of electricity per dollar of output in different sectors provides a rough idea of how light-intensities might vary. In the 2002 U.S. input-output tables, the electricity used per unit output of real estate was 200 times greater than that of software. Similar differences are seen across other sectors. This example suggests that industrial composition across countries is likely to make the output-luminosity relationship in (3) relatively noisy.

We want to construct a luminosity-output proxy from these relationships. We have measurements of all variables over time and space at the national and grid cell levels. However, we need to develop measures of the error of measurement of national and grid cell output, the coefficient on luminosity as well as the error in the structural relationship in equation (3).

Our procedure is first to estimate equation (3). This provides an estimate of the structural coefficient β. We then do an errors-in-variable correction for β from our prior estimates of the error in GDP to get a corrected estimate of β. The corrected coefficient is calculated as:

$$(4) \qquad \tilde{\beta} = \left(\frac{\sigma_{y^*}^2 + \sigma_\varepsilon^2}{\sigma_{y^*}^2} \right) \hat{\beta}$$

Where $\hat{\beta}$ is the estimated coefficient in equation (3); σ_ε^2 is a priori error of true output; $\sigma_{y^*}^2$ is the estimated variance of true output. (We obtain estimates of the error variances of both conventionally measured output and our luminosity-output proxy in a manner that is described below.)

We then estimate luminosity output as follows by inverting (3):

(5) $z_i = (1 / \tilde{\beta}) \, m_i$

where z_i is the log of our luminosity-output proxy and $\tilde{\beta}$ is the corrected coefficient in equation (3).

Next, we will construct an optimal synthetic measure of output by taking the optimal weights of conventional measures of output and our luminosity-output proxy:

(6) $x_i = \theta y_i + (1 - \theta) z_i$

where
        X_i = new synthetic measure of output
        $x_i = \ln(X_i)$
        θ = optimal weighting fraction on conventional output measures

The key variable of this study is θ, which is the optimal weighting share on conventional output measures. The central question we address is whether we can significantly improve conventional measures of output using luminosity. If the measurement error of synthetic output has been significantly reduced compared to the prior error ε, luminosity can be a useful proxy.

The error variance of x_i can be expressed as a function of θ as follows:

$$V(\theta) = E[\theta y + (1 - \theta) z - y^*]^2$$

$$= E[\theta(y^* + \varepsilon) + (1 - \theta)\gamma(\beta y^* + u) - y^*]^2$$

$$= E[\theta\varepsilon + (1 - \theta)\gamma u]^2$$

$$= \theta^2 \sigma_\varepsilon^2 + (1 - \theta)^2 \gamma^2 \sigma_u^2$$

Where $\gamma = 1 / \tilde{\beta}$

Maximizing with respect to θ:

$$V'(\theta) = 0 = 2\theta\sigma_\varepsilon^2 - 2(1-\theta)\gamma^2\sigma_u^2$$

$$\theta\sigma_\varepsilon^2 = \gamma^2\sigma_u^2 - \theta\gamma^2\sigma_u^2$$

(7) $$\theta = \frac{\sigma_u^2}{(\beta^2\sigma_\varepsilon^2 + \sigma_u^2)}$$

The expression for the optimal weight is intuitively appealing because the weight on measured output is 1 when (a) $\beta = 0$ (when luminosity is unrelated to true output); (b) when $\sigma_\varepsilon^2 = 0$ (when there is no measurement error in output); or (c) when σ_u^2 is very large (when the relationship between luminosity and output is extremely noisy.)

Differences in reliability by country

There are several modifications of equation (7) that can be pursued, but we use only one in the present study. We know both from direct observation and from studies that the reliability of national statistical systems vary greatly. We therefore will be particularly concerned about the applicability of our estimates to different countries. For this purpose, we will divide countries into different "grades," k = A, B, C, D, and E. (We discuss the methods further below.) This leads to the slight modification of (7) in which the optimal weights depend on the error output estimates by grade:

(8) $$\theta^k = \frac{\sigma_u^2}{\beta^2\left(\sigma_\varepsilon^k\right)^2 + \sigma_u^2}$$

where σ_ε^k is the error for measured output for country grade k.

Estimates of the Error of Measured Output and Luminosity

The last section showed that the statistical model for deriving the optimal weights on conventional GDP measures and luminosity is underidentified and requires estimates of the two error terms as well as the coefficient in the luminosity equation. We discuss each of these terms. The first two are relatively simple, while the third is extremely difficult to estimate.

Errors of measurement of luminosity

The errors in the luminosity equation (3) can be estimated from both different satellites and different years of the same satellite as well as from estimates of the equation. The estimates are straightforward, and we therefore describe them verbally. Satellites differ in their optical quality and may degrade over time. In addition, there is sampling variation because of physical factors. We can estimate the error of satellites by examining the difference between the luminosity of different satellites for the same year because we have 12 year-satellite overlaps (See Table A-5, Figure A-1 and Figure A-2). In addition, because the satellite noise is so large relative to output changes, we can look at adjacent years to measure the errors for individual satellites. These are estimated for both cross-sections and growth rates. To give a flavor of the estimates, we estimate that the measurement error of log luminosity using the same year but different satellites is between 0.136 and 0.369, depending upon year. The estimated error from year-to-year changes of the same satellite is between 0.122 and 0.338 depending upon year and satellite. (See Table A-5.) These indicate that the error of measurement of luminosity is in the range of 14 to 37 logarithmic percent. In our final estimates, we use mean square residual obtained from regression results of equation (3), assuming the error terms in the structural equation are primarily caused by errors in luminosity.

Structural coefficient in luminosity equation

Estimates of the coefficient β in equation (3) are a second ingredient in the estimation of the optimal weights on the two measures. The estimates here are straightforward. While the coefficient estimate in equation (3) will suffer from bias because of errors of measurement, the error is relatively small. We have also used an alternative approach of instrumental variables, using a suite of cell variables such as climate, proximity to ports, and population, and the results differ very little from the correction using classical measurement error except for the cross section output of grade-E countries (the last row in Table A-6). Table A-6 in the Appendix provides the IV estimates of the luminosity equation for all grid cells. Because the results of the IV approach are very similar to the other approaches, we ignore those results in the present discussion.

Errors of measurement of output

The most difficult parameters to estimate are the errors of conventional GDP measures. This discussion deals primarily with issues of errors in national economic accounts and is relatively technical. Those who are not concerned about the details can skip to the next section.

We begin with a definition of "error." Depending on country and time period, there are many errors in measuring national output (conceptual differences, data sources, index construction, sectoral definitions such as how to treat home

production, and so on). The question might be posed as follows: Relative to some ideal measure of national output, what would be an estimate of the measurement error of output and growth of output? Errors would include not only the accuracy of source data and imputations, but the scope of the exclusions, such as home production, subsistence farming, illegal activity, smuggling, and the like. For our purpose, we define the "ideal" measure of output as that one corresponding to the line delineating national output by the System of National Accounts (1993). As we noted above, it may sometimes occur that luminosity will more closely track the actual location of economic activity than conventional accounts. Examples of this phenomenon include road traffic (see Figure 2), fishing boats, and gas flaring associated with oil and gas production.

We can distinguish two different kinds of errors. The first is *time-series errors.* Measures of output here generally keep the conceptual basis of the measures as well as the data sources constant over time (at least for short periods); errors here will arise primarily from errors in the source data or errors in aggregation. Moreover, since there are two or three alternative methods of constructing national output (e.g., income and expenditures), we can examine the statistical discrepancy to make a first estimate of the size of the measurement error (conditional on the methods).

A second kind of measurement error is *cross-sectional level or density errors.* These would apply to comparisons among countries or regions. Cross-sectional errors will be much broader and include source-data errors, differences in concepts, price measurement errors, as well as errors in measuring exchange rates among different currencies. Moreover, there are no identities that can be relied upon to provide alternative estimates of the kind that produce the statistical discrepancy in country measures. We would therefore expect the cross-sectional errors to be much larger than the time-series errors.

In addition, the present study examines both country output data and grid-cell output data. We will therefore consider errors in both country and grid-cell as well as time-series and level estimates.

We must alert the reader at the outset that estimation of errors of conventional GDP measures is a vast topic that has received very only modest attention in the literature on national accounts. Our estimates therefore are very tentative. However, even order-of-magnitude estimates may be sufficient for the purpose at hand.

Errors in national-level data

Time series errors: general

We begin by surveying estimation errors for the U.S. The most systematic measure of error is the statistical discrepancy (SD) between income and expenditure

accounts. The SD is relatively "unmanaged" and is therefore a relatively clean measure of the measurement error conditional on the definitions. This absolute value of the change in the ratio of the annual SD to GDP averages around 0.56 logarithmic percentage points (per year) of GDP for the 1929-2008 period. If the true value is the average of income and product, this would indicate that the error in the growth rate is 0.28 percentage points. As an alternative estimate of the errors, BEA has examined the change from the third annual estimates to current methodology and found an average absolute revision of the annual growth rate of real GDP of 0.41 percentage points for 1983-2006 with 0.29 for nominal GDP (Fixler and Grimm, 2008).[3] This second calculation includes some changes in methodology as well. From these two calculations, this indicates a lower bound for the measurement error of the growth rate of real output of around 0.3 percentage points per year for countries with high-quality statistical systems.

Time-series errors from index-number differences

One of the methodological differences among countries involves the index-number techniques used in determining growth rates. Most high-quality systems currently use superlative techniques (such as Fisher's Ideal index), while other countries (such as China) continue to use Laspeyres indexes. BEA's calculations indicate that for the U.S., the average error is around 0.3 percentage points per year (this being always positive). Larger biases would be expected in countries with particularly rapid structural change.

Cross-sectional errors from revisions and methodological differences

A second set of estimates concerns the level of GDP (or GDP density per square km when used in conjunction with luminosity). For the U.S., we can use the average ratio of SD to GDP as a lower bound estimate of the measurement error conditional on the methodology. Over the period 1929-2008, the SD was 0.50 percent of GDP. Again, if the average of income and product sides is the correct estimate, this would indicate an average measurement error of 0.25 percent of GDP as a lower bound for high-quality statistical systems.

However, it is likely that cross-sectional errors will arise from other issues such as measurement error, sectoral inclusion, exchange rates, and even idiosyncratic country effects. One way to estimate the methodological differences is to examine the change in the level of nominal GDP in different vintages of estimates. Comparing current estimates with those of 1964 and 1973, BEA found an average error of between 3.1 and 3.3 percent. Most of these are probably definitional (such as the inclusion of software in investment) rather than measurement error, however. If we go back to the earliest estimates of national income by Simon Kuznets published

---

[3] All data comparisons are based on data from the BEA web site at www.bea.gov as of February 2010.

in 1934, we find much more substantial differences, however (Kuznets, 1934). The difference between the original estimates of national income by Kuznets and the current estimates by the BEA is 17 percent for 1929-32. The average absolute difference in the logarithmic growth rates of nominal national income was 4.5 percentage points for this period. This is a very demanding test, of course, because these first estimates were at the dawn of national income accounting, and the period was the descent into the Great Depression. These are suggestive of the very substantial cross-sectional differences that can arise in immature accounting systems as well as the measurement problems that can arise in economic crises.

Cross-section differences from exchange-rate calculations

One of the thorniest issues in country comparisons is the conversion from national currencies into a common unit. Common practice today is to use PPP, or purchasing-power parity, exchange rates rather than market exchange rate. While there is (in our view) no question about the appropriateness of these measures conceptually, the practice of calculating them has proven extremely difficult, and in some cases, such as the appropriate multilateral weights, unresolved.[4]

One point of clarification should be made about PPP-corrected estimates. In some but not all systems, introduction of PPP exchange rates should have no effect upon the country growth rates. To a first approximation, PPP adjustments are to the level of output, not to the growth rates.

In our estimates of measurement errors, we will rely on the country grading system defined by the Penn World Table and earlier authors. Countries are assigned subjective quality grades from A to D by the authors of PWT based on several criteria of the data. We will add grade E for those countries with essentially no statistical system which are missing from the PWT and other standard sources. Appendix Table A-1 gives the number of countries by grade while A-2 gives the grade sheet for country systems.

Very few countries receive the grade of "A" and a substantial number are C or D. Perhaps the most pertinent estimates are those provided by Robert Summers and Alan Heston on the margin of error of the estimates. These are most usefully interpreted as the cross-sectional errors that arise from price and aggregation estimates involved in moving from market exchange rates to PPP exchange rates. Originally, Summers and Heston provided estimates of the margin of error (which is defined as the root mean squared error) by different country grade. These are shows in Table 1.

---

[4] There is a vast literature on the subject. For a recent review, see the article by Deaton and Heston (2008).

Estimated margin of error in PPP cross section	
Country grade	Error in level
A	9%
B	15%
C	21%
D	30%

Table 1. Cross-sectional errors estimated by Summers and Heston

The A countries would be representative of countries such as the United States. Note that the margin of error is much greater than the average statistical discrepancy, but not as large as the difference in the Kuznets-BEA estimates. We will adopt the margin of errors in Table 1 for our estimates by countries.

The errors in the international data have been recently examined in a comprehensive study by Simon Johnson, William Larson, Chris Papageorgiou, and Arvind Subramanian (2009). They examined the revisions of estimates of both the level of the price index and the growth in real GDP across countries between Penn World Table 6.1 and 6.2, shown in Table 2. One feature of the study was to examine the changes or revisions in cross-sectional differences due to price changes. These revisions arise from several sources: revisions in source data, methodological changes, and most importantly from changes in the international price data. These can be interpreted as an approximation of the error in PWT 6.1. For example, if the PWT 6.2 was exactly correct, then the errors would be the errors in the PWT 6.1. There is no theoretical way to determine whether these should be higher or lower than the Summers-Heston margins of error, but they do tend to be considerably smaller. Note that the Summer-Heston grades applied to an earlier version of PWT and should presumably be smaller on average in the latest version.

Estimated margin of change in PPP price between PWT 6.1 and PWT 6.2	
Country grade	Change in level
A	2.9%
B	2.9%
C	10.9%
D	22.6%

Table 2. Cross-sectional errors as estimated by the change in country PPP level between PWT 6.1 and PWT 6.2

Johnson et al. focus primarily on the revisions of the growth rates. Based on their results, we have compiled estimates of the revisions to the growth rates of real GDP by country grade, and the results are shown in Table 3. This table shows the revisions in growth rates across countries between Penn World Table 6.1 and 6.2. We transformed the 29-year growth rate differences to 1-year growth rate differences by assuming that the differences by year were independent. As with the cross-sectional differences in Table 2, there is no necessary relationship between the errors and the revisions.

The number for the U.S. using this methodology is 0.3 percentage points per year, which is virtually identical to the number derived above. The errors for the A countries are consistent with the estimates for the U.S., indicating about double the measurement error for the U.S. Part of the difference, however, is probably due to the fact that the U.S. dollar is the numeraire. The revision numbers for the D countries are astoundingly high, indicating a revision that is actually more than double the mean growth rate. However, note that 135 of the poor quality countries are "C" while only 13 are "D."

At the same time, some of the worst statistical systems – Iraq, West Bank and Gaza, and Afghanistan – are not even in the PWT data set.

Difference between PWT 6.1 and 6.2			
Country grade	29-year growth rate	1-year growth rate*	
		Average	Median
A	0.108%	0.583%	0.350%
B	0.134%	0.720%	0.700%
C	0.576%	3.104%	1.642%
D	1.644%	8.852%	6.731%

* One-year growth rate is 29-year multiplied by square root of 29.

Table 3. Revision to growth rate of real GDP

Differences from compilation errors

A final issue comes from the "transmission errors" that come when those who compile databases make errors of omission or commission. We have generally relied on the World Bank compilation of national accounts statistics because it is the most complete. One set of errors would arise because the data are from earlier vintages of publications from the national statistical offices. For data available on the same date, the ratio of the World Bank measure of U.S. nominal GDP to the official version

erred by an average of 1.1 percent over the 1960-2006 period. The average growth rate of real GDP differed by 0.18 percentage points. By contrast, the IMF database associated with the *World Economic Outlook* had essentially no errors.

Grid-cell estimates

In our estimates below, we use grid cell as well as country output estimates. The grid cell output data have higher errors than the national data, but they may be useful because of the much higher resolution than country data. We have about 20,000 non-zero grid cells as compared to somewhat less than 200 countries.

The tradeoff is that estimating the grid-cell errors is more challenging because their estimation is in its infancy. We would consider the national level and growth estimates to be a lower bound on our grid-cell estimates. The major approach available to estimate potential error is similar to that used above -- to examine changes in estimates of levels of PPP output for individual grid cells across GEcon revisions. The revisions have added considerable accuracy by using improved maps, better population estimates, and improved imputations. In addition, the GEcon estimates have added output estimates for "E" quality countries for which data are not generally available, such as Somalia and Afghanistan. Revisions have been completed for 34 of the countries in the data set, with 7756 grid cells.

The revisions considered here are from the first published version (GEcon 1.3 from 2005) to the latest version (GEcon 3.32 in 2010). One comprehensive set of revisions has been in the gridded population data, which were a complete revision cycle. The economic data have been thoroughly revised only for a subset of countries, primarily the large countries such as the U.S., China, Brazil, and Russia.

Table 4 shows the revisions measured as the standard deviation of the ln of first to last estimates for GDP per grid cell and population per grid cell for countries with revisions. These should be compared with the cross-sectional results above. The results suggest very high potential error for the grid cell output estimates, even for grade A countries.

| Country grade | Standard deviation of | |
| | ln (First estimate/last estimate) | |
	Output	Population
A	36.1	33.4
B	24.4	20.9
C	79.5	33.4
D	57.4	41.8
E	na	na

Table 4. Revisions in estimates in grid-cell data from GEcon data

We will also be using growth rate estimates for the grid cell data. These have been developed from a combination of gridded population for different years and country data on regional or national per capita GDP. At present there are no revisions of the data for comparison purposes. Work on regional GDP estimates suggest, however, that relative per capita incomes in most countries are relatively stable over time, so the errors in the growth rates of gridded data are likely to be only slightly above the estimates of the errors for the population and for the national GDP data. We have made tentative estimates here, but we recognize that these error estimates have only a sparse empirical basis.

The last column in Table 4 provides the estimates of the estimated errors for grid cell output data for each of the five groups of countries. These are clearly very high, and are still tentative given the tentative nature of the GEcon data. The question is whether, given the very high potential errors in these data, the luminosity information can be used to improve the estimates.

Final Estimates

Table 5 shows our estimates that we have used in our empirical estimates below. The estimates for country growth and cross-sectional errors are based on the findings discussed above. For *countries,* the cross-sectional are largely consistent with the Summers-Heston grades, while the growth estimates are largely drawn from the Johnson et al results. For *grid cells,* we decided to double the estimates of the country errors for the cross-sectional errors. The assumption of doubling is based in a rough way on the revision estimates for the GEcon data discussed above. We also note that the results for the "E" countries are based on our own judgments as these data are generally not found in the international data bases. For the time-series numbers for grid cells, we have doubled the errors for the high-quality data, but have added a maximum of 1 percentage point to the cell growth errors. This number comes from estimates we have made about the relative divergence of regional from national data in several countries. In addition, we use weighted average over the number of observations to generate error estimates for all countries and all cells in the following analysis.

It must be emphasized that estimating the error of standard measures of GDP is extremely difficult, and our results should be taken with some caution because of this problem.

Country grade	Estimates for country output		Estimates for grid-cell output	
	1-year growth rate	Output level	1-year growth rate	Output level
A	0.600%	10%	1.200%	20%
B	0.800%	15%	1.600%	30%
C	3.000%	20%	4.000%	40%
D	5.000%	30%	5.000%	60%
E	6.000%	50%	8.000%	100%

Table 5. Estimates of errors of national and gridded GDP data used in estimates of combined measures of output

Some Preliminary Tests

We can get the basic point with the simplest test. For this test, we take a group of 103 grade-C countries (those with the moderate quality statistical systems) and construct growth rates of both real GDP and luminosity over the 1992-2008 period. A linear regression of the two growth rates has a R^2 of 0.22.

We estimate that the error of the real GDP growth rate of C countries is 3 percentage points per year. Using these data, we calculate that $\hat{\beta}$ = 0.25. The error variance of the 17-year GDP growth rate is 0.015, while the error variance of 17-year luminosity growth is 0.046. From the formula for the optimal weights, we get:

$$\theta = \frac{\gamma^2 \sigma_u^2}{(\sigma_\varepsilon^2 + \gamma^2 \sigma_u^2)} = \frac{(0.25)^{-2} \times 0.046}{(\ 0.015 + (0.25)^{-2} \times 0.046)} = \ 0.980$$

For this simple example, then, the optimal weight is 98 percent on measured GDP growth and 2 percent on luminosity growth. The intuitive reason for the low weight on luminosity is that it is too noisy a proxy to be useful for measuring output in countries with moderately accurate statistical systems. If we double the estimated error of national growth rates to 6 percentage points per year, the optimal weight on luminosity will increase to 0.24. This simple example suggests that luminosity will not be much help in countries with moderately accurate statistical systems unless the errors of growth rate of these countries are significantly underestimated.

What about the low-quality countries? For these countries, we take an average error of 6 percentage points per year. If we perform the same experiment for the 9 countries with grade E, we get:

$$\theta = \frac{\gamma^2 \sigma_u^2}{(\sigma_\varepsilon^2 + \gamma^2 \sigma_u^2)} = \frac{(0.285)^{-2} \times 0.0045}{(0.061 + (0.285)^{-2} \times 0.0045)} = 0.476$$

In this simplest case, luminosity data have approximately the same information content as conventional output data. These preliminary tests suggest that only for countries with the most rudimentary accounts (such as Angola, Chad, Democratic Republic of the Congo, Republic of Djibouti, Eritrea , Niger, Sudan, Uganda and Yemen), can the luminosity data be considered as reliable a measure of output growth as standard output measures.

Formal Tests of Luminosity and National Output

The informal tests in the last section are only suggestive because they rely primarily on one or two country grades. An alternative test is to look at data from all country grades and estimate the optimal weights on luminosities. For these countries, we conducted two sets of estimates. The first set is for the 17-year growth of output from 1992 to 2008, while the second is output density measured as output per unit area. We look at both country data and grid-cell data. The grid cell results generally have large samples (N > 250). The sample size for countries is small particularly for E countries (see Table A-1).

Our ultimate goal is to see how much luminosity can contribute in constructing the "true" GDP measures. We do this by estimating the weight θ in equation (7). We report θ estimates for countries and cells in both the 17-year growth rates and the output density. We also run a separate analysis for observation with low density GDP, because many of the low-quality statistical systems are also in low-density regions. For countries, we defined low-density observations as ones with 1990 GDP per unit area one standard deviation below the mean GDP density. For grid cells, we defined cells with ln output density above -9 in the year 2000 (about $100,000 per square km) as high density-observations, while the rest of the cells are designated as low-density observations.

Table 6 reports the results for countries, while Table 7 reports the results for grid cells. In addition, we show the estimated weights on luminosity graphically in Figures 6 and 7. Comparing the estimated θ weights leads to the following conclusions.

First, examine the results for the time series, shown at the top of Tables 6 and 7. There is essentially no information in the luminosity signal for A and B countries,

with the weight on luminosity being less than 1 percent. There is a small weight on luminosity in the C countries, with the weight being between 2 to 7 percent. The results on D and E countries are inconsistent, but the general finding is that there is substantial weight, with a maximum weight on luminosity being over 60 percent for E countries with low GDP density.

Second, consider the cross sectional values, or output density estimates, shown as the bottom panels in Tables 6 and 7. Since the areas are well measured, these involve uncertainties about output. The estimates here are consistent across both countries and grid cells. For A through D countries, there is a small value added of luminosity, with the range of weights on luminosity from 1.0 to 14.0 percent. By contrast, there is substantial weight on luminosity in the E countries, with the weights being over 30 percent for both cell and country data. The interesting finding here is that luminosity is considerably more useful than official data, having a 94 percent weight, for high-GDP-density cells in E countries. This result is confirmed by its highest weight for those observations in the IV analysis (the last row in Table A-6). One possible reason is that economic activity is more concentrated in urban areas in low-density and therefore poorer countries than in developed countries, and these cells are more likely to locate in populous, more developed cities, such as the national capitals. Thus, luminosity could be a good proxy for their output measures due to the lack of reliable regional data in these countries.

Third, consider the relative contribution of luminosity for grid cells and countries. We originally expected that there would be much more value added for luminosity in grid-cell output measures because of the low quality of regional economic data in most countries. This turns out not to be the case. In general, the information content of luminosity is approximately the same for grid-cell data and for country data. The only exception is that luminosity adds more information for E countries (about 60 percent) than for cells in the E countries (about 35 percent) in time series analysis.

Fourth, consider the results for low-density observations. We had expected that the value added of luminosity for the low-density observations would be much higher than for all observations. The results are more complicated. For *countries*, the weight on luminosity is almost uniformly higher for low-density observations. For *grid cells*, the results are mixed. The weights on luminosity are generally higher for the low-density cells than for all cells for time series, but the relative weights for cross section are reversed. We have not investigated the reasons why luminosity adds less information for low-GDP-density cells for cross section. We hypothesize that this phenomenon may arise because economic activity is more concentrated in large urban areas in poor and therefore low-density countries.

As a final test, we examine output and luminosity data for U.S. states. The reason to look here is that U.S. states probably have the most accurate cross sectional data set constructed with a uniform and high-quality methodology. They are less accurate than the U.S. national data, but probably better for comparative purposes than any other regional comparison. We obtained real annual GDP by state from the U.S. Bureau of Economic Analysis and reallocated grid cell luminosity to state luminosity. We assumed that the error of annual growth rate is 0.8 percent and the error of output density 15 percent (equivalent to the errors for grade B countries). Using the standard approach in equation (7), we calculation that the optimal weight on luminosity is 0.0012 for the growth rate and 0.102 for the density measure. This result further indicates that the contribution of luminosity is of no value for time series estimates but is of modest value for cross sectional estimates for high-quality data systems.

Summary

This project began with the hope that luminosity measures could be used to improve the estimates of output at the regional level, particularly for countries with low-quality data systems.

The results confirm that the use of luminosity as a proxy for output will be most beneficial for countries with the poorest statistical systems, those that receive a D or E grade. Luminosity has very limited value as a proxy variable for output for countries with moderate to excellent statistic systems because it is so noisy and the errors in the standard economic data are relatively low. But we do find that luminosity can significantly improve the cross-sectional output measures for rich cells in poor countries.

We draw one major conclusion for the development of economic statistics from this study. There are natural concerns about improving the quality of economic data in countries with low-quality statistical systems. Given the results here, we suggest that major improvements will come primarily from the standard procedures for designing and improving national economic accounts. These involve improving methods as well as increasing the depth and reliability of source data. We see relatively little utility in relying instead on proxy measures such as luminosity except in the regions with the most deficient data. The reason for this conclusion is that, even in the best of cases, the errors in the luminosity data at present are too large to provide accurate indicators of economic activity. Once a country has moved from the E or D grade to the C or B grade, there is almost no value added in using luminosity for estimates of real economic growth. However, in the interim, and particularly for periods or countries where economic data are extremely weak, luminosity data may be a useful supplement to current indicators.

| | Country sample for time series | |
Country grade	All countries	Low density countries
All	0.9747	0.9547
A	1.0000	na
B	0.9986	na
C	0.9791	0.9548
D	0.8911	0.9721
E	0.4745	0.3655

| | Country sample for cross section | |
Country grade	All countries	Low density countries
All	0.9773	0.9291
A	0.9829	na
B	0.9500	na
C	0.9754	0.9650
D	0.9778	0.9604
E	0.6809	0.6529

Table 6. Estimated optimal weight on measured output for countries of different grades using 17-year growth rates and for cross-sectional density

| | Cell sample for time series | | |
Country grade	All cells	High density cells	Low density cells
All	0.9676	0.9568	0.9681
A	0.9956	0.9995	0.9945
B	0.9995	0.9989	0.9971
C	0.9642	0.9284	0.9706
D	0.8352	0.6327	0.8173
E	0.8203	0.9273	0.8111

| | Cell sample for cross section | | |
Country grade	All cells	High density cells	Low density cells
All	0.9633	0.9266	0.9813
A	0.9850	0.9772	0.9916
B	0.9663	0.9514	0.9880
C	0.9702	0.8929	0.9839
D	0.9615	0.8609	0.9827
E	0.6948	0.0638	0.6671

Table 7. Estimated optimal weight on measured output for grid cells of different grades using 17-year growth rates and for cross-sectional density

Figure 1. Nighttime lights from India

(Source: Defense Meteorological Satellite Program.)
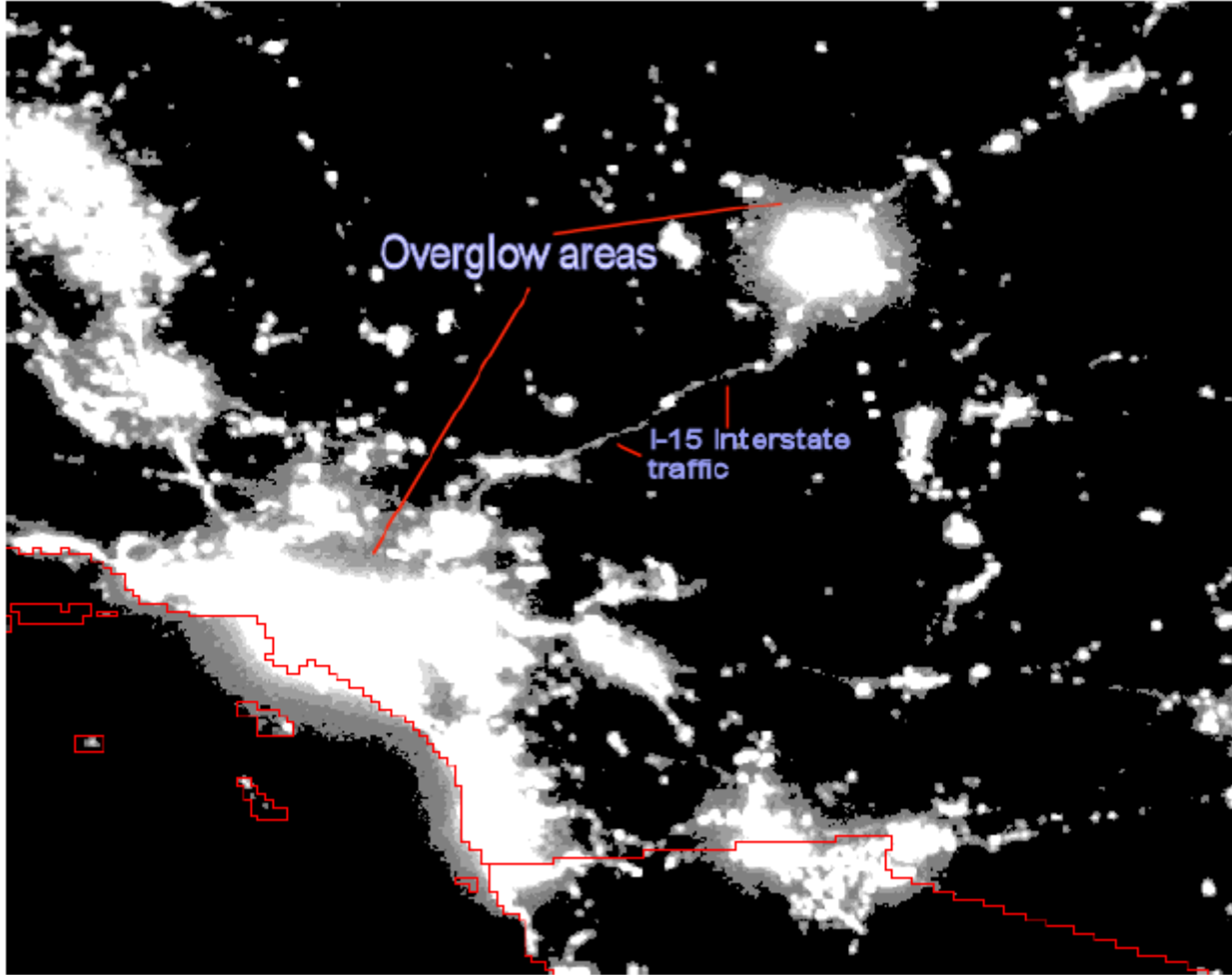
Figure 2. A detail from nighttime lights for Los Angeles for 2002

Note the lights from rural interstate I-15, which would not be measured by standard economic accounts. Note also the bleeding of lights beyond the coastline, which is shown in red. The picture covers an area of approximately 250 x 250 km.
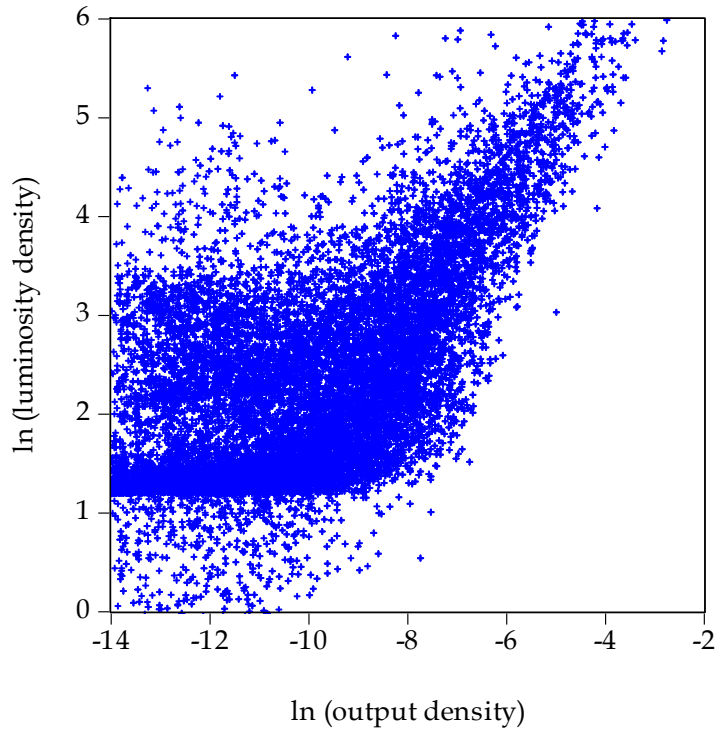
Figure 3. GDP from GEcon and luminosity data, all grid cells

This figure shows all grid cells with RIG > 0.1 and ln output density > -14.
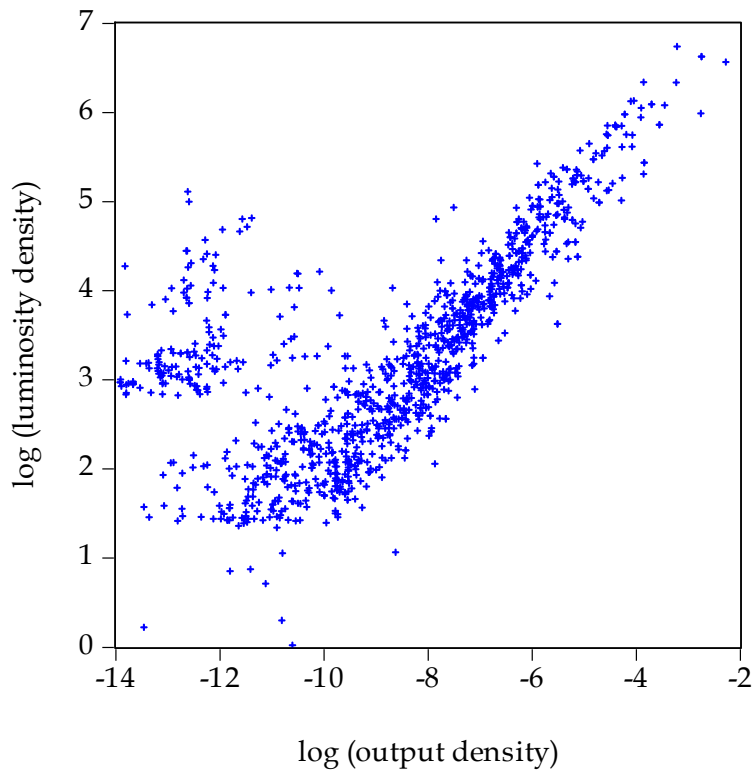
Figure 4. GDP from GEcon and luminosity data, US grid cells

This shows all U.S. grid cells. The cells at the upper left are Alaska.
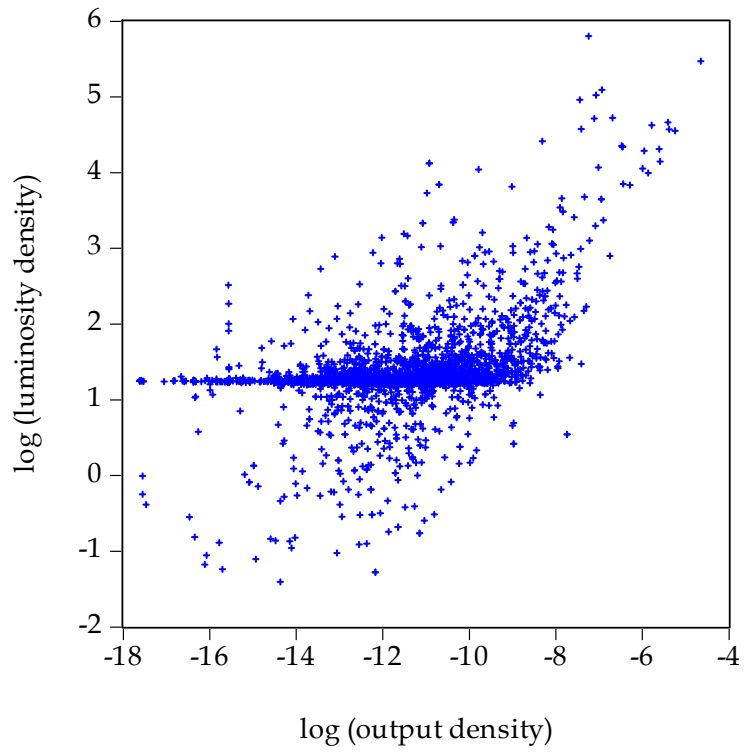
Figure 5. GDP from GEcon and luminosity data, African grid cells
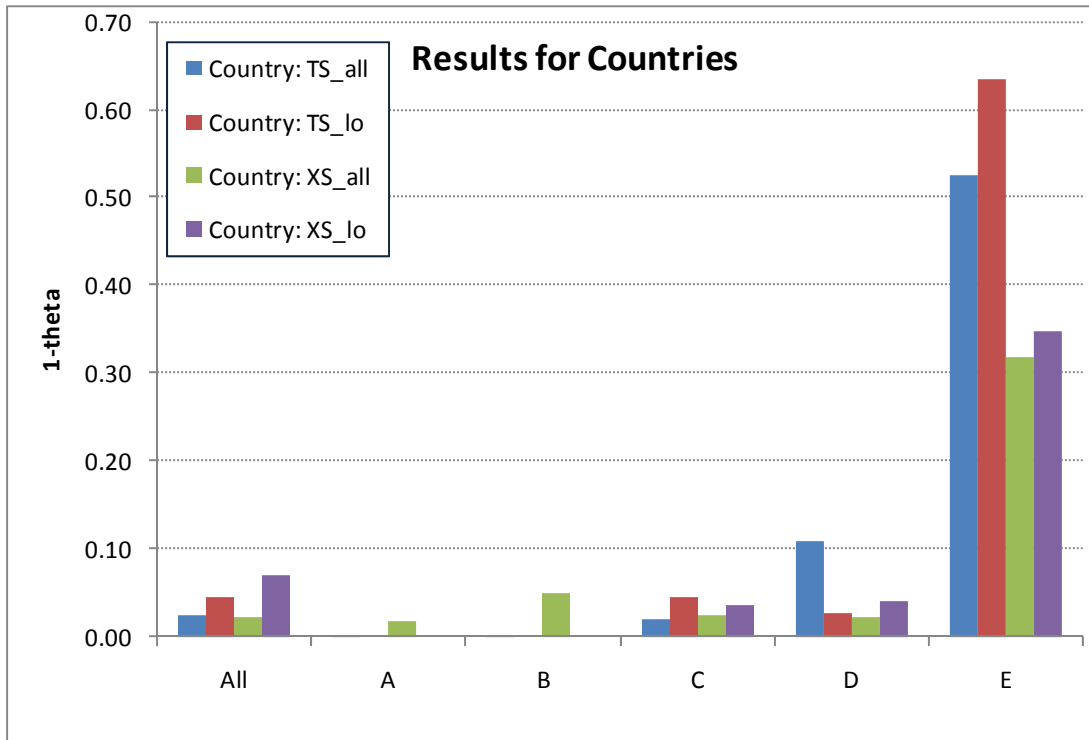
Figure 6. Summary estimates of the value of (1-θ) or the weight on luminosity for countries of different grades
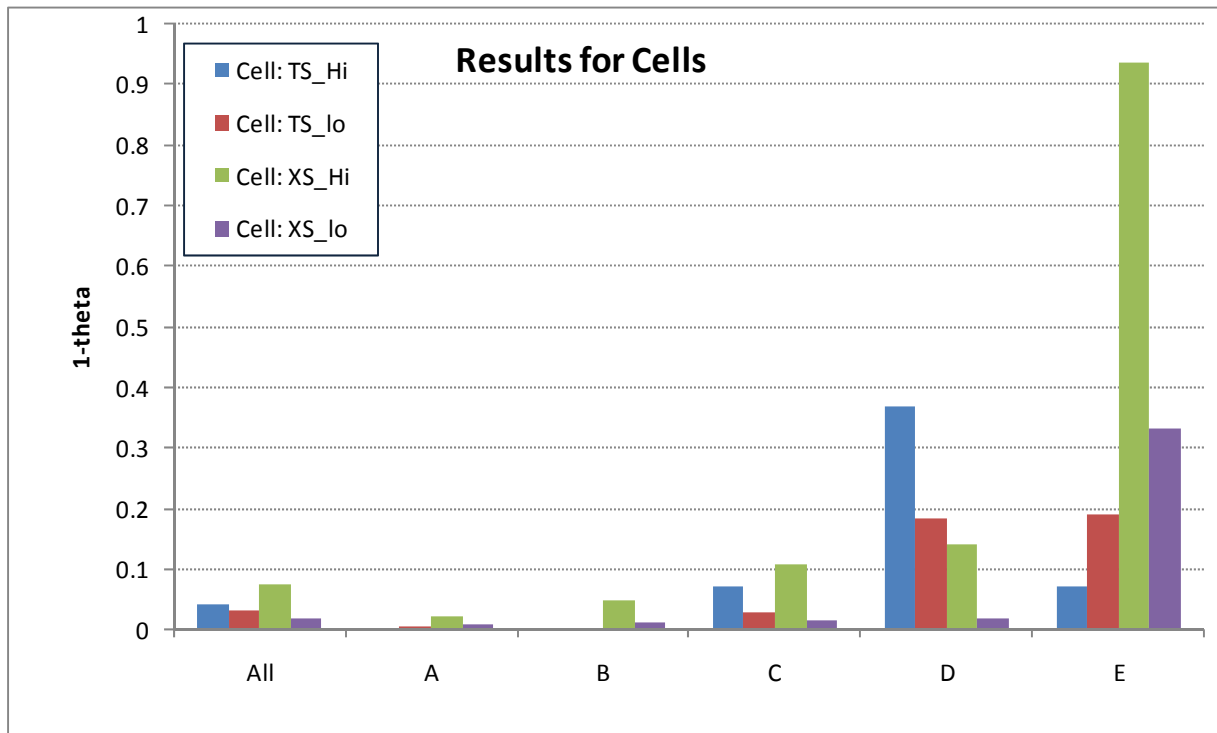
Figure 7. Summary estimates of the value of (1-θ) or the weight on luminosity for cells of different grades

References

Deaton, Angus S. and Heston, Alan. 2008. "Understanding PPPs and PPP-Based National Accounts." NBER Working Paper Series, Vol. w14499, 2008.

Doll, Christopher. 2008. *CIESIN Thematic Guide to Night-time Light Remote Sensing and its Applications*. Palisades, NY: Center for International Earth Science Information Network of Columbia University.

Doll, Christopher, Muller, J-P., and Christopher D. Elvidge. 2000. "Nighttime Imagery as a Tool for Global Mapping of Socio-economic Parameters and Greenhouse Gas Emissions." *Ambio*, 29, 157-162.

Ebener, Steve, Christopher Murray, Ajay Tandon, and Christopher Elvidge. 2005. "From Wealth to Health: Modeling the Distribution of Income per capita at The Sub-national Level Using Nighttime Lights Imagery." *International Journal of Health Geographics*, 4(5): 5-14.

Elvidge, Christopher D., Jeffrey Safran, Benjamin Tuttle, Paul Sutton and Pierantonio Cinzano, et al. 2007. "Potential for Global Mapping of Development via a Nightsat Mission." *GeoJournal*, 69: 45–53.

Elvidge, Christopher D., Kimberly E. Baugh, John B. Dietz, Theodore Bland, Paul C. Sutton and Herbert W. Kroehl. 1999. "Radiance Calibration of DMSP-OLS Low-Light Imaging Data of Human Settlements." *Remote Sensing of Environment*, 68(1):77-88.

Elvidge, Christopher D., K.E. Baugh, E.A. Kihn, H.W. Kroehl, E.R, Davis, and C.W. Davis. 1997. "Relation between Satellite Observed Visible - Near Infrared Emissions, Population, Economic Activity and Electric Power Consumption." *International Journal of Remote Sensing*, 18:1373-1379.

Fixler, Dennis J. and Bruce T. Grimm. 2008 "The Reliability of the GDP and GDI Estimates," *Survey of Current Business*, 88(2): 16–32.

Johnson, Simon H., William D. Larson, Chris Papageorgiou, and Arvind Subramanian. 2009. "Is Newer Better? Penn World Table Revisions and Their Impact on Growth Estimates" NBER Working Paper, No. w15455.

Henderson, Vernon, Adam Storeygard and David Weil. 2009. Measuring Economic Growth from Outer Space, NBER Working Paper Series, Vol. w15199, 2009.

Kuznets, Simon. 1934. *National Income, 1929-1932,* NBER: New York.

Sutton, P. C., C. D. Elvidge and T. Ghosh. 2007. "Estimation of Gross Domestic Product at Sub-national Scales Using Nighttime Satellite Imagery." *International Journal of Ecological Economics and Statistics*, 8 (SO7): 5 – 21.

Sutton, Paul C., Robert Costanza. 2002. "Global Estimates of Market and Non-market Values Derived from Nighttime Satellite Imagery, Land cover, and Ecosystem Service Valuation." *Ecological Economics*, V41(3): 509-527.

System of National Accounts. 1993. United Nations.
http://unstats.un.org/unsd/sna1993/toctop.asp (accessed March 21, 2010)

Grade level	Number of countries	Number of cells
A	16	3731
B	13	980
C	103	8736
D	26	1471
E	9	1063
Total	167	15981

Table A-1. Distribution of countries and cells without missing values by grade

Country	Grade	Country	Grade	Country	Grade
Afghanistan	E	Central African	D	Gibraltar	B
Albania	C	Chad	E	Greece	B
Algeria	D	Chile	B	Greenland	B
American Samoa	C	China	C	Grenada	C
Andorra	C	Christmas I.	D	Guadeloupe	C
Angola	E	Cocos Is.	D	Guam	C
Anguilla	C	Colombia	C	Guatemala	C
Antarctica	D	Comoros	D	Guinea	D
Antigua and Barb	C	Congo	E	Guinea Bissau	C
Argentina	B	Cook Is.	D	Guyana	D
Armenia	C	Costa Rica	C	Haiti	D
Aruba	C	Cote d'Ivoire	C	Heard I. & McDon	D
Australia	A	Croatia	C	Honduras	C
Austria	A	Cuba	D	Hong Kong	C
Azerbaijan	C	Cyprus	D	Hungary	C
Bahamas	C	Czech Republic	C	Iceland	C
Bahrain	C	Democratic Repub	D	India	C
Baker and Howlan	D	Denmark	A	Indonesia	C
Bangladesh	C	Djibouti	E	Iran	C
Barbados	C	Dominica	C	Iraq	E
Belarus	D	Dominican Republ	C	Ireland	A
Belgium	A	Ecuador	C	Isle of Man	C
Belize	C	Egypt	C	Israel	B
Benin	C	El Salvador	C	Italy	A
Bermuda	C	Equatorial Guine	C	Jamaica	C
Bhutan	D	Eritrea	E	Jan Mayen	D
Bolivia	C	Estonia	C	Japan	A
Bosnia&Herzegovi	C	Ethiopia	C	Jarvis I.	D
Botswana	C	Falkland Is.	C	Johnston Atoll	D
Bouvet I.	D	Faroe Is.	C	Jordan	C
Brazil	C	Federated State	D	Kazakhstan	C
British Indian O	D	Fiji	C	Kenya	C
British Virgin I	C	Finland	B	Kiribati	D
Brunei	C	France	A	Kuwait	C
Bulgaria	C	French Guiana	C	Kyrgyztan	C
Burkina Faso	C	French Polynesia	C	Laos	D
Burundi	C	French Southern	D	Latvia	C
Cambodia	D	Gabon	C	Lebanon	C
Cameroon	C	Gambia	C	Lesotho	D
Canada	A	Georgia	C	Liberia	D
Cape Verde	C	Germany	B	Libya	D
Cayman Is.		Ghana	C	Liechtenstein	B

Country	Grade	Country	Grade	Country	Grade
Liechtenstein	B			St. Pierre & Miq	C
Lithuania	C	Norway	A	St. Vincent and	C
Luxembourg	A	Oman	C	Sudan	E
Macau	C	Pakistan	C	Suriname	D
Macedonia	C	Palau	D	Svalbard	C
Madagascar	C	Panama	C	Swaziland	C
Malawi	C	Papua New Guinea	D	Sweden	A
Malaysia	C	Paraguay	C	Switzerland	A
Maldives	C	Peru	C	Syria	C
Mali	C	Philippines	C	Taiwan	C
Malta	D	Pitcairn Is.	D	Tajikistan	D
Marshall Is.	D	Poland	B	Tanzania	C
Marshall Islands	D	Portugal	B	Thailand	C
Martinique	C	Puerto Rico	B	Timor Leste	E
Mauritania	C	Qatar	C	Togo	D
Mauritius	C	Reunion	C	Tokelau	D
Mayotte	D	Romania	C	Tonga	D
Mexico	C	Russia	C	Trinidad and Tob	C
Micronesia	C	Rwanda	C	Tunisia	C
Midway Is.	C	Saint Helena	C	Turkey	C
Moldova	C	Samoa	C	Turkmenistan	D
Monaco	B	San Marino	C	Turks & Caicos I	C
Mongolia	D	Sao Tome and Pri	C	Tuvalu	D
Montserrat	D	Saudi Arabia	D	Uganda	E
Morocco	C	Senegal	C	Ukraine	C
Mozambique	D	Serbia and Monte	C	United Arab Emir	C
Myanmar	E	Seychelles	C	United Kingdom	A
Namibia	D	Sierra Leone	C	United States	A
Nauru	D	Singapore	B	Uruguay	B
Nepal	C	Slovakia	C	Uzbekistan	D
Netherland Antil	C	Slovenia	C	Vanuatu	D
Netherlands	A	Solomon Islands	C	Vatican City	D
Netherlands Anti	C	Somalia	E	Venezuela	C
New Caledonia	C	South Africa	C	Vietnam	C
New Zealand	B	South Georgia &	E	Virgin Is.	C
Nicaragua	C	South Korea	B	Wake I.	C
Niger	E	Spain	B	Wallis and Futun	D
Nigeria	C	Sri Lanka	C	West Bank and Ga	E
Niue	D	St. Helena	C	Yemen	E
Norfolk I.	D	St. Kitts and Ne	C	Zambia	C
North Korea	E	St. Lucia	C	Zimbabwe	D
Northern Mariana	D				

Table A-2. Country grades

year	f10	f12	f14	f15	f16
1992	130901.1	.	.	.	.
	215777.5	.	.	.	.
	17575.0	.	.	.	.
1993	116858.5	.	.	.	.
	204683.3	.	.	.	.
	17571.0	.	.	.	.
1994	111849.4	139993.3	.	.	.
	214893.1	224159.5	.	.	.
	17326.0	17489.0	.	.	.
1995	.	130891.2	.	.	.
	.	227916.0	.	.	.
	.	17539.0	.	.	.
1996	.	123192.8	.	.	.
	.	219661.5	.	.	.
	.	17575.0	.	.	.
1997	.	135643.5	96053.0	.	.
	.	234274.8	192918.7	.	.
	.	17575.0	17552.0	.	.
1998	.	191499.8	105024.5	.	.
	.	253795.5	197161.8	.	.
	.	19355.0	17565.0	.	.
1999	.	177086.1	103215.6	.	.
	.	263062.1	202037.6	.	.
	.	18378.0	17549.0	.	.
2000	.	.	109679.3	181436.7	.
	.	.	216877.1	240954.9	.
	.	.	17469.0	19823.0	.
2001	.	.	125058.7	145177.2	.
	.	.	227136.4	232601.6	.
	.	.	17550.0	17575.0	.
2002	.	.	124119.7	152794.8	.
	.	.	236180.5	239186.9	.
	.	.	16365.0	17575.0	.
2003	.	.	128392.2	99541.4	.
	.	.	222748.3	191665.7	.
	.	.	19138.0	17575.0	.
2004	.	.	.	110133.6	140468.3
	.	.	.	192524.2	213234.6
	.	.	.	19826.0	19826.0
2005	.	.	.	104591.3	110601.9
	.	.	.	199367.8	195375.6
	.	.	.	19826.0	19826.0
2006	.	.	.	102925.4	133114.0
	.	.	.	198590.5	212744.5
	.	.	.	19826.0	19826.0
2007	.	.	.	103532.5	153832.3
	.	.	.	198575.4	238442.9
	.	.	.	19826.0	19826.0
2008	.	.	.	.	137462.6
	.	.	.	.	234186.5
	.	.	.	.	19826.0

Table A-3. Statistics of luminosity raw data (mean, standard deviation, and number of observations)

| Correlation Coefficent	f10	f12	f14	f15	f16
1992	0.519				
1993	0.521				
1994	0.546	0.534			
1995		0.561			
1996		0.542			
1997		0.537	0.561		
1998		0.369	0.529		
1999		0.486	0.545		
2000			0.549	0.356	
2001			0.550	0.515	
2002			0.635	0.496	
2003			0.466	0.502	
2004				0.350	0.403
2005				0.420	0.428
2006				0.416	0.417
2007				0.424	0.410
2008					0.469

Table A-4. Correlation of log GDP and log luminosity for grid cells by year

| errors of regressions	f10 (t+1 on t)	f12 on f10	f12 (t+1 on t)	f14 on f12	f14 (t+1 on t)	f15 on f14	f15 (t+1 on t)	f16 on f15	f16 (t+1 on t)
1992	0.205								
1993	0.260								
1994		0.369	0.273						
1995			0.150						
1996			0.172						
1997			0.156	0.211	0.249				
1998			0.157	0.188	0.185				
1999				0.182	0.267				
2000					0.338	0.288	0.158		
2001					0.320	0.255	0.122		
2002					0.315	0.357	0.144		
2003						0.254	0.203		
2004							0.158	0.176	0.169
2005							0.138	0.136	0.168
2006							0.160	0.162	0.136
2007								0.175	0.127

Table A-5. Errors of regressions of log grid cell luminosity over satellites or years

The entries are the standard errors of estimates of regressions of the logarithm of luminosity for grid cells across satellites or across time. The unshaded columns show the SEE for regressions of luminosity in t and t+1, while the shaded columns show the SEE for luminosity of satellite k on satellite j.
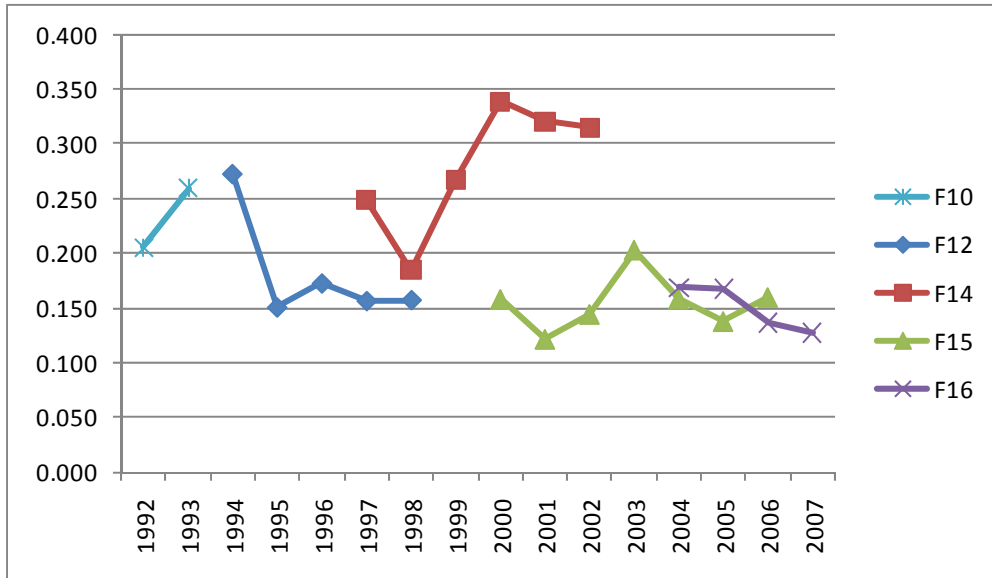
Figure A-1. Time-series error of log grid cell luminosity

This figure shows the standard error of estimate or SEE (described in Table A-5) for individual satellites and adjacent time periods. For example, the first line shows the SEE for two pairs of years for satellite F10. These data correspond to the second column of Table A-5.
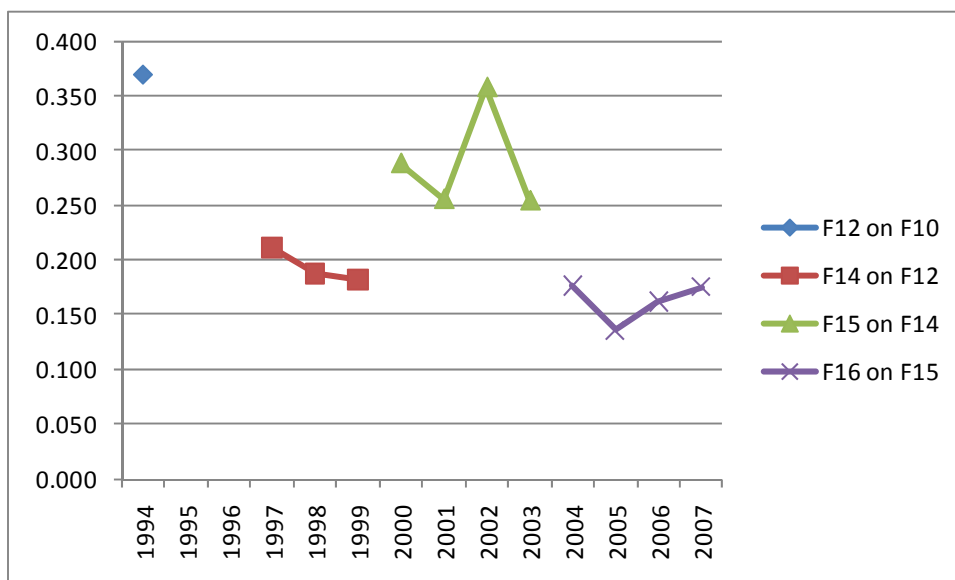
Figure A-2. Cross sectional errors of log grid cell luminosity

This figure shows the standard error of estimate or SEE (described in Table A-5) for the same year and different satellites. For example, the second line with three observations shows the SEE comparing three pairs of years for satellite F12 and satellite F14. These data correspond to the fifth column of Table A-5.

IV method	Cell sample for time series		
Country grade	All cells	High density cells	Low density cells
All	0.9164	1.0000	0.9161
A	0.9935	0.9479	0.9931
B	0.9945	1.0000	0.9908
C	0.9347	0.9791	0.9418
D	0.9693	0.9864	0.9361
E	0.8942	0.9961	0.8944

IV method	Cell sample for cross section		
Country grade	All cells	High density cells	Low density cells
All	0.9640	0.9273	0.9810
A	0.9802	0.9604	0.9811
B	0.9637	0.9411	0.9826
C	0.9707	0.9022	0.9829
D	0.9730	0.8772	0.9929
E	0.8641	0.4064	0.9348

Table A-6. IV regressions of luminosity equation for all grid cells