Measuring Economic Growth from Outer Space

Vernon Henderson
Brown University and NBER

Adam Storeygard
Brown University

David N. Weil*
Brown University and NBER

July 21, 2009

Abstract

GDP growth is often measured poorly for countries and rarely measured at all for cities. We propose a readily available proxy: satellite data on lights at night. Our statistical framework uses light growth to supplement existing income growth measures. The framework is applied to countries with the lowest quality income data, resulting in estimates of growth that differ substantially from established estimates. We then consider a longstanding debate: do increases in local agricultural productivity increase city incomes? For African cities, we find that exogenous agricultural productivity shocks (high rainfall years) have substantial effects on local urban economic activity.

Keywords: economic growth, remote sensing, urbanization, income measurement
JEL Codes: E01, O47, Q1, R11

* Vernon_Henderson@Brown.edu, Adam_Storeygard@Brown.edu, David_Weil@Brown.edu. We thank Chris Elvidge for advice and auxiliary data; Andrew Foster, Stefan Hoderlein, Blaise Melly, Daniel Orenstein, and seminar participants at Brown, the 2008 BREAD/CEPR/Verona Summer School on Development Economics, and the 2009 Annual Meeting of the Population Association of America for comments; and Joshua Wilde and Isabel Tecu for research assistance. Storeygard acknowledges support from Award Number T32HD007338 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development.

Introduction

Gross Domestic Product (GDP) is the most important variable in analyses of economic growth. The conceptual problems in defining GDP, let alone using it as a measure of welfare, are the stuff of introductory economics courses. Just as serious, however, is the problem that GDP itself is often badly measured, especially in developing countries. Relative to developed countries, in a typical developing country a much smaller fraction of economic activity is conducted within the formal sector, the degree of economic integration and price equalization across regions is lower, and the government statistical infrastructure is much weaker. These factors make the calculation of nominal GDP (total value added, in domestic prices) difficult. Making useful comparisons of real GDP, either over time or between countries, also requires the construction of price indices: either a domestic price index to measure real income growth within a country, or purchasing power parity exchange (PPP) rates based on prices for a comparable set of goods to make inter-country comparisons.

In the Penn World Tables (PWT), one of the standard compilations of cross-country data on income, countries are given grades corresponding to subjective data quality, with a grade of A indicating a margin of error of 10%, B indicating 20%, C indicating 30%, and D indicating 40%. The grading is based in part on the ability to construct good PPP measures, but also reflects a country's capacity to produce reliable national income accounts and domestic price indices. Almost all industrialized countries receive a grade of A. By contrast, for the 43 countries of sub-Saharan Africa, 17 get a D and 26 get a C. (Deaton and Heston 2008)

An illustration of the degree of measurement error in the PWT comes from Johnson et al.'s (2009) study of revisions to the PWT data. Specifically, the authors compared version 6.1 of the PWT, released in 2002, with version 6.2, released in 2006. The standard deviation of the change in countries' average growth over the period 1970-1999 was 1.1% per year – an enormous change in comparison to the average growth rate over this period of 1.56% per year. To give a particularly striking example: the authors calculated the ten worst growth performers in Africa based on the 6.1 data and similarly based on the 6.2 data. Only five countries were on both lists.[1] Measurement error in GDP data can easily lead researchers to erroneous conclusions.

[1] Changes in data between different versions of the PWT can result from changes in the pricing survey used to establish purchasing power parities (known at the International Comparisons Project or ICP) as well as revisions in underlying national income accounts data and changes in

For example, Dawson *et al.* (2001) claim that the empirical link between output volatility and income growth in the PWT data is purely a product of measurement error in annual income.

In the worst case, some countries simply have no national accounts data available at all. For example, Iraq, Myanmar, Somalia, and Liberia are among the countries not included in the most recent (6.2) version of PWT. Finally for most developing countries and many developed ones, reliable data on output at the sub-national level, particularly cities but even larger regions, is not regularly available.

In response to the problems of measuring GDP, there is a long tradition in economics of considering various proxies that cover periods or regions for which GDP data are not available at all or not available in a timely fashion. For example, until the year 2005, the Federal Reserve Board based its monthly index of industrial production in part on a survey of utilities that measured electricity delivered to different classes of industrial customers. Similarly, an IMF study examining electricity consumption in Jamaica over the decade of the 1990s concluded that officially measured GDP growth, which averaged 0.3% per year, understated true output growth by 2.7% per year, the gap being explained by growth of the informal sector (IMF, 2006). Economic historians have also employed a variety of proxies for studying economic outcomes in the period before the creation of national income accounts and in order to examine growth in sub-national units. For example, Good (1994) estimates output in 22 sub-regions of the Habsburg Empire in the period 1870-1910 using proxies such as the number of letters mailed per capita. The essays in Steckel and Rose (2002) use skeletal remains to measure both the average standard of living and the degree of inequality in the Americas over the last two millennia.

In this paper we explore the usefulness of a different proxy for economic activity: the amount of light that can be observed from outer space. More particularly, our focus will be on using changes in "night lights" as a measure of economic growth. There are two reasons to do so. First we can use the change in night light intensity as an additional measure of income growth. Even if changes in light from space are subject measurement error, it is well known that several error-prone measures are better than one, especially if there is no reason to think that the measurement errors are correlated (e.g., Browning and Crosley, 2009). In the paper, we develop

---

methodology. Versions of the PWT within the same "generation," for examples versions 6.1 and 6.2, use the same ICP data. Johnson *et al.* report that changes in national income accounts data are the dominant source of differences between the two versions.

a simple framework showing how to combine our lights measure, which is in a different metric than income (c.f., Browning and Crosley, 2009 or Krueger and Lindahl, 2001), with income measures to improve estimates of true economic growth. We illustrate the methodology with an application to countries that are perceived as having low capacity in generating reliable national income accounts and price indices, those that receive a grade D in the PWT. For these countries we provide new estimates of their economic growth over the period 1992/3 to 2002/3.

Second, there are many circumstances where we have changes in night lights data that inform us about economic growth, but no corresponding measures of income growth. Most significantly, night lights data are available at a far greater degree of geographic fineness than is attainable in any standard income and product accounts. As discussed later, we can map data on light observed from space on approximately one-kilometer squares and aggregate them to the city or regional level. This makes the data uniquely suited to spatial analyses of economic activity. Economic analysis of growth and of the impacts of policies and events on cities and regions of many countries is hindered by a complete absence of any regular measure of local economic activity. While population data are sometimes regularly available for cities above a certain size, almost no countries have city level GDP data.[2] Night lights data give us such a measure. Note also that data from satellites are available at a much higher time frequency than standard output measures. Thus they are available well in advance of income measures from national accounts and provide an early signal of country growth changes. Also, as will be illustrated below, they allow us to assess the time patterns on regional income growth of events such as discovery of minerals, construction of roads, civil strife, and the like.

To illustrate the application of night lights to measuring economic growth at sub-national levels and at the same time contribute to a long-standing debate in economics, we examine the extent to which productivity in the agricultural hinterland of a city affects city income. Urban economists tend to model cities as either divorced from their hinterland (e.g., Black and Henderson, 1999) or as source of demand for local agricultural crops (von Thunen, 1826 and Nerlove and Sadka, 1991). Traditional development economics views the rural sector as simply a source of surplus labor (dual sector models following Lewis 1954 and Harris-Todaro 1970). The new economic geography allows agriculture to be a source of demand for urban products, but the interaction plays a limited role in analysis (Krugman, 1991, with some empirical application in

[2] For an exception, see Au and Henderson (2006) on China.

de Mata et al., 2007). Only a handful of agricultural growth economists (e.g, Irz and Roe, 2005 and Tiffin and Irz, 2006) seem to seriously consider that productivity gains in local agriculture play a strong role in stimulating city economic activity. The idea that agricultural activity spurs urban economies is hard to test because it require sub-national data on both city incomes and incomes in the agricultural hinterland of cities, as well as a context to make inferences about which way causality runs. In this paper, we make use of the natural experiment of rainfall shocks to examine the extent to which productivity gains in local agriculture engender increases in economic activity as measured by night lights, for 541 African cities served by local agricultural hinterlands.

The rest of this paper is organized as follows. Section 1 gives a brief introduction to the night lights data, discusses more obvious examples of how they represent differences in income levels or growth across countries and the effects of shocks on growth or income levels, and estimates simple baseline specifications where changes in lights over time may be used to predict income growth. In Section 2 we develop the statistical framework to show how information on changes in lights can be combined with existing measures of income growth to get improved estimates of true income growth. In Section 3 we turn to the application where we estimate the impact of agricultural productivity shocks on urban economic activity for a large sample of African cities. Section 4 concludes.

1. Night lights data

Several US Air Force weather satellites circle the earth 14 times per day, recording the intensity of earth-based lights. Each satellite observes every location on the planet (between 65 degrees S latitude and 65 degrees N latitude) every night at some time between 8:30 and 10:00 pm. Using night lights during the dark half of the lunar cycle in seasons when the sun sets early removes intense sources of natural light, leaving mostly man-made light. Readings affected by auroral activity (the northern and southern lights) and forest fires are also removed both manually and using frequency filters. Observations where cloud cover obscures visible light are also excluded. Intensity of lights is measured as a six-bit (0-63) digital number calculated for every 30-second output pixel (approximately 0.86 square kilometers at the equator), which is averaged across overlapping raw input pixels and all valid evenings in a year. The values are not direct measures of physical luminance, because sensor settings vary over time. However, they can be relatively calibrated over time to get a reasonable approximation of trends in luminance, in part because of several years in which multiple sensors on different satellites were collecting data at the same time. The recalibrated data, which we use throughout the paper, is on a scale of 0-65. Because pixel size varies by latitude,[3] below in statistical analysis for each relevant region (e.g., a country), we calculate a weighted average of lights across pixels within a country, with the weights being a pixels' shares of a region's land area.

Intensity of night lights reflects outdoor and some indoor use of lights. However, more generally, consumption of nearly all goods in the evening requires lights. As income rises, so does light usage per person, in both consumption activities and many investment activities. Obviously this a complex relationship, and we abstract from such issues as public versus private lighting, relative contributions of consumption versus investment, and the relationship between daytime and nighttime consumption and investment. Because we will be looking at *growth* in lights in statistical work, cross-country level differences in these ratios will not be important. Growth in lights is just another proxy measure for true growth in income, where the advantage of lights data over other proxies is that they are readily available.

Table 1 gives some sense of the data, describing the distribution of digital numbers across pixels for ten countries covering a broad range of incomes and population densities. For reference, we also include data on GDP per capita, population density, and the fraction of the population living in urban areas. One measure of interest is the fraction of pixels for which no light at all is registered. In the United States, 67.7% of pixels are unlit. In Canada that percentage is over 90, while in the Netherlands it is under 1. The percentage of unlit pixels falls with income holding density constant; Bangladesh, with higher population density than the Netherlands, has 68% of pixels unlit. Among poor, sparsely populated countries like Mozambique and Madagascar, over 99% of pixels are unlit.

Among the countries in Table 1 (and more generally in the sample) there are remarkably few pixels with digital numbers of 1 or 2. Among middle and lower income countries, the most commonly observed range for the digital number is from 3-5; for the US and Canada, it is 6-10; and for the Netherlands, it is 21-62. The minimal fraction of pixels with digital numbers of 1 or

[3] Data for lights (and rainfall) are reported on a latitude-longitude grid. Because of the curvature of the Earth, grid cell size varies in proportion to the cosine of latitude. Thus all grid cell sizes are reported at the equator; sizes at other latitudes can be calculated accordingly. For example a grid cell in London, at 51.5 degrees latitude, is 0.62 square kilometers.

2 reflects, we think, the effect of software designed to filter out noise in the sensor. More generally the censoring of data at the low end means some low density-low income pixels do not get counted, so to some extent we will undercount lights nationally. Pixels with values of 63-65 are mostly top-coded.[4] This affects small, densely-populated areas of rich countries and almost nowhere in poor countries.

Table 1 also shows the mean digital number and the within-country Gini for the digital number. The mean ranges from 22 in the Netherlands to 0.03 in Madagascar. The Gini varies enormously across countries as well. Below in the empirical work we will explore whether dispersion measures like the Gini additionally contribute to our ability to predict income growth.

1.1 Simple examples of what night lights data reflect

A global view

A quick look at the world in Figure 1 suggests that lights reflect human economic activity as pointed out in Croft (1978), Elvidge *et al.* (1997), Sutton and Costanza (2002), Ebener et al (2005), Doll *et al.* (2006) and Sutton et al (2007), among others.[5] In the figure unlit areas are black, and lights appear with intensity increasing from gray to white. Lights in an area reflect total intensity of income, which is increasing in both income per person and number of people. In the United States, where living standards are fairly uniform nationally, the higher concentration of lights in coastal areas near the oceans and the Great Lakes reflects the higher population densities there. The comparison of lights in Western Europe and India reflects huge differences in per capita income, as does the comparison between Brazil and the Democratic Republic of Congo.

Eastern Europe and the Former Soviet Republics Over Time

---

[4] Because of relative calibration across years, the top-coded value ranges from 63-65; in the raw data, it is always 63. The distribution of the data is such that it is much rarer to find pixels with a value of 63 in a satellite-year in which the top-coded value is 64 or 65.

[5] Indeed, several of these authors estimated the cross-sectional lights-GDP relationship for countries and subnational units of developed regions. However, to our knowledge only Ebener *et al* (2005) and Sutton *et al* (2007) have considered sub-national units of developing countries, both with very small numbers of units per country. Sutton *et al* (2007) is the only paper with quantitative analysis of data for multiple (two) years, but they do not produce panel estimates.

To see mostly pure income effects, we examine the differential effects of the economic transition on income and lights in Eastern Europe versus the neighboring former Soviet republics. Specifically, we compare the former Soviet republics of Moldova and Ukraine, where per capita income fell in the wake of the USSR's breakup, with their neighbors Hungary, Poland, and Romania, which went through a much smoother transition process.

Although our satellite data only start two years into transition the differences in lights are dramatic (Elvidge *et al.*, 2005). In Figure 2 the more brightly lit areas in 2002 are in the Eastern European countries, where light intensity increases dramatically from 1992 to 2002. The dimming of lights over the same 10 years for their neighbors who were formerly part of the Soviet Union is distinct. In Moldova and Ukraine, income per capita fell by 30% and 35% respectively, while population fell by 3% and 8% respectively, and light intensity dropped by 68% and 47% respectively. In Hungary, Poland and Romania, where incomes rose by 41%, 56%, and 23%, the respective rises in lights were 46%, 80%, and 112%.

Gemstones in Madagascar

As mentioned above, a strength of night light data is that they can be used to examine changes in economic activity at a very local scale. In late 1998, large deposits of rubies and sapphires were accidentally discovered in southern Madagascar, near the towns of Ilakaka and Sakaraha. The region is now thought to contain the world's largest sapphire deposit, accounting for around 50% of world supply, and Ilakaka and Sakaraha have become major trading centers for sapphires. Previously little more than a truck stop, Ilakaka's population is now estimated at roughly 20,000.[6] The story of these developments can clearly be seen in the night lights data in Figure 3. In 1998 (and all but one of the previous six years) there were no lights visible in either Ilakaka or Sakaraha. Over the next five years there was a sharp growth in the number of pixels for which light is visible at all, and in the intensity of light per pixel. The other town visible in the figure, Ihosy, shows no such growth. If anything, Ihosy's light gets smaller and weaker, as it suffers in the competition across local cities for population.

[6] Hamilton, Richard BBC News Online, "Madagascar's Scramble for Sapphires," 1 August 2003, http://news.bbc.co.uk/2/hi/africa/3114213.stm Accessed 18 January 2008
Hogg, Jonny. BBC News Online, "Madagascar's Sapphire Rush," 17 November 2007, http://news.bbc.co.uk/2/hi/programmes/from_our_own_correspondent/7098213.stm Accessed 18 January 2008

1. 2. Lights as a measure of economic activity

In this sub-section, we analyze the use of lights as a measure of growth in national economic activity. If Y_j is true real income and $\tilde{X}_j$ is total lights summed across all pixels country j with area A_j, as a level relationship we expect:

$$\ln(\tilde{X}_j / A_j) = f(\ln(Y_j / A_j)), \qquad f' > 0 \tag{1}$$

As a "structural" relationship, increased income generates increased light usage, so lights in an area are an increasing function of total income in the area. As written and in this paper, we assume the latter is increasing at the same rate in number of people and per capita income. It is not clear what the curvature of the $f(\cdot)$ function should be, although we will generally assume log-linearity. There could be some diminution in the rate of increase of light as income rises: with more urbanization there is a greater likelihood of people living above one another, so that some light is blocked from reaching space. Also, with urbanization, there could be economies of scale in the use of lights, such as street lamps. On the other hand, there are large fixed costs associated with electricity distribution, which could lead to a convex relationship between income and light output around some income threshold. Of course the shape of the relationship will also be affected by the nature of the sensors used. The functional relationship between true luminance and recorded digital numbers is unknown.

Our data's capacity to measure true luminance varies across countries by climate and is affected by changes in light sensor technology and specific satellites over time. Also the composition of income between consumption and investment, the division of economic activity between night and day, population density, and land area vary across counties. To mitigate all these problems, we restrict attention to growth formulations, where these variations across counties are differenced out, and we include year fixed effects in our analysis. Also, we are not so interested in the structural relationship in (1) *per se*. Rather, our goal is to predict income growth using light growth data. Similarly, by focusing on income growth, we can reduce error by avoiding PPP measures of income (Nuxoll, 1994). Instead, we look at the income growth rate in a country in constant local currency units (LCU), which tells us the real internal growth for the bundle of goods relevant to the country in question.

For purposes of predicting income growth using data on changes in lights, we difference a log-linear version of (1) and rearrange to estimate an equation of the form

$$y_{1jt} = \psi x_{jt} + e_{jt}, \qquad (2)$$

where $y_{1jt} \equiv \ln Y_{1jt} - \ln Y_{1jt-1}$ and $x_{jt} \equiv \ln X_{jt} - \ln X_{jt-1}$. X_{jt} is the weighted average of lights across pixels in a country. The parameter ψ is the inverse of the elasticity of lights with respect to income. We experimented with different functional forms and controls for changes in light dispersion. Those experiments, some of which we report below, suggest (2) is appropriate.

We estimate (2) for a panel of countries, in two ways. First, we look at annual data for 1992-2003 on income and light, and estimate a levels specification with a full set of country fixed effects. We add time fixed effects to help control for differences in light calibration across different aging satellites in different years, as well as sweeping out worldwide income growth effects. Identification is from within-country relative variation in lights and income over time. Second, we estimate (2) directly, with a long differenced relationship between 1992/93 and 2002/03. In our application in the next section we rely on the long differenced model.

Our measure of GDP is in constant local currency units and taken from the World Development Indicators (WDI). The lights data are collected by US Air Force weather satellites. Data for the years 1992-2003 are processed and distributed by the National Oceanic and Atmospheric Administration's (NOAA) National Geophysical Data Center. In years with data for two satellites, simple averages across satellites are calculated for each pixel. Details are in the Appendix.

1.3 Basic Results

Table 2 presents some basic results for a slightly unbalanced panel of 187 countries over 12 years, where we drop Equatorial Guinea as an outlier (see below).[7] An average of 179 countries appear in each year. The smallest number in any year is 174. Column 1 shows the fixed effect results, where the within R^2 is very high at 0.66. Column 2 suggests a quadratic

[7] The panel is primarily unbalanced due to missing GDP data.

specification does not fit the data, while Figure 4, looking non-parametrically at the y_{jt}, x_{jt} relationship suggests a linear specification in the growth rates is appropriate.

The estimated coefficient on lights, ψ, is 0.29 and highly significant in column 1. As discussed below, this estimate is seriously biased and we are primarily interested in using it in prediction rather than using it to infer a structural parameter. However, if one ignores the bias and interprets the coefficient structurally, it implies that the elasticity of light with respect to income ($1/\psi$) is greater than one, which would be consistent with light being a luxury good over the relevant range in the data.

In column 3 we control for dispersion of lights within a country by using the Gini coefficient for lights among pixels within a country. Given that the estimated relationship between lights and income in column 1 is concave, one would expect that a greater dispersion of lights, holding the average level of light constant, would be associated with lower average income. However, in column 3, the coefficient on lights is the same as in column 1 and the Gini has a zero coefficient. We also tried interactions of the Gini with lights and a translog formulation of the two, but the results suggest the simple log-linear model fits the data better.[8]

In column 4 we estimate the relationship in long differences, averaging the first two and last two years of data.[9] The elasticity is somewhat higher, though within one standard error, of the estimate in column 1. Figure 5 plots the long difference data points for 171 countries adding back in Equatorial Guinea. The figure shows why the linear approximation in Table 2 does so well, and also illustrates why we dropped Equatorial Guinea as an extreme outlier.[10] We also estimated a long difference version adding in a quadratic term and then the change in the Gini; again both coefficients are zero.

2. Using night lights data to improve estimates of growth in true income

[8] To measure dispersion one could also use the standard deviation of lights within a country. However, even after factoring out country and year fixed effects the simple correlation between the standard deviation and mean of lights is 0.89. Note the Hirschman-Herfindahl index can be decomposed into a part related to the standard deviation and a part to do with number of pixels per country; the latter is already controlled for by country fixed effects.

[9] For the Bahamas, Barbados and Cambodia, income data for one of the four relevant years is missing. In these cases, we simply use the other three.

[10] The WDI data imply an annual growth rate of GDP in Equatorial Guinea of over 23%. During this period, oil production in the country went from almost nothing to the third highest level in Africa. The population is less than one million.

As mentioned above, the PWT gives grades to countries corresponding to degree of error in measurement of PPP GDP. While we are not using PPP numbers, we might expect that the quality gradient in the PWT would apply to our GDP data (built up from national income accounts and domestic price indices) as well. Variations in the degree of measurement error should show up as heteroskedasticity in our regressions, with the error variance depending on the quality group a country is in. To test whether this is the case, we calculated the mean squared errors for the four quality groups for the long difference equation (column 4 of Table 2). These are, in grade order (A through D): 0.022, 0.037, 0.024, and 0.041. The A and B groups are very small, however, with only 18 and 13 countries, respectively. Combining A and B into one group results in values of 0.028, 0.024, and 0.041 for A and B, C, and D groups, respectively. The mean squared error for D countries is much larger than for the other country groups.

This exercise suggests, not surprisingly, that it is in the D countries where an alternative to national income accounts data would be most valuable. We now proceed to show how such an alternative can be constructed using the lights data.

2.1 The statistical model

We have an unobserved magnitude, y_{jt}, the growth rate in true income in country *j*, for which we wish to obtain the best estimate possible. We have two measures that relate to y_{jt}: (1) the growth in measured income y_{1jt} and (2) the growth in lights x_{jt}. The relationships are

$$y_{1jt} = y_{jt} + \varepsilon_{1jt} \tag{3a}$$
$$x_{jt} = \beta y_{jt} + \varepsilon_{2jt} \tag{3b}$$

Note the units of the dependent variables in (3a) and (3b) differ and that (3b) is a specific functional form adapted from the growth version of (1). We assume the error terms ε_{1jt} and ε_{2jt} are uncorrelated with y_{jt} and with each other. The variances of the error terms (σ_1, σ_2) and that of y (σ_y) are unobserved.

Combining (3a) and (3b), we can write the relationship between growth of lights and growth of measured income as

$$x_{jt} = \beta y_{1jt} + \eta_{jt} \tag{3c}$$

where $\eta_{jt} = \varepsilon_{2jt} - \beta\varepsilon_{1jt}$. An estimate of the structural parameter β in (3c) via OLS would be biased by classical measurement error. Specifically,

$$\text{plim}\,\hat{\beta} = \frac{\sigma_y^2}{\sigma_y^2 + \sigma_1^2}\beta .$$

As described in equation (3a), y_{1jt}, which is what we have in the data, is an imperfect estimate of true income growth y_{jt}. We can potentially improve on this estimate as follows. We estimate equation (2) by OLS to then get fitted values $\hat{y}_{1jt}$. As detailed below, the estimated parameter $\hat{\psi}$ from equation (2) is a highly biased estimate of $1/\beta$, but for the exercise at hand we simply wish to get the best fitted values, $\hat{y}_{1jt}$. We now have two imperfect measures of y_{jt}, namely y_{1jt} and $\hat{y}_{1jt}$. We form a linear combination of the two

$$\hat{y}_{jt} = \lambda y_{1jt} + (1-\lambda)\hat{y}_{1jt} \ , \tag{4}$$

and choose λ to minimize the error with which $\hat{y}_{jt}$ measures y_{jt}. The λ that minimizes var $(\hat{y} - y)$ is given by[11]

$$\lambda^* = \arg\min_{\lambda} \text{var}(\hat{y} - y) = \frac{\sigma_y^2\sigma_2^2}{\sigma_1^2(\beta^2\sigma_y^2 + \sigma_2^2) + \sigma_y^2\sigma_2^2} . \tag{5}$$

λ^* is a function of four unknown parameters ($\sigma_y^2, \sigma_1^2, \sigma_2^2,$ and β), but the observed data provide only three sample moments:

[11] $\text{var}(\hat{y} - y) = \text{var}(\lambda y_1 + (1-\lambda)\hat{y}_1) + \sigma_y^2 - 2\,\text{cov}(\lambda y_1 + (1-\lambda)\hat{y}_1, y)$, where $\hat{y}_1 = [\text{cov}(y_1, x)/\sigma_x^2]x$.

$$\mathrm{var}(y_1) = \sigma_y^2 + \sigma_1^2 \tag{6}$$

$$\mathrm{var}(x) = \beta^2\sigma_y^2 + \sigma_2^2 \tag{7}$$

$$\mathrm{cov}(y_1, x) = \beta\sigma_y^2 \tag{8}$$

As with classical measurement error, there are two ways to proceed. One path would be to estimate the structural parameter β by regressing growth in lights on growth in measured income, using instrumental variables to correct for measurement error in income. Eligible instruments in this case would any variables that drive income growth, such as investment in physical or human capital, changes in institutions, and so on. In general, we were concerned about the validity and power of any instrument for y_1 . For D countries in particular, we could not find variables that were sufficiently good predictors of income growth and were available for a large enough number of countries.

The alternative path, which is the one we chose to follow, is to make an assumption about the ratio of signal to total variance in measured GDP growth, y_1. Define this ratio as

$$\phi = \frac{\sigma_y^2}{\sigma_y^2 + \sigma_1^2}. \tag{9}$$

Note that equation (9) is also the expression for the degree of bias in the estimate of β in equation (3c) under OLS estimation, as described above. If we assume a specific value for ϕ then the optimal λ is given by

$$\lambda^* = \frac{\phi\,\mathrm{var}(y_1)\,\mathrm{var}(x) - \mathrm{cov}(y_1, x)^2}{\mathrm{var}(y_1)\,\mathrm{var}(x) - \mathrm{cov}(y_1, x)^2} = \frac{\phi - \rho^2}{1 - \rho^2}\ , \tag{10}$$

where ρ is the correlation between y_1 and x.

Identification can also be achieved by assuming a value for the ratio of signal to total variance for the second measure: $\theta = \beta^2\sigma_y^2 / (\beta^2\sigma_y^2 + \sigma_2^2)$. We do not know either θ or ϕ, but the data impose a relationship between the two to give a locus of the two possible signal to variance ratios:

$$\theta\phi = \frac{\text{cov}(y_1, x)^2}{\text{var}(y_1)\,\text{var}(x)} = \rho^2 \tag{11}$$

2.2 Application to the D Countries

For the application we proceed as follows. We are going to estimate true income growth from 1992/93-2002/03 by combining information on measured income growth with lights information. The first issue concerns the optimal weight on measured GDP growth from equation (10). The data give us estimates for these countries for $\text{cov}(y_1, x)$, $\text{var}(y_1)$, and $\text{var}(x)$, which are 0.0806, 0.0751 and 0.1704 respectively. From those we get a value of ρ in equation (10) of 0.7124. For values of signal to total variance ratio measures of ϕ = 0.6, 0.75, and 0.9, we would get weights on measured income growth of 0.19, 0.49, and 0.80, with the rest of the weight being on fitted income growth. For purposes of the illustration, we will use ϕ=0.75. This value implies that $\theta = .677$. Thus the illustration assumes that the GDP growth data is a little less noisy than growth as measured by lights.[12]

The next step is to estimate equation (2) to get fitted values of $\hat{y}$ for those countries. Table 3 gives estimates for equation (2) for this sample of countries. The estimated values of ψ for the fixed effect and long difference specifications are respectively 0.396 and 0.473. The long difference ψ is higher than the fixed effect estimate, even after accounting for country coverage differences.[13] We will utilize the long difference formulation, since, in the end, we want to predict 10 year growth rates. Note that the estimates of ψ for the D countries in Table 3 are higher than for the full sample in Table 2. They could be higher because in the structural relationship (3b) the true β differs; that is, there could be a different relationship between income and lights in the less developed countries that make up group D. Alternatively, the

[12] The "40% margin of error" for D countries in the PWT might suggest a signal to total variance ratio of 0.6. However, it is not clear from the PWT literature whether this number applies to levels or growth rates. Our assumption is that growth as measured in our WDI data has less measurement error than cross sectional GDP at purchasing power parity as measured in PWT.

[13] For the 36 countries in column 2 of Table 3 the panel estimate of ψ is 0.403, little different than the 0.396 for 41 countries.

degree of bias in the estimate due to measurement error could differ between the full country sample and the D sample. Note if we wanted ψ for the purposes of estimating $1/\beta$ (rather than just for prediction or fitted value purposes), we would have to account for the degree of bias where

$$\text{p}\lim\hat{\psi} = (1/\beta)\frac{\sigma_x^2 - \sigma_2^2}{\sigma_x^2} \ . \tag{12}$$

Finally, by assuming ϕ=0.75, we can solve also for all unknown parameters of the model in equations 6-9. Foremost is β = 1.43. For σ_y^2, σ_1^2, and σ_2^2, we have 0.056, 0.019 and 0.055. This suggests that in a structural interpretation in equation (3b) the elasticity of lights with respect to true income is 1.43.[14]

2.2.1 Results

Applying the weights to the reported WDI growth rates in local currency units and our fitted values, we can get an estimate of the average annual growth rate of true income, $\hat{y}$, for each of the 36 D countries. These rates are recorded in Table 4 for comparison with WDI estimates. Figure 6 presents a graphical version of the comparison. The horizontal axis records the annualized growth rate of GDP over the decade as measured in the WDI while the vertical axis shows the same thing as measured by the lights data. Points near the 45 degree line in Figure 6 are countries where the two measures give similar results. The further above (below) the 45 degree line is a data point, the higher (lower) is growth in lights data in comparison to growth in the WDI data. The figure also shows a set of iso-composite growth lines, where each iso-composite growth line shows the combinations of lights and WDI based growth rates for which our calculated true growth rate is the same. The slope of these iso-composite growth lines (but not the position of the data points on the graph) will vary with the assumed value of λ; as the weights on lights based growth rates decline, lines become steeper but the points at which they intersect the 45 degree line do not change.

[14] A regression of lights on measured income correspondingly yields an estimate of 1.07, consistent by construction with the 0.75 degree of bias. Note in equation (10) this implies a true ψ of 0.70, while the estimate is 0.47, consistent by construction with the bias in equation (12).

The figure and table suggest that, as would be predicted by a standard analysis of measurement error, growth is more likely to be 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. But there is a lot of variation across countries in the adjustment. By reading the true growth rates versus WDI and lights based numbers in Table 4 or by viewing the divergence between the WDI versus lights based numbers in Figure 6, one can see, that after adjustment, countries like Surinam (SUR) and Papua New Guinea (PNG) have noticeably higher growth rates, but countries like Uzbekistan (UZB) and Central African Republic (CAF), which have similar recorded growth rates, show little change. We downgrade higher growth rate countries like Mozambique (MOZ) and Sudan (SDN), but not Cambodia (KHM), Lao PDR (LAO), or Bhutan (BTN). For D countries at the tails of high or low recorded growth such as Myanmar (MMR), Liberia (LBR), and the Democratic Republic of Congo (COD), lights strongly amend recorded growth rates. For example, in Congo, the WDI data imply an annual average growth of GDP of -2.6% per year while the satellite data imply growth of 2.4% per year. The optimally weighted average is almost exactly zero. In Myanmar, the WDI data say that GDP grew at an annual rate of 8.6% while the lights data imply an annual growth rate of 3.4%. In both these cases, there is reason, beyond the night lights data, to suspect that GDP is particularly poorly measured in the WDI. The Democratic Republic of Congo experienced civil war for much of the period for which we have satellite data, while the economy in Myanmar was largely autarkic and non-market.

3. Application: Does local agriculture contribute to local city growth?

As noted in the introduction, urban economists model city growth as a process disconnected from agriculture both in theory and empirically (Glaeser *et al* 1992 and Glaeser and Saiz 2004). Development economists have long recognized the rural-urban interaction in two-sector models dating back to Lewis (1954), but most modeling assumes that the rural sector is just a source of labor for the growing urban sector. On the empirical side, Brueckner (1990) looks at city sizes as they relate to rural-urban income gaps. Using aggregate country data, he finds that higher rural incomes retard urbanization and the growth of the largest city in a country.

Da Mata *et al* (2007) find that higher rural incomes in city hinterlands also retard city population growth in Brazil.

What these approaches generally miss is the positive side: higher rural incomes can contribute to local urban economic growth, something that is hinted at in the new economic geography literature (Krugman, 1991), as well as in da Mata *et al* (2007) for Brazil. This notion has long been pursued by agricultural economists, as well as a few growth economists (e.g., Kuznets, 1955; Kogel and Prskawetz 2001, Irz and Roe 2005, Tiffin and Irz, 2006). Local agricultural growth can generate local savings and investment in manufacturing and services, which are more urbanized activities. Farmers with increased incomes in a city hinterland demand more urban output such as farm machinery, household items, and personal and business services.

However no studies have had the data to do a convincing empirical analysis to show that *exogenous* increases in farm incomes in a city's hinterland causally spur urban income growth in that city. In this section we examine a panel of 541 cities in 18 African countries over 9 years. As explained in the Appendix, the selection of countries is in part dictated by needing city population data and co-ordinates so as to identify cities. For 14 of the countries, data cover all cities with populations over 10,000 in 2008 within 3 km of a night light source, while for the other countries the minimum population size is 5,000- 20,000 (see Table A2). We have annual data on rainfall and on lights. Rainfall is an exogenous source of increases in agricultural yields and incomes in many African contexts (Miguel, Sergenti and Satyanath, 2004; World Bank, 2005). We don't have income data for these cities at all, and we have population data for at most one year in the time period for which we have detailed rainfall. However we have lights for every year. Our presumption is that increased rain increases agricultural productivity and thus income in hinterland areas of cities. Farmers' spending increases demand for urban goods, raising urban income. The rise in urban income leads to an increase in lights. We test the net result directly—increased hinterland rainfall spurs urban lights.

The formulation we use is

$$\ln(x_{jt}) = \sum_{i=0}^{k} \beta_i r_{j,t-i} + \alpha_j + \lambda_t + \varepsilon_{jt} \tag{13}$$

where x_{jt} is lights in city *j* in time *t* and $r_{j,t-i}$ is rainfall in the hinterlands of city *j* at time *t-i*. In equation (13) current and prior years' rainfall affect current lights after allowing for city and time fixed effects. The lag structure in (13) implies that productivity shocks in agriculture persist in changing urban incomes beyond the current year. So for example, farmers who get windfall income in a year may smooth spending in urban areas over several successive years. Also, income windfalls in agriculture may result in increased investments in agricultural production (seeds, fertilizer and equipment) which generate agricultural income gains in succeeding years, which in turn increase demand for urban products. We will find that effects attenuate at $k = 4$; and we will look at the falsification test of adding a lead year of rain. Also in interpreting equation (13), lights could increase with rainfall because urban incomes rise due to either per capita urban income growth, population growth, or both. While we can't distinguish the two, in this case it seems likely to be per capita income growth. City population effects likely go in the opposite direction: other studies suggest that improved agricultural incomes reduce migration from rural hinterlands to cities (Brueckner, 1990 and da Mata *et al*, 2007).

An issue in estimation of (13) concerns the distribution of the ε_{it}. We allow for clustering of the ε_{it} by city, but the process may be more distinct. We might expect serial correlation along the lines of an AR[1] process. Other conditions facing a city that vary over time may be serially correlated in a common fashion across cities. We will look at both fixed effects and AR[1] estimates. A second concern is that in 7% of city-years, x_{jt} equals zero, so $\ln(x_{jt})$ is undefined. Generally we rely on OLS, but replace $\ln(x_{jt})$ with $\ln(x_{jt} + \delta)$, where $\delta=1$.[15] Note that 2 is the smallest nonzero value of lights in the data We also present a Tobit specification for $\ln(x_{jt})$, with truncation when the light measure falls below 2 and is not recorded. The Tobit results are almost identical to OLS ones. Fixed effects Tobits are biased for short panels, but our panel is not that short and most observations are not censored.

In application of equation (13), the impact of agricultural rain may differ according to the urban context. Large industrialized cities may be more independent of local agricultural conditions, relying more on national and international trade in industrial goods. Smaller cities may be more grounded in local hinterland economies and more sensitive to changes in agricultural productivity. We explore this by looking at whether effects vary between primate

[15] Results with $\delta=0.5$ and $\delta=2$ produce coefficients 10-20% larger than and smaller than, respectively, results with $\delta=1$ for all $k \geq 0$, but with correspondingly different standard errors, so t-statistics are within 5% of their counterparts when $\delta=1$. Results are very similar to those cited if δ is only added to city-years with light values of zero, instead of all cities, before logging. Using unlogged lights values produces effects in the same directions, except for k=3 in some specifications, but most coefficients are no longer significant. This makes sense – one would not expect a linear effect of the same amount of rain across all city sizes. No approach is available for AR[1] errors that is equivalent to Honoré's (1992) censored fixed effects method for models using clustered errors.

cities and other cities in the sample. We define primate cities as the largest or the effective capital cities in each of our 18 countries. For all but Malawi, the capital and largest city are the same. We will also look at whether results differ for cities of less than versus more than 200,000 people.

City Data

We have two main sources of data for our African cities. First are the lights. We have no city boundaries, so we define cities as contiguous lit areas. Figure 7a illustrates our methodology for an area of southern Ghana. The boundaries of contiguous sections of lights on the landscape are marked for different years. We draw the outer envelope of contiguous lit pixels across all years and define this as the potential urban area. Then, as shown in Figure 7b, we map in jurisdictional cities as points, based on geo-coordinates identified with each city (see Appendix). The population for each lit area is the sum of the city populations in that area. In the overwhelming majority of cases (502 of 541), there is only one city per lit area (as in the south-east corner of Figure 7b). In the other 39, larger urban areas consist of several jurisdictions as pictured in the northwest portion of Figure 7b, where 3 cities fall within the same light. The second data source is annual rainfall estimates (Love *et al* 2004), recorded on a 0.1 degree grid (approximately 124 sq. km at the equator). The rain data only exist starting in 1995, so we cannot use the first three years of the lights. We draw a 30 km buffer around each lit area (i.e. the green area pictured in Figure 7b) to create a catchment area. We measure average rainfall over all grid entries that are in the catchment area but outside the lit area.[16]

Basic results for rainfall effects on urban incomes

Columns 1-5 in Table 5 state the basic results. With clustered robust errors and no AR[1] structure, columns 1-4 show different lag structures. Column 1 includes only rain in the contemporaneous year; column 2 allows for three years of effects; column 3 for four years; and column 4 for five years. It is clear rain from two years before the present still has a significant effect on urban income. In columns 3 and 4 coefficients for rain from three years prior to the

[16] Results were broadly similar when radii from 20 to 70 km were used.

current year are smaller, and in column 4, rain from four years prior has an insignificant (negative) coefficient. We generally use a lag structure with three or four years of rain, including the current year, in further specifications. In column 5 we give Tobit results for column 4; they are almost identical. In column 6, we re-estimate column 3 imposing an AR[1] process.[17] That slightly reduces the rain effect of the first two lags. In columns 7 and 8 we conduct a falsification test by adding a lead year of rain, which should have no effect. With an AR[1] process modeling serial correlation, the lead year has no effect. The fact that the lead shows some effect for the ordinary panel estimation suggests that not modeling the serial correlation in the data can result in misleading estimates.

Rainfall effects are arguably large. Each one standard deviation increase in rain (0.90 mm/day), in the current or either of the prior two years, leads to a roughly 14% increase in lights. From Table 2 a 14% increase in lights represents about a 4% increase in GDP for a city. Thus, a sustained increase in rain over several years would have a very strong effect on urban incomes.

However, the effect of hinterland rain on city growth is heterogeneous and differs by type of city. Bigger, more industrialized cities are less dependent on their hinterlands, as are political centers. Table 6 shows that primate cities have much lower rainfall effects. For one year of rain the coefficient of 0.155 is just 0.054 for primate cities. When three years of rain are included, the coefficients for year t, t-1 and t-2 are 0.16, 0.15, and 0.15 for ordinary cities, while for primate cities they are 0.084, 0.073 and 0.053. Treating the 29 cities with a population over 200,000 in 1995 as primate cities, in column 3, the differential in coefficients is almost the same as column 2. Allowing for an AR[1] structure in columns 4 and 5 yields similar effects.

Robustness

Finally, we consider two alternative explanations of the results. First, more rain results in cheaper hydroelectric power, which drives increased usage of electricity. Hydroelectric power is very common in sub-Saharan Africa. While hydroelectric power accounted for only about 20% of electricity generation in sub-Saharan Africa in 2003, it represented more than half of

[17] We use Foster and Lee's (2009) version of the method of Hansen (2007).

generation for more than half of the countries, including 12 of 18 in our sample.[18] In order to test this hypothesis, we construct a crude national measure of hydro dependence, hydro generation divided by total electricity consumption averaged across all years in the sample. It is crude because some countries import and export a lot of electricity, and we cannot identify imports or exports by country pair or by generation type. When this measure of hydro dependence is interacted with rainfall, countries more dependent on hydro have smaller rainfall effects, not larger ones, and the interacted term is not significant. Thus hydroelectric power costs do not seem to be at the heart of what is going on. That conclusion is also consistent with the Table 6 result that primate cities are less affected by rainfall changes than non-primate cities, since inhabitants of primate cities are likely to have a much greater reliance on electric power, in terms of both household coverage and intensity of lighting per household.

Second, the satellite only takes data when there are no clouds over a place, and rain requires clouds, so one might expect that high rainfall is associated with noisy lights measurements averaged over fewer nights. We can confirm empirically that more rain is correlated with fewer nights of lights data. However, controlling for the number of nights of data has little effect on our results. Similarly, controlling for the 3% of lights that contain at least one top-coded pixel saturating the sensor has little effect.

4. Conclusion

Satellite night-lights data are a useful proxy for economic activity at temporal and geographic scales for which traditional data are of poor quality or unavailable. We developed a statistical model to optimally combine data on changes in night lights with data on measured income growth to improve estimates of true income growth. We applied the methodology to countries with low quality national income data, the D countries in the PWT. For these 36 countries, we get a new set of income growth numbers for the 10 years 1992/3 – 2002/3. These estimates differ from data in WDI by as much as several percent per year. As a second application in which no income measures are available, we considered the interaction between the economies of urban areas and their rural hinterlands in Africa, and demonstrated that

[18] Energy Information Administration (USA; EIA). 2007. *International Energy Annual 2005*. http://www.eia.doe.gov/pub/international/iea2005/iea_2005.zip Accessed 9 April 2008

productivity shocks in the form of rainfall in agriculture contribute strongly to economic growth of the cities serving agriculture. This is the first empirical contribution to the debate about whether rural hinterlands contribute to urban growth.

Appendix: Data

A. Lights

The Version 2 Defense Meteorological Satellite Program Optical Linescan System (DMSP-OLS) Nighttime Lights Time Series data are available from the National Oceanic and Atmospheric Administration's (NOAA) National Geophysical Data Center (NGDC) as a set of annual composites, currently for 1992-2003.[19] This most recent version of the data is a series of 18 annual composites from 4 satellites each operating for overlapping periods of 3 to 6 years between 1992 and 2003.

Each annual composite is a raster (grid) dataset with values every 30 seconds of latitude and longitude (approximately 0.86 km^2 at the equator, decreasing with the cosine of latitude) between 65 degrees north and 65 degrees south latitude. The exclusion of high latitude zones affects approximately 3 million people, or 0.05% of the global total, in 7 countries. Each grid value is an eight-bit integer (0-63) called a digital number (DN), averaged for over all nights fitting certain criteria (i.e. not too much moonlight, sunlight, aurora activity or cloud cover). They were compiled and cleaned, removing temporary features such as forest fires, by NGDC. A calibration has been applied to ensure greater comparability across satellite-years, but they cannot be interpreted directly as physical units of light (Chris Elvidge, personal communication).

Global lights data have several problems besides this lack of true calibration. First, the sensor saturates at a level of light that is very common in the cities and towns of rich countries, resulting in topcoded values. At high latitudes no summer data can be used because sunlight is still contaminating images at local pass times of 8:30 to 10 pm. This effect is diminished closer to the equator. The data are subject to overglow or blooming, which means that lights tend to appear larger than they actually are, especially for bright lights and over water. Snow tends to magnify light values. Humidity, which varies significantly across the continent, is known to affect the performance of other sensors but has never been studied in relation to the DMSP-OLS.

[19] Available at http://www.ngdc.noaa.gov/dmsp/global_composites_v2.html

Many of these problems are not likely to be important in the Africa city examination, as there are few instances of top-coding, no long summer nights, and no snow. Further details about the lights data and processing can be found in Elvidge *et al*. (1997, 1999, 2002, 2003, 2005), Lieske (1981), and Small, Pozzi and Elvidge (2005).

For the Africa section of the present paper, lights were processed as follows. In any given year, the overwhelming majority of land in Africa is unlit (i.e. it has a DN value of zero). Lit areas corresponding to cities, oil wells or other features thus form a set of polygons. This set varies from year to year, but there is substantial overlap, as individual features tend to persist over time. In order to have coherent units of analysis we create a combined map, in which the value of a given pixel is the maximum value of all 18 individual satellite-years. The result of this is a set of 9189 non-contiguous polygons on the African continent in which all pixels were lit for at least one year. For each of these, the total calibrated digital number for each satellite-year, as well as the minimum and maximum pixel, were reported.

B. African cities

City location and population

In order to identify cities we need a data source with cities and their populations (which also allows us to separate effects by city size). Cities and their population are obtained from www.citypopulation.de. Only countries for which information is available for at least one census after 1994 are used. Island states were also dropped. While population figures are not necessarily taken directly from the official census bureaus, spot checks suggest that they are consistent with the official figures, where available. Five countries (Algeria, Egypt, Morocco, South Africa, and Tunisia) were dropped because massive agglomerated lights containing significant proportions of their populations make them qualitatively different than the rest of the continent. Three more, (Republic of Congo, Swaziland and Lesotho) were dropped because of significant contamination across their borders by lights from other countries, namely Democratic Republic of Congo, Angola and South Africa). While this is in itself an interesting phenomenon, it would render interpretation too difficult for the present exercise. Lastly, Western Sahara was removed because its sovereignty has been contested over the course of the study period. This left 18 countries (listed in Table A2) and 782 identified cities above some threshold size as reported by www.citypopulation.de. Of the 9,189 African lights, 2,323 have centroids falling within the

18 countries selected above. In 13 of these countries, all settlements of more than 10,000 are purported to be reported by www.citypopulation.de. However, Mozambique and Ghana's nominal cutoffs are 20,000, Mauritania's is 15,000, Central African Republic's is 5,000, and Rwanda lists no cutoff. Furthermore, the benchmark year for these cutoffs is never specified, and in practice, 14 of 782 cities have lower populations than their nominal cutoff for every year up to 2008.

Latitudes and longitudes for African cities

Latitudes and longitudes were assigned from three sources: www.citypopulation.de, the Gridded Rural Urban Mapping Project,[20] and www.world-gazetteer.com. Locations were validated with respect to satellite imagery in Google Earth to ensure that they indeed fell in or very near a city. However, no further information was available to ensure that it was the named city, other than the three original sources. In a few instances, one of the three coordinate sources was chosen because it placed the city within a light, whereas another source did not. We consider this appropriate because we are not attempting to demonstrate the well-known collocation of cities and lights (e.g. Welch 1980), but rather to use this fact for further analysis. For fifteen cities in three countries (Tanzania, Mauritania, and Ghana) no coordinate information was available. This reduced the sample to 767 cities.

Combining lights and population

Each light in the sampled countries is assigned the population of all cities within 3 kilometers. The three kilometer buffer is used because of measurement error in the latitude/longitude data and the georeferencing of the lights, following Balk *et al*. (2004). 111 city points fell farther than 3 km from the nearest light in the sample. In most cases, the points that fell within 3 km fell within 1 km, as would be expected from simple rounding of coordinates to the nearest hundredth of a degree. This reduced the set of city-points from 767 to 656. However, only one of these 111 has a population over 25,000, and it is a border city that would have fallen within a light whose centroid fell in another country if such lights were included in the sample.

[20] Center for International Earth Science Information Network (CIESIN), Columbia University; International Food Policy Research Institute (IFPRI); The World Bank; and Centro Internacional de Agricultura Tropical (CIAT). 2004. Global Rural-Urban Mapping Project (GRUMP), Alpha Version. Palisades, NY: Socioeconomic Data and Applications Center (SEDAC), Columbia University. At http://sedac.ciesin.columbia.edu/gpw.

Only six more have populations over 20,000, and one of these would have fallen within a light whose centroid fell in another country if such lights were included in the sample. So it is plausible that nearly all of these 111 points are just too small to be seen by the satellites.

Of the 2323 lights, 541, or 23% contain at least one city for which we have population. However, the others are far less bright and/or extensive lights on average, consistent with the idea that they correspond to smaller settlements not included in the population data. They could also correspond to mines or other industrial facilities. Of the 541 lights with populated places, thirty-five touch a border, at least according to one common set of international boundaries.[21] Of these, seven contain city points on both sides of the border.[22]

Rainfall

Rainfall data for each 0.1 degree grid cell (approximately 124 km^2 at the equator) are from the NOAA Climate Prediction Center's Africa Rainfall Climatology (ARC; Love *et al*. 2004). Unlike most commonly used rainfall data, these are estimates based on both rain gauges and satellite measurements. The addition of satellite measurements is especially important in Africa, where actively reporting rain gauges are sparsely located. It means that neighboring observations are significantly less dependent than those based on stations alone. Ideally we would calculate rainfall for years corresponding to agricultural seasons, like Maccini and Yang (2008). However, seasons vary across Africa, and the lights composites are only available for calendar years anyway. Data are available for 1995 to the present.

References

Au, Chun-Chung and J. Vernon Henderson. 2006. “Are Chinese Cities Too Small?” *Review of Economic Studies*, 73(3): 549-76.

Balk, Deborah, Francesca Pozzi, Gregory Yetman, Uwe Deichmann, and Andy Nelson. 2004. "The Distribution of People and the Dimension of Place: Methodologies to Improve the

[21] ESRI Data & Maps 2002 : World Administrative Units. Using a different set of international boundaries might result in a slightly different list.

[22] This of course requires both countries to be in the sample – in a few other cases it is possible that a city in an unsampled country falls within the same light.

Global Estimation of Urban Extents."
http://sedac.ciesin.columbia.edu/gpw/docs/UR_paper_webdraft1.pdf.

Black, Duncan and J. Vernon Henderson. 1999. "The Theory of Urban Growth," *Journal of Political Economy*, 107: 252-84.

Browning, Martin and Thomas Crossley. 2009. “Are Two Cheap, Noisy Measures Better Than One Expensive, Accurate One?” *American Economic Review*, 99(2): 99-103.

Brueckner, Jan K. 1990. “Analyzing Third World Urbanization: A Model with Empirical Evidence.” *Economic Development and Cultural Change*, 38(3): 587-610.

Croft, T.A., 1978. “Night-time Images of the Earth from Space.” *Scientific American*, 239: 68-79.

da Mata, Daniel, Uwe Deichmann, J. Vernon Henderson, Somik V. Lall, and H. G. Wang. 2007. “Determinants of City Growth in Brazil.” *Journal of Urban Economics*, 62(2): 252-72.

Dawson, John W., Joseph P. DeJuan, John J. Seater, and E. Frank Stephenson. 2001. “Economic Information versus Quality Variation in Cross-Country Data.” *Canadian Journal of Economics*, 34(3): 988-1009.

Deaton, Angus and Alan Heston. 2008. “Understanding PPPs and PPP-based National Accounts.” NBER Working Paper No. 14499.

Doll, Christopher N.H., Jan-Peter Muller and Jeremy G. Morley. 2006. Mapping Regional Economic Activity from Night-time Light Satellite Imagery”, *Ecological Economics* 57(1): 75-92.

Ebener, Steeve, Christopher Murray, Ajay Tandon and Christopher D. Elvidge. 2005. “From Wealth to Health: Modeling the Distribution of Income per Capita at the Sub-national Level Using Night-time Light Imagery.” *International Journal of Health Geographics*, 4(5): 1-17.

Elvidge, Christopher D., Kimberley E. Baugh, Eric A. Kihn, Herbert W Kroehl, and Ethan R. Davis. 1997. “Mapping of City Lights Using DMSP Operational Linescan System Data.” *Photogrammetric Engineering and Remote Sensing*, 63: 727-34.

Elvidge, Christopher D., Kimberley 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., Ruth Hobson, Ingrid L. Nelson, Jeffrey M. Safran, Benjamin T. Tuttle, Kimberley E. Baugh, and John B. Dietz. 2002. “Global Observation of Urban Areas Based on Nocturnal Lighting.” *Land Use and Land Cover Change Newsletter,* 8: 10-12.

Elvidge, Christopher D., Ruth Hobson, Ingrid L. Nelson, Jeffrey M. Safran, Benjamin T. Tuttle, John B. Dietz, and Kimberley E. Baugh. 2003. “Overview of DMSP-OLS and Scope of Applications.” In *Remotely Sensed Cities*, ed. Victor Mesev, 281-99. London: Taylor and Francis.

Elvidge, Christopher D., Kimberley E. Baugh, Jeffrey M. Safran, Benjamin T. Tuttle, Ara T. Howard, Patrick J. Hayes, and Edward H. Erwin. 2005. “Preliminary Results From Nighttime Lights Change Detection.” *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(8).

Foster, Andrew D. and Yongsuk Lee. 2009. “Staffing Subsidies and the Quality of Care in Nursing Homes: FGLS Estimates – How to Effectively Double the Number of States in your Analysis.” Unpublished.

Glaeser, Edward L, Hedi D. Kallal, Jose A. Scheinkman, and Andrei Shleifer. 1992. "Growth in Cities." *Journal of Political Economy*, 100(6): 1126-52.

Glaeser, Edward L. and Albert Saiz. 2004. "The Rise of the Skilled City." *Brookings-Wharton Papers on Urban Affairs*, 47–94.

Good, David F. 1994. “The Economic Lag of Central and Eastern Europe: Income Estimates for the Habsburg Successor States, 1870-1910.” *The Journal of Economic History*, 54(4): 869-91.

Hansen, Christian. 2007. “Generalized Least Squares Inference in Panel and Multilevel Models with Serial Correlation and Fixed Effects.” *Journal of Econometrics*, 140: 670-94.

Harris, John R. and Michael P. Todaro. 1970. “Migration, Unemployment and Development: a Two-Sector Analysis”, *American Economic Review*, 60, 126-42.

Henderson, J. Vernon., and Hyoung Gun Wang. 2005. “Aspects of the Rural-Urban Transformation of Countries.” *Journal of Economic Geography,* 5(1): 23-42

Honoré, Bo E. 1992. “Trimmed LAD and Least Squares Estimation of Truncated and Censored Regression Models with Fixed Effects.” *Econometrica*, 60: 533–565.

International Energy Agency (IEA). 2006. *World Energy Outlook 2006*. Paris: IEA.

International Monetary Fund (IMF). 2006. “Jamaica: Selected Issues.” IMF Country Report No. 06/157.

Irz, X and T. Roe. 2005. “Seeds of Growth? Agricultural Productivity and the Transitional Dynamics of the Ramsey Model.” *European Review of Agricultural Economics*, 32(2): 143-65.

Johnson, Simon, William Larson, Chris Papageorgiou, and Arvind Subramanian. 2009. “Is Newer Better? The Penn World Table Revisions and the Cross-Country Growth Literature.“ Unpublished.

Kögel, Tomas and Alexia Prskawetz. 2001. “Agricultural Productivity Growth and Escape from the Malthusian Trap.” *Journal of Economic Growth*, 6(4): 337-57.

Krueger, Alan B. and Mikael Lindahl. 2001. “Education for Growth: Why and for Whom?” *Journal of Economic Literature*, 39(4): 1101-36.

Krugman, Paul. 1991. "Increasing returns and economic geography.” *Journal of Political Economy*, 99: 483-99.

Kuznets, Simon. 1955. “Economic Growth and Economic Inequality.” *American Economic Review*, 45: 1-28.

Lewis,W. Arthur. 1954. “Economic Development with Unlimited Supplies of Labor.” *The Manchester School*, 22: 139–91.

Lieske, R.W. 1981. “DMSP primary sensor data acquisition.” *Proceedings of the International Telemetering Conference*, 17:1013-20.

Love, Timothy B., Vadlamani Kumar, Pingping Xie, and Wassila Thiaw. 2004. “A 20-year Daily Africa Precipitation Climatology Using Satellite and Gauge Data.” Paper presented at the American Meteorological Society Conference on Applied Climatology, Seattle, WA.

Maccini, Sharon and Dean Yang. 2009. “Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall.” *American Economic Review*, 99(3): 1006-26.

Miguel, Edward, Shanker Satyanath and Ernest Sergenti. 2004. “Economic Shocks and Civil Conflict: An Instrumental Variables Approach.” *Journal of Political Economy*, 112(4): 725-53.

Nerlove, Marc and Efraim Sadka. 1991. "Von Thünen's Model of the Dual Economy." *Journal of Economics*, 54(2): 97-123.

Nuxoll, Daniel A. 1994. “Differences in Relative Prices and International Differences in Growth Rates.” *American Economic Review*, 84(5): 1423-36 (December).

Small, Christopher, Francesca Pozzi and Christopher D. Elvidge. 2005. “Spatial Analysis of Global Urban Extent from DMSP-OLS Nighttime Lights.” *Remote Sensing of Environment,* 96: 277-91.

Steckel, Richard H. and Jerome C. Rose. 2002. *The Backbone of History: Health and Nutrition in the Western Hemisphere*. Cambridge, UK: Cambridge University Press.

Sutton, Paul C. and Robert Costanza. 2002. “Global estimates of market and non-market values derived from nighttime satellite imagery, land cover, and ecosystem service valuation.” *Ecological Economics*, 41(3): 509-27.

Sutton, Paul C., Christopher D. Elvidge and Tilottama Ghosh. 2007. “Estimation of Gross Domestic Product at Sub-National Scales Using Nighttime Satellite Imagery.” *International Journal of Ecological Economics & Statistics,* 8(S07): 5-21.

Tiffin, Richard and Xavier Irz. 2006. “Is Agriculture the Engine of Growth?” *Agricultural Economics,* 35(1): 79-89.

von Thunen, Johann Heinrich. 1826. *von Thunen's Isolated State*. Tr. Peter Hall. London: Pergamon, 1966.

Welch, R., 1980. “Monitoring Urban Population and Energy Utilization Patterns From Satellite Data.” *Remote Sensing of Environment*, 9: 1-9.

World Bank. 2005. Managing Agricultural Production Risk: Innovations in Developing Countries. Report No. 32727-GLB.

Table 1: Night Lights Data for Selected Countries, 1992-2003 average

Digital Number (DN)	USA	Canada	Netherlands	Brazil	Costa Rica	Guatemala	Bangladesh	Madagascar	Mozambique	Malawi
0	67.74%	93.38%	0.89%	94.07%	69.10%	82.37%	68.20%	99.74%	99.56%	97.65%
1-2	0.00%	0.00%	0.00%	0.01%	0.00%	0.01%	0.30%	0.00%	0.01%	0.00%
3-5	6.36%	0.46%	0.38%	2.20%	11.33%	9.78%	20.02%	0.13%	0.23%	0.84%
6-10	13.42%	3.24%	17.15%	2.13%	13.01%	5.13%	7.99%	0.07%	0.11%	0.95%
11-20	5.89%	1.68%	32.05%	0.79%	3.56%	1.57%	2.02%	0.03%	0.04%	0.29%
21-62	5.56%	1.15%	46.37%	0.71%	2.54%	0.99%	1.36%	0.03%	0.04%	0.27%
63-65	1.02%	0.09%	3.16%	0.09%	0.45%	0.16%	0.10%	0.00%	0.00%	0.01%
% area unlit	64.87%	92.14%	0.85%	94.28%	69.53%	82.89%	68.04%	99.74%	99.58%	97.16%
Avg. DN	5.0249	0.8947	22.3948	0.6664	3.1691	1.4412	2.2637	0.0257	0.0398	0.3135
Gini(DN)	0.8286	0.9597	0.3636	0.9682	0.8229	0.8958	0.7929	0.9985	0.9977	0.9864
Population density (per sq. km)	30	3	463	20	73	98	1021	26	22	116
% urban	78%	79%	75%	80%	57%	44%	23%	26%	28%	14%
GDP per capita, PPP (2005 $)	36126	29675	30502	7728	7575	3785	839	833	475	663

Notes:

1) values of 64 and 65 are possible because of relative calibration across years.

2) % area unlit accounts for differences in cell area, whereas the percentage of cells having digital number 0, 1-2, etc. does not.

3) each figure is calculated within satellite-years, averaged across satellites within a year, and then across years.

Table 2. Baseline results for the world: 1992-2003; growth in real GDP (local currency units)

	Fixed effects specifications			Long differences
	(1)	(2)	(3)	(4)
	ln(GDP)	ln(GDP)	ln(GDP)	ln(GDP)
ln(lights/area)	0.287***	0.270***	0.286***	0.324***
	[0.046]	[0.044]	[0.050]	[0.041]
$\ln(\text{lights/area})^2$		-0.01		
		[0.011]		
gini(lights)			-0.005	
			[0.199]	
Constant				0.227***
				[0.018]
Observations	2149	2149	2149	170
Number of countries	187	187	187	170
(Within-country) R-sq	0.661	0.664	0.661	0.315
Country fixed effects	yes	yes	yes	no
Year fixed effects	yes	yes	yes	no
Error treatment	Robust, clustered by country	Robust, clustered by country	Robust, clustered by country	Robust

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in brackets
In column 4 long differences are calculated averaging the first and last two years of levels data.

Table 3. Results for "D" countries: 1992-2003; growth in real GDP (local currency units)

	Fixed effects	Long differences
	(1)	(2)
	ln(GDP)	ln(GDP)
ln(lights/area)	0.396***	0.473***
	[0.107]	[0.066]
Constant		0.220***
		[0.039]
Observations	466	36
Number of countries	41	36
(Within-country) R-sq	0.634	0.507
Country fixed effects	yes	no
Year fixed effects	yes	no
Error treatment	Robust, clustered by country	Robust

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in brackets
In column 2 long differences are calculated averaging the first and last two years of levels data.

Table 4. Ten-year growth rates in true income, $\hat{y}$, for "D" countries (1992/93-2002/03)

Country	ISO code	WDI (LCU)	fitted lights	optimal combination of WDI and fitted lights	difference
Myanmar	MMR	0.826	0.337	0.578	-0.248
Liberia	LBR	0.922	0.643	0.780	-0.141
Mozambique	MOZ	0.719	0.499	0.607	-0.112
Angola	AGO	0.484	0.285	0.383	-0.101
Sudan	SDN	0.549	0.367	0.457	-0.093
Togo	TGO	0.398	0.223	0.309	-0.089
United Arab Emirates	ARE	0.554	0.393	0.473	-0.082
Malta	MLT	0.366	0.236	0.300	-0.066
Uganda	UGA	0.661	0.548	0.604	-0.058
Yemen, Rep.	YEM	0.518	0.410	0.463	-0.055
Belarus	BLR	0.162	0.059	0.110	-0.052
Algeria	DZA	0.277	0.176	0.226	-0.051
Mongolia	MNG	0.312	0.238	0.275	-0.037
Niger	NER	0.332	0.279	0.305	-0.027
Guinea-Bissau	GNB	0.001	-0.051	-0.025	-0.026
Cyprus	CYP	0.377	0.330	0.353	-0.024
Seychelles	SYC	0.274	0.236	0.255	-0.020
Uzbekistan	UZB	0.216	0.187	0.201	-0.015
Central African Republic	CAF	0.191	0.170	0.180	-0.011
Chad	TCD	0.412	0.392	0.402	-0.010
Comoros	COM	0.154	0.137	0.145	-0.009
Namibia	NAM	0.379	0.379	0.379	0.000
Cambodia	KHM	0.684	0.722	0.703	0.019
Lao PDR	LAO	0.617	0.665	0.642	0.024
Bhutan	BTN	0.637	0.713	0.676	0.038
Guyana	GUY	0.332	0.412	0.373	0.041
Cape Verde	CPV	0.595	0.687	0.642	0.047
Lesotho	LSO	0.307	0.399	0.354	0.047
Saudi Arabia	SAU	0.179	0.277	0.229	0.049
Haiti	HTI	-0.031	0.192	0.082	0.113
Suriname	SUR	0.200	0.436	0.320	0.119
Eritrea	ERI	0.450	0.686	0.570	0.120
Djibouti	DJI	-0.040	0.215	0.090	0.130
Papua New Guinea	PNG	0.105	0.407	0.258	0.153
Tajikistan	TJK	-0.227	0.102	-0.060	0.167
Congo, Dem. Rep.	COD	-0.264	0.241	-0.008	0.256

Table 5: Results for African rainfall and city growth, 1995-2003

	(1) ln(light(t)+1)	(2) ln(light(t)+1)	(3) ln(light(t)+1)	(4) ln(light(t)+1)	(5) ln(light(t))	(6) ln(light(t)+1)	(7) ln(light(t)+1)	(8) ln(light(t)+1)
rain(t)	0.152***	0.159***	0.201***	0.149***	0.158***	0.162***	0.223***	0.251***
	[0.041]	[0.043]	[0.050]	[0.051]	[0.055]	[0.041]	[0.056]	[0.060]
rain(t-1)		0.150***	0.160***	0.183***	0.193***	0.137***	0.153***	0.178***
		[0.035]	[0.045]	[0.059]	[0.063]	[0.042]	[0.049]	[0.052]
rain(t-2)		0.146***	0.156***	0.165***	0.176***	0.144***	0.123**	0.132**
		[0.040]	[0.042]	[0.052]	[0.057]	[0.040]	[0.051]	[0.053]
rain(t-3)			0.074*	0.090*	0.095*	0.098***	0.090**	0.079*
			[0.042]	[0.049]	[0.053]	[0.038]	[0.040]	[0.044]
rain(t-4)				-0.051	-0.051			
				[0.043]	[0.046]			
rain(t+1)							0.061	0.107**
							[0.046]	[0.050]
Observations	4869	3787	3246	2705	2705	2705	2164	2705
Cities	541	541	541	541	541	541	541	541
(Within-city) R-sq	0.046	0.055	0.041	0.048	0.048	0.032	0.036	0.043
City fixed effects	yes	yes	yes	yes	yes	yes	yes	yes
Year fixed effects	yes	yes	yes	yes	yes	yes	yes	yes
Error treatment	robust, cluster on city	robust, cluster on city	robust, cluster on city	robust, cluster on city	cluster on city in Tobit	AR[1], Foster and Lee (2009)	AR[1], Foster and Lee (2009)	robust, cluster on city

*** p<0.01, ** p<0.05, * p<0.1

Table 6: Rainfall: differential effect on primate cities, 1995-2003

	(1) ln(lights(t)+1)	(2) ln(lights(t)+1)	(3) ln(lights(t)+1)	(4) ln(lights(t)+1)	(5) ln(lights(t)+1)
rain(t)	0.155***	0.161***	0.163***	0.132***	0.134***
	[0.042]	[0.045]	[0.046]	[0.036]	[0.037]
primate*rain(t)	-0.102**	-0.077*	-0.085**	-0.076**	-0.068*
	[0.045]	[0.043]	[0.042]	[0.039]	[0.036]
rain(t-1)		0.152***	0.153***	0.126***	0.128***
		[0.036]	[0.037]	[0.035]	[0.036]
primate*rain(t-1)		-0.079*	-0.079**	-0.083*	-0.075**
		[0.043]	[0.037]	[0.045]	[0.037]
rain(t-2)		0.148***	0.148***	0.116***	0.117***
		[0.041]	[0.042]	[0.036]	[0.036]
primate*rain(t-2)		-0.095**	-0.062	-0.096**	-0.075*
		[0.045]	[0.046]	[0.042]	[0.040]
Observations	4869	3787	3787	3246	3246
Cities	541	541	541	541	541
(Within-city) R-sq	0.046	0.056	0.056	0.053	0.053
Primate definition	political	political	pop>200k	political	pop>200k
Error structure	robust, cluster on city	robust, cluster on city	robust, cluster on city	AR[1] Foster and Lee (2009)	AR[1] Foster and Lee (2009)
City fixed effects	yes	yes	yes	yes	yes
Year fixed effects	yes	yes	yes	yes	yes

*** p<0.01, ** p<0.05, * p<0.1

Table A1: Descriptives

Variable	N	mean	sd	min	max	Sample
change in ln(GDP), LCU	170	0.315	0.211	-0.384	0.922	all countries (except GNQ)
change in ln(lights)	170	0.271	0.365	-1.119	1.525	all countries (except GNQ)
change in ln(GDP), LCU	36	0.351	0.274	-0.264	0.922	grade D countries
change in ln(lights)	36	0.276	0.413	-0.573	1.061	grade D countries
gini(lights)	2149	0.820	0.215	0.045	1.000	all countries (except GNQ)
ln(std. dev.(lights))	2149	1.439	0.966	-1.430	3.085	all countries (except GNQ)
gini(lights)	466	0.907	0.191	0.189	1.000	grade D countries
ln(std. dev.(lights))	466	0.718	1.052	-1.430	2.982	grade D countries
ln(lights(t)+1)	4869	5.548	2.126	0	11.426	African cities
rain(t)	4869	1.903	0.904	0.007	5.111	African cities
rain(t-1)	4328	1.886	0.893	0.007	5.111	African cities
rain(t-2)	3787	1.896	0.899	0.007	5.111	African cities
rain(t-3)	3246	1.896	0.900	0.007	5.111	African cities
rain(t-4)	2705	1.943	0.921	0.007	5.111	African cities
rain(t+1)	4328	1.893	0.894	0.007	5.111	African cities
primate city dummy (political)	4869	0.035	0.184	0	1	African cities
primate city dummy (population > 200,000)	4869	0.054	0.225	0	1	African cities

Table A2. African countries with city population data

Country	census year 1	census year 2	census year 3	Unit	Pop. cutoff	WUP 2007 cutoff	# city points	# city lights	# lights
Benin	1992	2002		urban localities	10,000	10,000	64	29	56
Burkina Faso	1985	1996	2006	urban localities	10,000	10,000	44	38	58
Botswana	1991	2001		towns	10,000	5,000	27	21	128
Central African Republic	1988	2003		cities	5,000	3,000	37	14	27
Ghana	1984	2000		urban localities	20,000	5,000	69	34	256
Guinea	1983	1996		urban areas	10,000		27	23	66
Kenya	1989	1999		towns	10,000	2,000	62	47	220
Mozambique	1980	1997	2007	principal cities	20,000		34	32	136
Mauritania	1988	2000		communes	15,000	5,000	25	16	33
Malawi	1987	1998		towns	10,000		19	19	87
Namibia	1991	2001		towns	10,000		19	16	190
Niger	1988	2001		urban centers	10,000	2,500	36	31	135
Rwanda	1991	2002		principal municipalities	none		15	12	13
Senegal	1988	2002		urban communes	10,000	10,000	51	38	143
Tanzania	1988	2002		urban localities	10,000		104	74	255
Uganda	1991	2002		towns	10,000	2,000	60	39	67
Zambia	1990	2000		localities	10,000	5,000	37	30	135
Zimbabwe	1992	2002		towns	10,000	2,500	37	28	318
Subtotal							**767**	**541**	**2,323**
All other African countries									6,866
Africa Total									**9,189**

Robinson projection

Figure 1: Lights at night, 2003

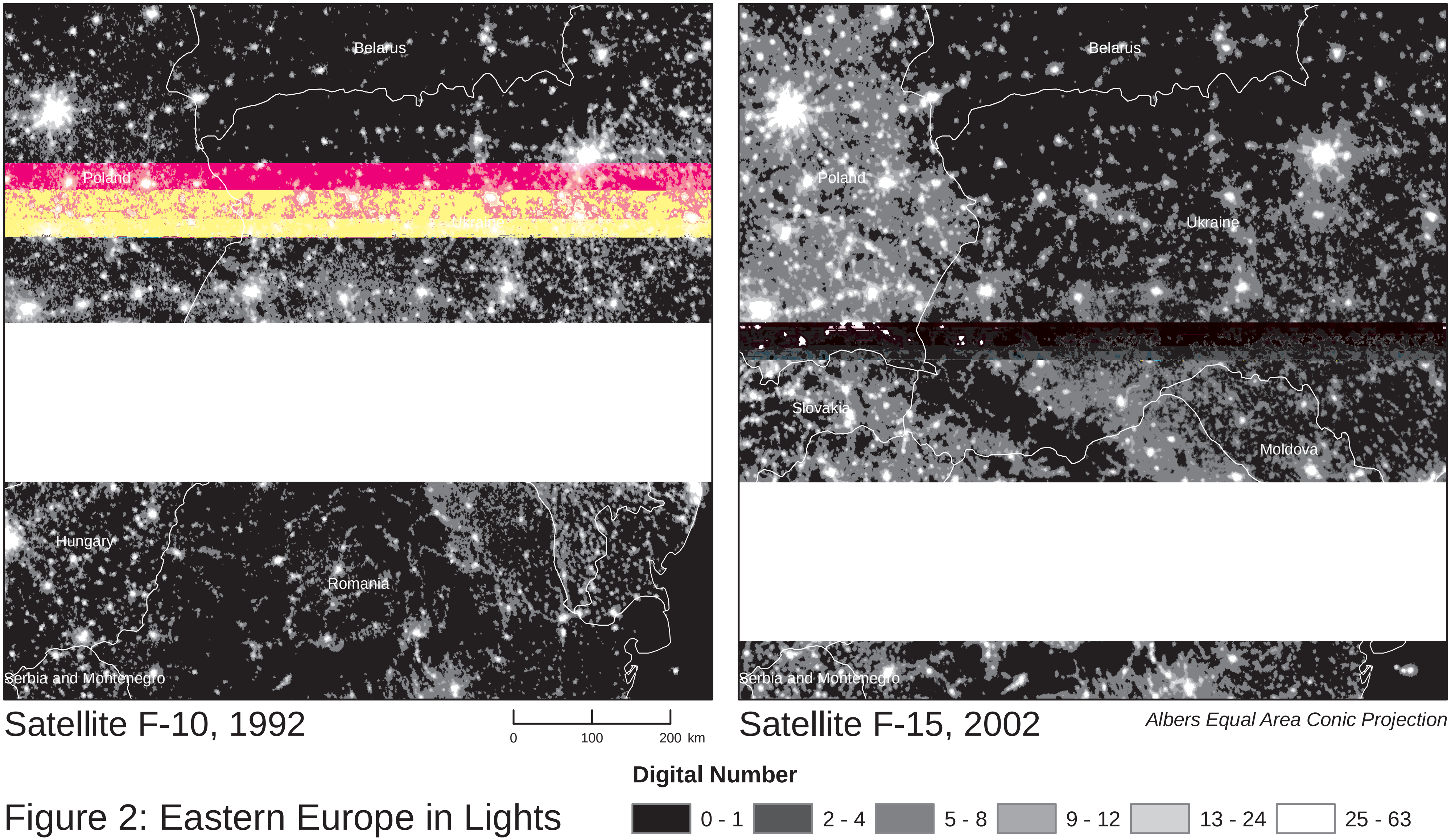


Figure 2: Eastern Europe in Lights

Figure 3: Discovery of sapphire and ruby deposits in Madagascar

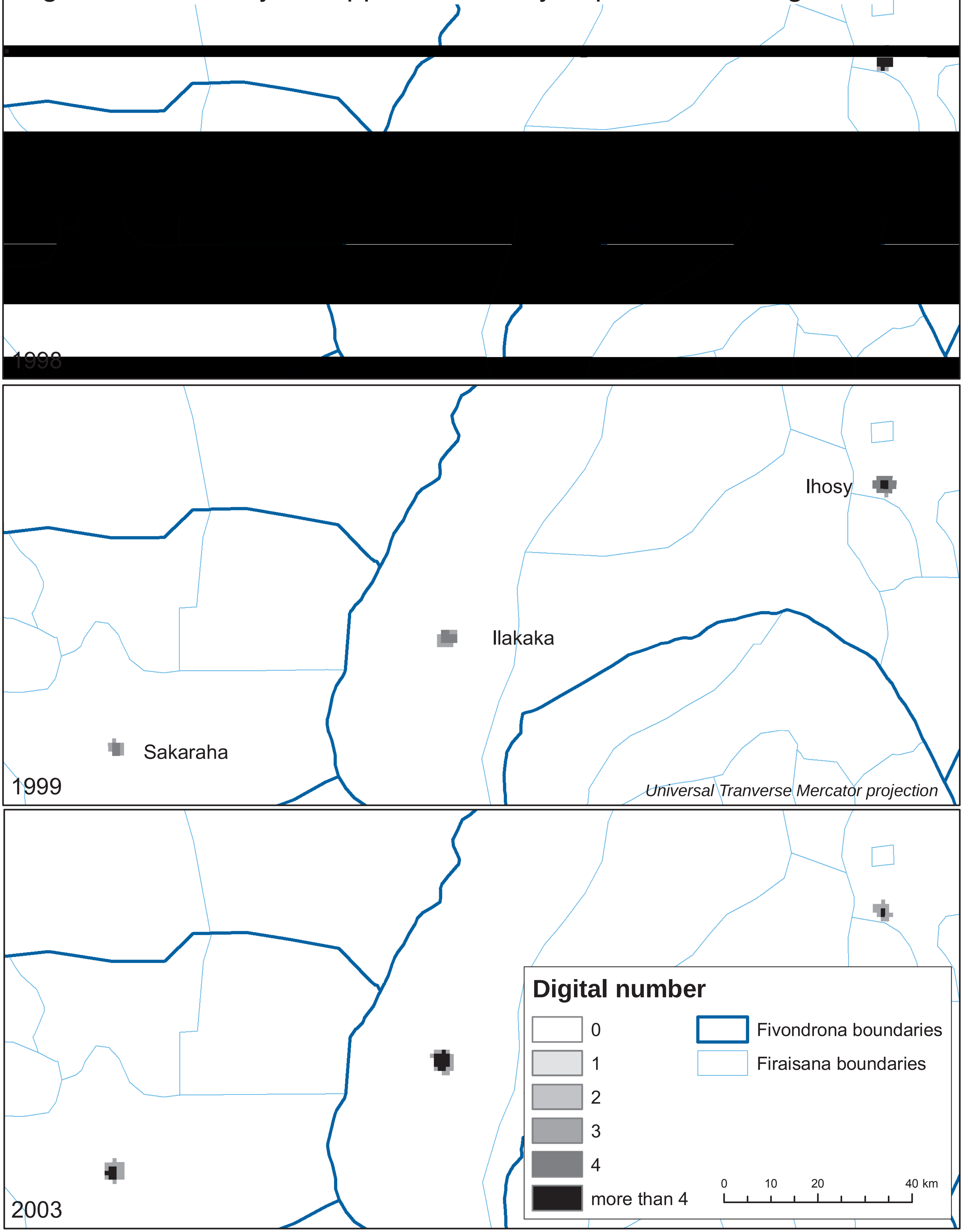
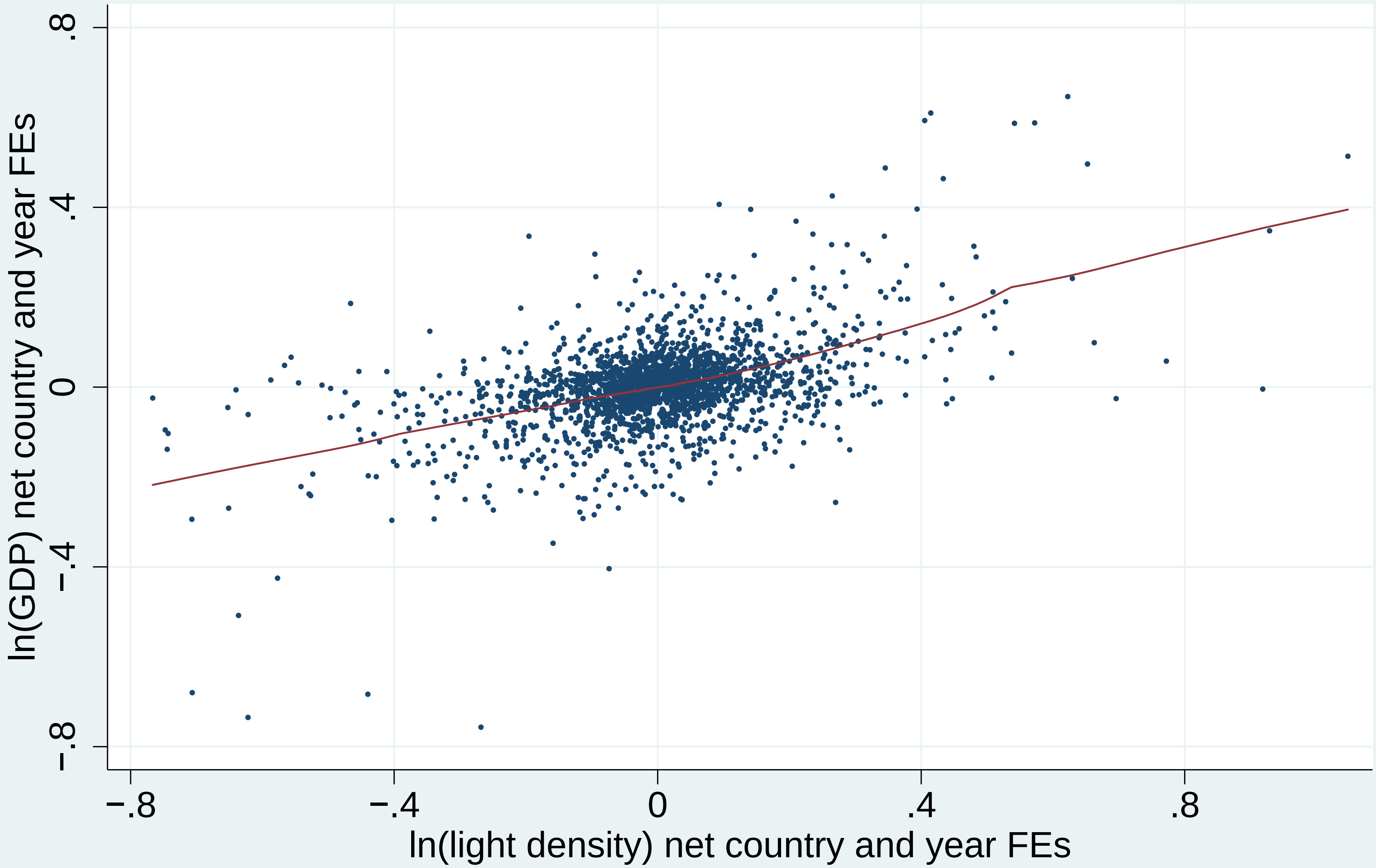


Figure 4. GDP versus lights: panel



note: excludes 3 points to the left

Figure 5: GDP versus lights: long differences

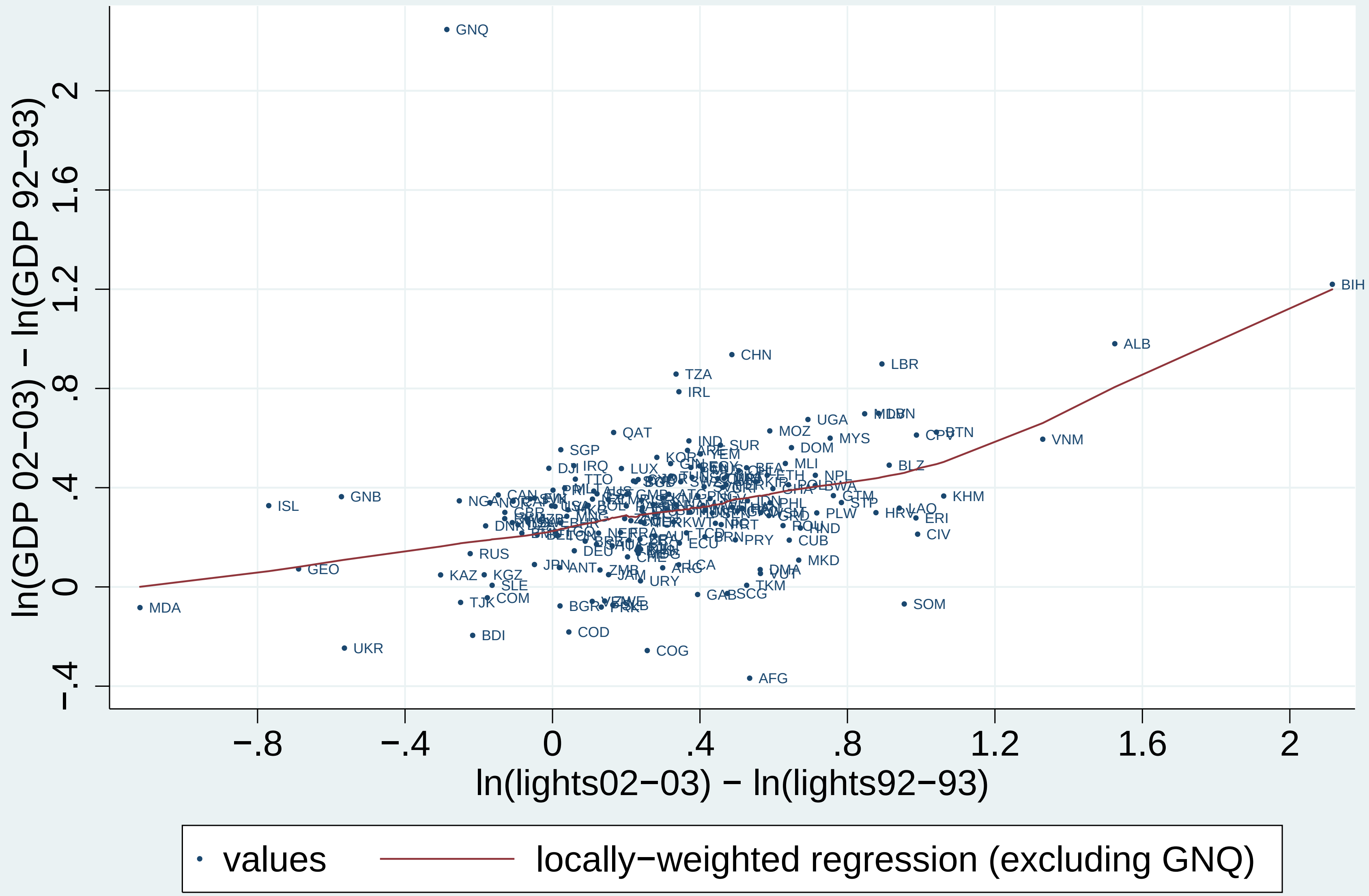
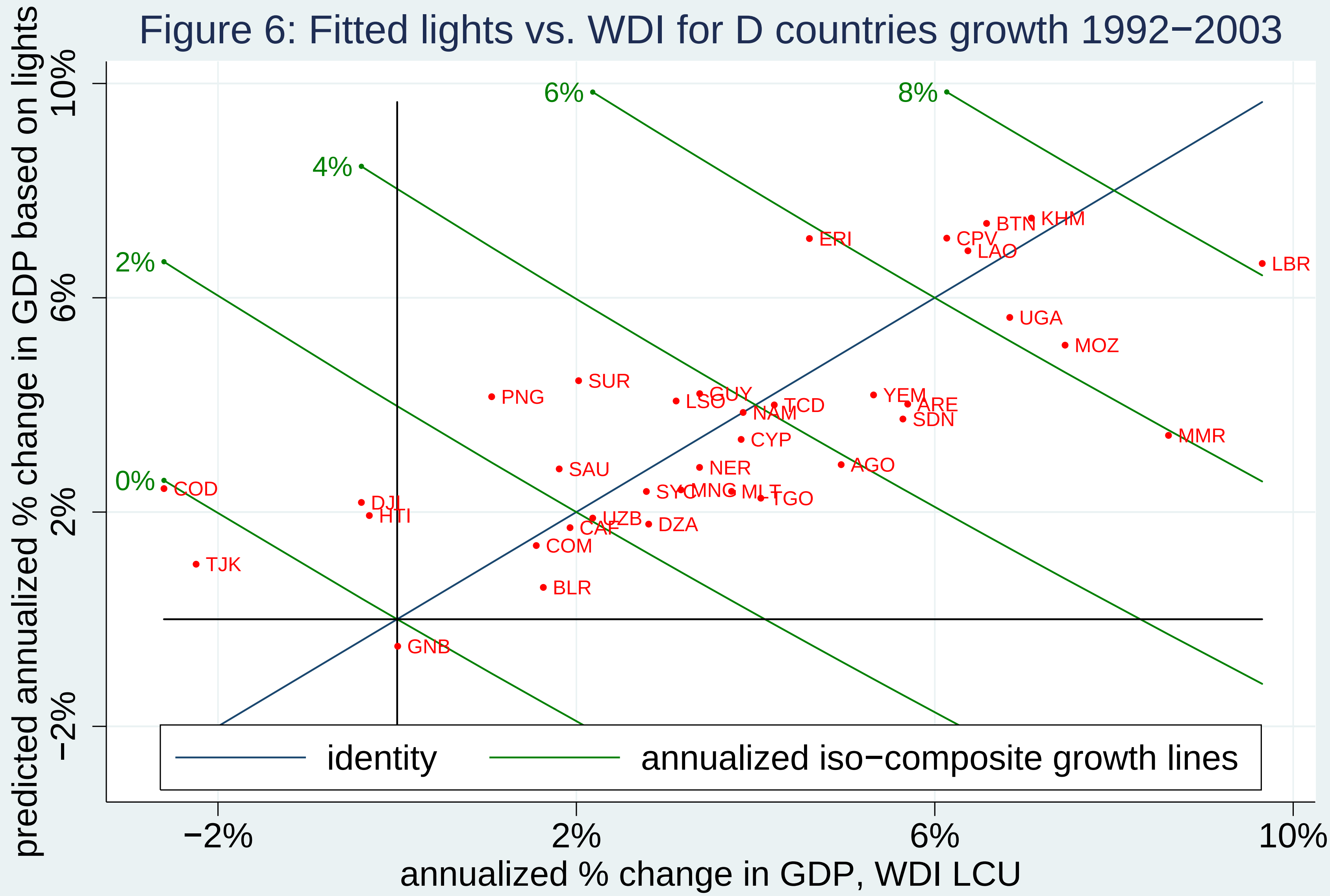


Figure 6: Fitted lights vs. WDI for D countries growth 1992−2003



lambda=.4924, phi=.75

Figure 7a: Outlines of city lights for selected years

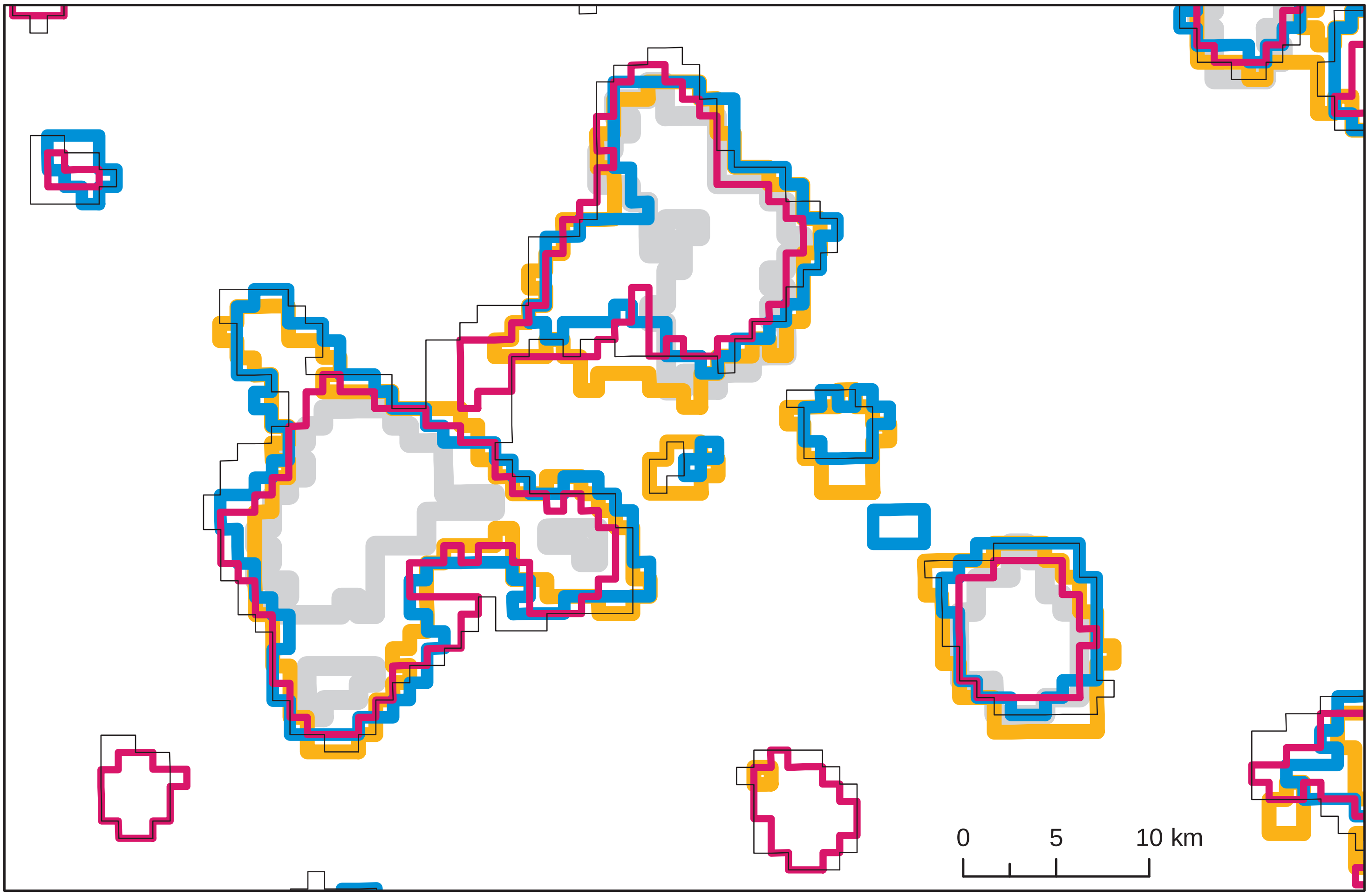


Figure 7b: Light outlines with outer envelope and city point populations