Measuring Economic Growth from Outer Space

Vernon Henderson, Adam Storeygard and David Weil[*]
Brown University

September 24, 2008

Abstract

We propose a simple model that allows us to consider satellite data on lights at night as a proxy for local economic activity, distinguishing between population density and consumption per person. The model is tested on a 12-year panel of countries, as well as a panel of subnational administrative units in Guatemala. To our knowledge this is the first global analysis using panel night lights data (1992-2003). We find that changes in lights explain a large portion of the variation in changes in consumption. We then apply the lights data to a developing country debate about whether growth in local agricultural activity in a city's hinterland itself spurs city growth, as opposed either to being irrelevant or to there being a one way street where local agriculture growth is dependent on city growth. We find for African cities that exogenous productivity shocks in agriculture as represented by years of high rainfall have a significant and substantial effect on the level of local urban economic activity..

0. Introduction

Economists studying macroeconomics and growth generally focus on Gross Domestic Product (GDP) as the dependent variable in their analyses. The conceptual problems in defining GDP, much less using it as a measure of welfare, are the stuff of introductory economics courses. Just as serious, however, is the problem that GDP is terribly measured. For example, in the United States, the standard deviation of the gap between the advance estimate of real quarterly GDP growth (which is available one month after the end of the quarter) and the latest revised estimate (recalculated every July for three years, and every five years thereafter) is 1.0% for 1983-2002. This is a substantial portion of the standard deviation of measured growth, 2.4% over the same period (BEA 2006, 2008).

The measurement problems in GDP are far more serious in developing countries, for several reasons. Compared to developed countries, a much smaller fraction of economic activity in developing countries is conducted within the formal sector, the degree of economic integration among regions is low, and the government statistical infrastructure is often quite weak. Comparing real GDP among countries requires not only the compilation of nominal GDP (total

[*] We thank Chris Elvidge for advice and auxiliary data, Andrew Foster and participants at the 2008 BREAD/CEPR/Verona Summer School on Development Economics for comments, and Joshua Wilde for research assistance.

value added, in domestic prices), but also information on prices that can be used to construct purchasing power parity (PPP) exchange rates. The difficulty of measuring GDP in developing countries was illustrated in 2005 when the World Bank carried out new surveys of prices in a large number of countries (the previous round had been in 1993, although many countries had not participated). The values of PPP GDP in both China and India were reduced by approximately 40% in the new calculation. In the Penn World Tables (PWT), for one of the standard compilations of cross-country data on income, countries are given grades corresponding to data quality, with a grade of A indicating a margin of error of 10%, B indicating 20%, C indicating 30%, and D indicating 40%. Almost all industrial countries receive a grade of A. By contrast, for the 43 countries sub-Saharan Africa (Madagascar and the mainland countries that don't touch the Mediterranean), 17 get a D and 26 a C. Measurement error in GDP data can easily lead researchers to erroneous conclusions. 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. 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 version of the Penn World Tables. Finally, reliable data on output at the sub national level are regularly available for very few countries, and most of those are highly developed.

In response to the problems of measuring GDP, there is a long tradition in economics of looking to various proxies that are more precisely measured, cover periods or regions for which GDP data are not available, or are available more quickly than standard GDP data. For example, up until the year 2005, the Federal Reserve Board based its monthly index of industrial production in part on a survey of utilities which 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 in inequality in the Americas over the last two millennia.

In this paper we explore the usefulness of a new 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 growth.

One reason for looking at data on change in night light intensity is simply as an additional measure of output growth. Even if changes in light from space are also subject to significant 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. The night lights data have several advantages, however. The data from satellites are available at a much higher time frequency than standard output measures. Although measurement considerations would make it unreasonable to look at frequencies as short as days or weeks, the satellite data allow for measurements of seasonal patterns of activity that would be unobtainable in most countries.

Most significantly, the 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 down to a unit smaller than a one-kilometer square, which can be matched with map data to aggregate to the town, city, or regional level. This makes the data uniquely suited to spatial analyses of economic activity. Economic analysis of growth and the impacts of policies and events on cities and regions of many countries is hindered by a complete absence of any regular measure of the level of local economic activity. While population data are typically regularly available for cities above a certain size, almost no countries have city level GDP[1] data. Night lights data give us such a measure and we will analyze applications and methodology in the paper. Besides looking at total economic activity, by decomposing economic activity into population density and per capita consumption components, we can potentially back out the rate of per person consumption growth and the impact of policies on living standards at the sub-national level.

To illustrate an application of lights data to an economic problem, we turn to the debate about the extent to which productivity in the agricultural hinterland of a city affects city economic growth. Urban economists tend to model cities as either divorced from their hinterland

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

(e.g., Black and Henderson, 1999) or as source of demand for local agricultural crops (von Thunen, 1826 and Nerlove and Sadka, 1987). 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 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 city growth. The idea that agriculture can spur urban growth is hard to test because of lack of detailed 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. Whose growth spurs whom? In this paper, we make use of a "natural" experiment using night lights data on 541 African cities to examine the extent to which productivity gains in local agriculture engendered by rainfall increases affect city economic activity.

The rest of this paper is organized as follows. Section 1 gives a brief introduction to the night lights data and discusses more obvious examples of how they represent differences in income levels across countries and the effects of shocks on growth or income levels. In section 2 we estimate baseline models where, first, lights are a measure of national GDP and then where national GDP is decomposed into per person GDP and population density components. Later in Section 2 we turn to estimating models at the sub-national administrative level, to see how local per person consumption growth could be estimated from a combination of night lights and local population data. Finally in section 3 we turn to an application for a large sample of African cities, where we estimate the impact of agricultural productivity shocks on urban growth.

1. Night lights data

A US Air Force weather satellite circles the globe 14 times per day as the earth rotates, recording the intensity of earth-based lights. The satellite records on a swath wide enough so that it covers virtually the entire earth between about 8:30 and 10pm (depending on location), when it is night time but people are typically still active, at least once every 24 hour time period. 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. The number of valid nights of data per year for an area varies by location and year depending on evenings of cloud cover, and seasons in high-latitude vs. equatorial places. Readings affected by auroral activity (the northern and southern lights) and forest fires are also removed, manually and using frequency filters. Our measure of intensity of lights is a six-bit (0-63) digital number calculated for every 30-second output pixel (approximately 0.86 square kilometers at the equator)[2] , which is averaged across overlapping raw input pixels and all valid evenings in a year. The values are not a direct measure 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 the same information. The recalibrated data, which we use throughout the paper, is on a scale of 0-65.

We believe intensity of night lights is a reflection of intensity of consumption, as reflected in the per person indoor and outdoor use of lights. Consumption of all goods in the evening requires lights. As income rises, so does light usage per person. Obviously there are issues about what types of public lighting versus private lighting go into the intensity measure. Lights are just a proxy for consumption, but we argue that, if no income data are available, we can back out a reasonable estimate of income using lights data. The advantage of lights data as noted above 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 income and population density. 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 percent is 90, while in the Netherlands it is under 1. The percentage of unlit pixels unlit 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, 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 low fraction of pixels with digital numbers of 1 or 2

[2] 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. All densities are calculated taking this into account.

reflects, we think, the effect of software designed to filter out noise in the sensor. This censoring of data will be one of the issues we discuss further in the analysis below.

Related, the last two rows of Table 1 show the mean digital number, the within country standard deviation of the digital number across pixels and the within country Gini for the digital number. The mean ranges from 22 in the Netherlands to 0.03 in Madagascar. In the Table and in the data more generally, both the standard deviation and the Gini rise with income, the latter directly reflecting some divergence in spatial equality of lights with economic growth. Below in the empirical work we will explore the extent to which these dispersion measures additionally contribute to our ability to predict income growth using changes in mean digital lights.

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.[3] 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 consumption in an area, which is increasing in both consumption per person and number of people. In the United States, where living standards are fairly uniform nationally, the higher concentration of lights in ocean and inland coastal areas reflects the higher population densities along those coasts. The comparison of lights in Western Europe and India reflects huge differences in per capita income, as does the comparison between sub-Saharan Africa and the low density inhabited parts of Canada. We focus first on using changes in observed lights as a measure of total local economic growth, and then later on decomposing that growth into population and per capita income components.

Eastern Europe and the former Soviet Republics over time

To see mostly pure income effects, we examine the differential effects of transition on income and lights in the former Soviet republics. We compare the former Soviet Republics of Moldova and Ukraine, where per capita PPP-adjusted (WDI) income fell by over 30% from 1992 to 2002, with their Eastern European counterparts of Hungary, Poland and Romania, who went

---

[3] 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 subnational 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.

through a much smoother transition process with incomes rising by 23 – 56 % in the same time period.

Unfortunately our satellite data only start two years into transition; nevertheless the differences in lights are dramatic (Elvidge *et al* 2005). In Figure 2 the more brightly lit areas in 2002 are 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, corresponding to the 30 and 35% drops in income per capita, along with modest population drops (3 and 8% respectively), light intensity drops 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

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 trading centers for sapphires. Previously little more than a truck stop, the population of Ilakaka is now estimated at roughly 20,000 (Hamilton 2003, Hogg 2007). 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 is 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 towns visible in the figure, Tuléar and Ihosy, show no such growth. If anything, Ihosy's light gets smaller and weaker, as it suffers in the competition across cities for population.

2. Lights as a measure of economic activity

In this section, first we analyze lights as a measure of local economic activity and then analyze the decomposition of economic activity into consumption per person and population at the national and sub-national level.

2.1 Lights and total economic activity

We start with a simple specification of the relationship between lights and consumption, which we then test and refine, depending on circumstances. We hypothesize that

$$light\,/\,area = \phi\,(consumption/area),\ \phi' > 0\,. \qquad (1)$$

Lights in an area are an increasing function of total consumption in the area. The latter is increasing in both the number of people and their per capita consumption, which effects we will later allow to differ. *A priori*, it is not clear what the curvature of the $\phi(\cdot)$ function should be. There could be some diminution in the rate of increase of light as the number of people in an area increases. First, higher population density implies a greater likelihood of people living above one another, so that some light is blocked from reaching space. Second, there could be economies of scale in the use of lights, such as street lamps. On the other hand, since there are strong economies of scale in electricity distribution. Dense areas are more likely, holding income constant, to have electricity which is key determinant of night lights; and, once there is electricity, household amperage could increase as costs fall. Of course the shape of relationship will be affected by the nature of the sensors used. For example, when lights are too faint the sensor registers no light at all; and the general relationship between "true luminance" and recorded digital numbers may be non-proportional. Finally, data used in estimation involves aggregating over pixels as noted earlier, with different distributions of light intensity across pixels for the same mean light in an area. We are agnostic about the shape of $\phi(\cdot)$ and the use of specific controls for dispersion of lights across pixels within the same area.

If, as a local approximation, we restrict $\phi(\cdot)$ to be a power function, then *light / area* = $\tilde{\beta}\,(consumption/area)^{\tilde{\alpha}}$. Taking logs and re-arranging,

$$ln(consumption\,/\,area) = \beta + 1/\tilde{\alpha}\ (lights\,/\,area)\,. \qquad (1)$$

We begin by using this specification to analyze data from a panel of countries. In practice, we use data on GDP, rather than consumption, since the two are highly correlated (and it is not clear that the non-consumption parts of GDP should produce observable light in a fashion different than consumption). We also allow for a full set of country and time fixed effects. Country fixed effects will allow for differences in climate and other local factors that may affect the relationship between consumption and observable light. They also allow for non-uniformity of consumption measures across countries (i.e. the inability of World Bank income measures to account properly for differences in purchasing power parity). Finally, use of country fixed effects also can account for the fact that, in the lights data, different countries have different land masses which involves averaging lights across sub-national areas with very different population densities and extent of uninhabited sections of the country (Japan versus Canada). As long as such variations don't impact the α coefficient, fixed effects are appropriate. A basic question is the extent to which fixed effects adequately control for differences in dispersion patterns of lights across countries, in using lights to predict consumption. Time fixed effects help control for any regime changes in the way purchasing power incomes are measured worldwide, and again help control for issues 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.

The specification is thus

$$ln(GDP)_{it} = \kappa_i + \delta_t + \alpha\ ln(lights\ /\ area)_{it} + \varepsilon_{it,} \qquad (1a)$$

where i indexes country, t indexes year, and $\alpha = 1/\tilde{\alpha}$. Our measure of GDP is PPP total national income is taken from the World Development Indicators (WDI) online. The lights data are collected by US Air Force weather satellites, and 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.

Basic Results

Table 2 presents some basic results for a (modestly unbalanced) panel of 166 countries over 12 years, where we drop Equatorial Guinea as an outlier (see below). Column 1 is the specification in (1) without country fixed effects, while column 2 adds in country fixed effects. Column 2 suggests a smaller elasticity to the lights- consumption relationship than the simple pooled specification. In terms of explanatory power, the pooled R^2 in column 1 is .13; the within R^2 in column 2 is very high at .69; and the total R^2 treating country dummies as nuisance parameters is over .99. Column 3 suggests a quadratic specification is not appropriate. Figure 4, looking non-parametrically at the $\ln(lights/area)\text{-}\ln GDP$ relationship (after factoring out time and country fixed effects) suggests a log-linear specification does well.

In column 4 we look at robustness of the relationship to long differences, averaging the first and last 2 years of data. The elasticity is noticeably higher; but we note that any difference between fixed effect and long difference specifications goes away once we decompose GDP into population and per person consumption components in the next section. Figure 5 plots the long difference data points for 167 countries adding back in Equatorial Guinea. The figure shows why the linear approximation in Table 2 does so well; it also illustrates why we dropped Equatorial Guinea[4].

Column 5 estimates the relationship for Africa, a continent of particular interest given poor GDP measures over time. In the last section we also focus on African cities, so having a baseline lights-GDP relationship to refer back to will be useful. The lights-GDP elasticity and the within R^2 are both higher for Africa alone. If we estimate the elasticity for low and lower-middle income countries (as defined by the World Bank for the base year 1992) we also see a rise in the elasticity from .23 in Table 2 to .27, while for upper-middle and upper income countries it is only .088. This is an issue we explore more in the next sub-section.

The estimated parameter α (0.23 in column 2) implies that consumption is an increasing, concave function of observed lights and thus that observed lights are a convex function of consumption, where the elasticity of light with respect to economic output is between 4 and 5. This could simply reflect a functional relationship in the sensors between "true" and observed lights. But suppose that relationship is proportional (ignoring the local convexity induced by censoring of low light levels by the recording sensors), *a priori,* there is no reason that the true relationship between output and light could not be convex. Consumption of light may be a luxury good. And as mentioned above, there are large fixed costs to the installation of electricity, so that only areas of sufficient population and income have access to electricity (approximately 75 percent of the world's population is estimated to have access to electricity (IEA 2006)).

Dispersion of lights within a country

This property of the $\phi(\cdot)$ function has implications in aggregation. In the regressions, the measure of light that we use is the digital number per pixel, averaged over all pixels in a country. As mentioned above, for most poor countries, the vast majority of pixels have a value of zero. This suggests the country average will be determined not only by the average level of economic

[4] Equatorial Guinea is dropped as an extreme outlier of high consumption for low lights, particularly in long differencing, given oil discoveries over the 12 years.

activity per unit area in the country, but also by its spatial variation. Some countries like Canada have a high proportion of unlit or low intensity pixels, with a small portion of intensely lit pixels. In some other countries, most pixels are lit and the variation across pixels is more modest. In the estimation in columns 1-5, country fixed effects control for these differences across countries. But the issue seemed important enough to explore, especially since there can be noticeable within country changes in the dispersion of lights over time.

If consumption is a concave function of lights, then for the same average, in principle, greater dispersion should be associated with lower consumption. We experimented with two indices of the dispersion (or bunching) of lights.[5] First in column 6 of Table 2 we add a control for the ln (standard deviation) of lights across pixels by country-year. The sign is positive, rather than negative as could be expected; and adding in the control sharply reduces the elasticity with respect to mean lights. At the same time, the within R^2 is little improved. We tried a translog specification in lights and its standard deviation to the same effect; in general the marginal effect of increasing variation in lights is associated with increased GDP.[6] A basic problem with the standard deviation as a measure of dispersion is that it is not scale invariant. Indeed there is an extremely high correlation between mean lights and its standard deviation: the simple correlation coefficient after factoring out time and country fixed effects is .89, suggesting why including the standard deviation of lights adds little to explanatory power. Then in column 7 we tried the Gini as a measure of dispersion. In column 7, the coefficient on lights is the same as in column 2 and the coefficient on the Gini is 0; the Gini coefficient is also 0 in the long differencing version.

2.2 Decomposing economic activity: population density versus consumption per person

Using light data to estimate total economic activity, as in the previous section, is useful in a number of contexts. However, there are many applications in which economists are interested in measuring income or consumption *per capita*. Since population is more easily measured, is

[5] 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.

[6] The explanation may be simple. As countries develop, they also urbanize, drawing population out of rural areas into dense urban areas. Additionally in this phase counties typically experience rising income inequality. In such contexts, increasing lights are strongly associated with greater dispersion in lights across the country. In the next section when we decompose GDP into population and per person income measures, we explored this issue in great detail looking at how mean lights' effects vary in translog specifications and by quintile of standard deviations, as well as examining other measures controlling for light variation. The basic problem of finding positive effects of variation in lights on consumption remains, as does the very high partial, within country correlation between mean lights and its standard deviation. In the rest of the paper, we do not worry about trying to control separately for variation in lights, as well as its mean.

more available in general, and is much more available for much smaller geographic units than is GDP, this suggests that one could combine data on lights and population to produce estimates of GDP per capita in a number of contexts when it is not measured. In this section we pursue this issue, starting with a discussion of why marginal increases in GDP per capita and population density may have differential effects on the measure of lights observed from space.

Our starting point is a structural relationship between consumption per capita and "true" light output.

$$light / person = \tilde{c}(consumption / person)^{a}, \qquad (2)$$

Multiplying by people and dividing by jurisdictional area, we get

$$light / area = \tilde{c}(consumption / person)^{a}(pop.\ density) \quad (3)$$

In some circumstances, no lights in a jurisdictional area are observed. Then we would define observed lights as[7]

$$obs.light / area = \max[0,\ \tilde{c}(consumption / person)^{a}(pop.\ density) - K]. \quad (3a)$$

For our benchmark case where all countries are lit, we work with equation (3). Later in discussing sub-national areas, we will examine the implications of (3a).

[7] In analyzing light data, as noted earlier there is an issue in measurement of aggregation across geographic units. Our readings are for pixels and when we give measured lights per jurisdictional area it involves averaging across pixels. Total lights for a jurisdiction area should reflect the following aggregation across pixels (*j*):

$$light / area = \sum_{j} [\tilde{b}(consumption / person)_{j}^{a}\ (pop.\ density_{j})]$$

If through labor mobility within a jurisdiction, consumption per person is approximately equal then what varies across pixels is population. How uneven population densities (and hence light intensities) are within the jurisdictional unit doesn't matter if the exponent on population density in the equation is one, but estimates suggest it is very different than one. As noted earlier in the text, attempts to control for this problem by controlling for measures of light dispersion, such as the standard deviation of light intensity across pixels in a jurisdictional area were thwarted by the very high correlation between mean and standard deviation of lights within jurisdictions over time.

We arrange (3) to focus on the variable which we want to predict in circumstances where it is not measured: consumption per person.[8] Second, we don't impose a restriction on the power to which population density is raised.

$$ln(consumption/person)_{it} = c_i + d_t + (1/a)ln(lights/area)_{it} - (b/a)ln(pop.\ density) + e_{it} \quad (4)$$

In (4) as in (1a), we allow for time and jurisdictional fixed effects (δ_t and c_i) and we will experiment with different functional forms and controls for light dispersion.

Country level results

As for total GDP estimates, we estimate equation (4) for a panel of countries covering 166 countries over 12 years and for a sub-sample of 50 African countries in that time period. Table 3 presents results for the world. Column 1 pools the data, with time effects (only). Relative to the specification in equation (2) as estimated in (4), coefficients have anticipated signs and the coefficients on lights and population density have similar absolute values, although the lights coefficient absolute value significantly exceeds that for density. However in column 2, once we add in country fixed effects, results change dramatically. Now the density coefficient is much larger in absolute value than the lights one, with the absolute value of the density coefficient being a little over three times the lights coefficient, a relative difference that is at least maintained for all fixed-effect and long difference results.

Column 2 is our preferred baseline to use in comparisons. The coefficients interpreted in equation (2) suggest lights are strongly convex in both consumption and density dimensions. As before this could reflect aspects of the relationship between "true" and observed lights, as well as the possibility of strong increasing returns to lights in both per person consumption and population.

Columns 3 and 4 of Table 3 explore other functional specifications. Column 3 adds an interactive term, which makes the lights coefficient insignificant. Column 4 explores the full translog, with little support for that functional specification. While a translog could be used, it adds little to the within R^2. Figure 6 plots lights against income, after factoring out the effect of

[8] While we are just estimating correlations, we prefer the specification in (4) per se. Non-parametric plots of $ln(consumption/person)_{it}$ on $ln(lights/area)_{it}$ after partialing out fixed effects and density versus plots of $ln(lights/area)_{it}$ on $ln(consumption/person)_{it}$ both suggested an estimate of a consistent with that in (4), as opposed to the estimate we get when estimating (3).

density, and fixed effects. The plot suggests a simple log-linear relationship, which we tend to rely on. In general the explanatory power of the model is high. The pooled model in column 1 has an R^2 of .74 and the within R^2's in later columns are about .40. Finally we note that controls for the two measures of dispersion from Table 2 produced similar results: the standard deviation of light remains highly collinear with mean lights and the Gini has an insignificant effect. Fixed effects seem to be a reasonable control for differences in geographic dispersion across countries.

In column 5 of Table 3 we check robustness of fixed effects results identified off annual variations of data within countries by using long differences calculated from the first and last 2 years of data for each country. A comparison of column 5 with 2 shows very stable results with very similar coefficients for both light and density.

Finally in column 6, we look at the sub-sample of 50 African countries over the 12 years. The relative difference between absolute lights and density coefficients is smaller than for the whole world, possibly suggesting heterogeneity of effects. We estimated the model to see if coefficients differ more generally between the set of low and low-middle income countries as defined in the base year (1992) by the World Bank, and the set of upper middle and high income countries. The results suggest that our point estimates on coefficients of covariates in column 2 are the same as those for low and low-middle income countries. Higher income countries have the same proportional difference between coefficients on lights and density, but smaller absolute magnitudes (.145 and -.555 respectively). Once total consumption is decomposed into population and per person consumption, low density Africa differs from other low income regions, with smaller negative effects on consumption for changes in population density holding lights fixed.

2.3 Sub-national estimates of consumption per person

Next we turn to an analysis of sub-national units to estimate equation (4) at finer geographic scale. We focus on Guatemala where we have consumption and population data for two time periods and nearly all jurisdictions are lit.

Data

For Guatemala, harmonized *municipio*-level consumption data for 1994 and 2002, originally estimated by SEGEPLAN (2001, 2005) from census and survey data using the small area estimation methods of Elbers *et al* (2003), are from CIESIN (2005). In 2002 when there are readings from 2 satellites, we pick the one (F-15) which had the best area coverage of Guatemala. Results are not substantially different when the other satellite (F-15) is used.

Results for Guatemalan municipios

In Figure 7a, we show the changes in lights in the sub-national areas of Guatemala from 1994 to 2002. There appears to be a modest increase in light intensity in the north, with more substantial increases over most parts of the south. These light changes seem to go along with little change in population densities except possibly for increased densities around the major urban centers in the south (Figure 7c). Figure 7b seems to suggest overall income growth, particularly movement of hinterland areas in the south out of the lower income brackets. But the figures are hard to read. The comment on Figure 7b is consistent with a decline in spatial inequality; and we note there was a decline in the coefficient of variation in consumption per person across municipios from .52 to .45 in the 8 years. However, the idea that population clustered more near the urban centers is at odds with changes in the overall coefficient in variation for population across the municipios nationally, which drops from 1.99 to 1.84. The question is whether regression analysis of consumption, lights and population inter-relationships can sort out the influences of population and consumption per capita on lights.

In Table 4, we present results for the 1994 and then the 2002 cross-section of municipios in Guatemala. The results in columns 1 and 2 for the 2 different years are close, suggesting we could mechanically predict 2002 municipio consumption levels, with the 1994 data. As with the pooled world data, cross-section results show little gap in absolute values of the light and density coefficients, with the lights coefficient again exceeding that of density. However time differencing in column 3 again dramatically changes estimates. As for the world, now the absolute value of the density coefficient exceeds the lights one by a large amount— for Guatemala by over fivefold. Again this suggests in equation (2) sharp returns to population density in producing lights across municipios.

2.4 The problem of unlit jurisdictions

In Guatemala in 2002 all sub-national jurisdictions are lit; and, in 1994, 89% are lit. In estimation in Table 4, in both the overtime comparison and the 1994 cross-section, we discarded the unlit jurisdictions. Is that inappropriate? To try to look at this question we turned to an area of the world for which we have data, albeit just cross-section data, where much higher percentages of sub-national jurisdictions are unlit. These are the African countries of Madagascar, Mozambique, and Malawi, for the years 1993, 1998, and 1997 respectively. Since the data are only cross-sectional, our reporting is brief and highlights just two key findings.

In equation 3a, we indicated that lights are only recorded if light output is sufficient. In the absence of a way to independently define a cut-off point, we model this as a Heckman (1979) selection problem. We pooled the 2019 3rd level administrative regions of the three African countries of which 22% are lit. Then we jointly estimated by MLE the likelihood of areas not being lit, along with the likelihood for the level of lighting if lit, both as functions of population density and per person consumption (and country dummies). We note the error drawings on observed lights and the event of whether lights are observed or not are somewhat different. For lit areas, the difference between observed and "true" lights may vary by satellite and equipment, as discussed above. Whether lights are "observed" or not involves an additional element: low readings are censored by filtering as noted in Section 1.

We note two results from the MLE estimation of the selection model for the administrative areas of the three African countries.[9] First, modeling selection itself is not important. In the MLE results, a Hausman test could not reject equality of the coefficients on a lights equation (relevant version of (3)) between the MLE results and the OLS results on the sub-sample of lit areas; in fact the coefficients were virtually identical. Second, the correlation coefficient between the error terms while positive is not significant at the 5% level (t-stat. = 1.77). Thus a tentative conclusion is that one can model outcomes for just lit areas without worrying about selection. The second result is that, at least in cross-sectional work, if one wants to model whether an area is lit or not, the simple Probit model predicts well, with both consumption and population density effects being highly significant.[10] For example, in Mozambique about 26% of jurisdictions are lit. At minus 2 standard deviations of the two covariates, the probability of a jurisdiction being lit is 0.43% while at plus 2 standard deviations, it is 83%.

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 in both theory (e.g., Black and Henderson, 1999, with Henderson

[9] Estimates for Madagascar (Mistiaen *et al* 2002), Malawi (Benson *et al* 2002) and Mozambique (Simler and Nhate 2005) use the same method and source as Guatemala, but only one year of data per country is available: 1993, 1998, and 1997, respectively.

[10] Marginal effects of ln (consumption per person) and ln (population density) are respectively .164 and ..084 (with s.e.'s of .0243 and .0066).

and Wang 2005 and Brueckner 1990 as exceptions) 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 many of these models take aggregative approaches with the idea that a poor rural sector is 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) using data on Brazilian city population growth find that higher rural incomes in city hinterlands also retard city population growth.

What these approaches generally miss however 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 pushed by agricultural economists, as well as a few growth economists (e.g., Kuznets, 1964; 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 in a city hinterland with increased incomes 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 for that city. In this section we examine a panel of 541 cities in 18 African countries over 9 years. For 14 of the countries this covers all cities for 2008 over 10,000 within 3 km of a night light source, while for the other countries the minimum population size is 5,000- 20,000. 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 consumption data for these cities at all, and we have at most population data for one year for the time period for which we have detailed rainfall. However we have lights for every year. Our presumption is that increased rain increases agricultural income in urban hinterlands. Farmers' spending increases demand for urban goods and thus urban incomes. This rise in urban consumption leads to an increase in lights. We test the net result directly—increased rainfall spurs urban income growth.

The formulation we use is

$$\ln(light_{it}) = \sum_{j=0}^{k} \beta_j rain_{i,t-j} + \alpha_i + \lambda_t + \varepsilon_{it}. \qquad (5)$$

In equation (5) after allowing for city and time fixed effects, this and prior years' rainfall affect current lights. We will find that effects attenuate at $k = 4$; and we will look at the falsification test of adding a lead year of rain. An issue in estimation of (5) 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 weather conditions and urban 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.

In application of equation (5), the impact of agricultural rain may differ according to the urban context. Large industrialized cities may operate more on their own, being more reliant for growth 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 regular cities in the sample and primate cities. 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 how results differ for cities less than versus more than 200,000 people.

While, in interpreting equation (5), we assign lights the role of measuring growth in local urban incomes, there is another channel. Increased urban lights could reflect increased city population (noting we control for city and time fixed effects). We believe that city population effects go against our results, because improved agricultural incomes deter migration from rural hinterlands to cities, as found in Brueckner (1990) and da Mata *et al* (2007).

City Data

We have two sources of data for our Africa cities. First are lights data for pixels covering Africa from 1995-2003. We have no city boundaries, so we define cities in the first pass as contiguous lit areas. Figure 8a illustrates for a hypothetical situation. The 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 8b, we map in jurisdictional cities as points, based on geo-coordinates identified with each city (see data 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 8b) but larger urban areas may consist of several jurisdictions as pictured in the northwest portion of Figure 8b, where 3 cities make up the urban area. The second data source is annual rainfall estimates (Love *et al* 2004), recorded on a 0.1 degree grid (approximately 124 km^2 at the equator). For each lit area we draw a 30 km buffer around the green envelope pictured in Figure 8b to create a catchment area and we measure average rainfall over all grid entries in that catchment area.[11]

Basic results for rainfall effects on urban incomes

Columns 1-5 in Table 5 state the basic results. With just clustered robust errors, columns 1-4 show different lag structures. Column 1 is rain in the contemporaneous year; column 2 allows for 3 years of effects; column 3 for 4 years; and column 4 for 5 years. It is clear rain from 2 years before the present still has a significant effect on urban income. In columns 3 and 4 coefficients for rain from 3 years prior to the current year are weak and in column 4 rain from 4 years prior has an insignificant (negative) coefficient. We generally use a lag structure with 3-4 years of rain, including the current year. In column 5 we re-estimate column 3 imposing an AR[1] process. That process reduces the rain effect of the first year. In columns 6 and 7 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, but it appears to do so for the ordinary panel estimates.

Finally in columns 8 and 9 we look at long differences to test robustness. In column 8, we show the effect on long differences in rainfall on long differences in lights with no lag structure. In column 9, we allow for difference in lights between years t and s to be influenced by differences in rain between t and s, between t-1 and s-1 and between t-2 and s-2. Despite the noise in the data, the effects are very clear. Increases in rain and inferred agricultural productivity lead to increases in light as in the panel.

Rainfall effects are arguably large. A one standard deviation increase in rain (.90 mm/day) in the current or any of the prior two years each leads roughly to a 14.4% increase in lights. From Table 2 a 14.4% increase in lights represents about a 3.3% increase in GDP for a

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

city. A sustained increase in rain over several years in Africa would have a strong effect on urban incomes.

However the effect of hinterland rain on city growth 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 .155 is just .054 for primate cities. For 3 years of rain the coefficients for year t, t-1 and t-2 are .16, .15, and .15 for ordinary cities, while for primate cities they are .085, .075 and .051. For the 29 cities over 200,000 in 1995 versus smaller 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 reduces the net impact of rain on large cities to zero, but we note that estimates on the primate city-rain interaction terms are imprecise.

4. Conclusions

Data appendix:

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 series of annual composites, currently for 1992-2003. 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, of 0.05% of the global total, in 7 countries. Each grid value is an eight-bit integer (0-63), 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.[12] 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. Many of these problems are less in the Africa city examination: less instances of top-coding, no long summer nights, 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. First, all 18 light-years were combined to produce a set of 9189 non-contiguous polygons in which all pixels were lit for at least one year, of which 2323 have centroids falling within the 18 countries with population data. For each, the total calibrated digital number for each satellite-year, as well as the minimum and maximum pixel, were reported.

B. Consumption data

Harmonized *municipio*-level consumption data for 1994 and 2002 for Guatemala, originally estimated by SEGEPLAN (2001, 2005) from census and survey data using the small area

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

estimation methods of Elbers *et al* (2003; hereafter ELL), are from CIESIN (2005).[13] The 1994 and 2002 data rely on censuses in their nominal years but surveys from 1998-99 and 2000, respectively.

The ELL method proceeds as follows. First, as large a set as possible of comparable attributes from a census and a survey for the same population and similar years are selected. In the 2002, this set was supplemented with *municipio*-level averages. Next these attributes are used in a set of reduced form regressions predicting consumption in the survey. In the case of the 1994 Guatemala work, 12 separate regional regressions were run, 4 rural and 8 urban. In 2002, 7 urban, 7 rural, and 1 integrated (Metropolitana province) regression were run. The estimated coefficients are applied to all the census households to generate predicted household consumption. Lastly, the household consumption measures are aggregated to regional averages, with standard errors produced by Monte Carlo simulation.

The regression variables for the 1994 map include electrical connection, type of lighting system and cooking fuel. Elbers *et al* (2005) argue that this does not create endogeneity problems for regressions that use consumption estimates to predict e.g. lighting that are any more significant than those that would be found in regressions of lighting on true consumption. The use of estimates does change the standard error calculation though, an issue that should be considered in future work. Because the errors reported under the ELL method have recently been called into question by Tarozzi and Deaton (2007), this will require careful study.

C. Africa data

City location and population

City population data are from citypopulation.de. Only countries for which information is available from at least two different post-1979 censuses (i.e. not projections or estimates) are used. In addition, I required that at least one census was after 1995. Island states were also dropped. While these figures are not taken directly from the official census bureaus, spot checks suggest that they are consistent with the official figures, where available. Four countries (Algeria, Egypt, Morocco 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 767 cities.

While all countries in this sample have city populations for most cities for at least two censuses, both after 1979 and one after 1995, a minority of cities in these countries have population data for only one year after 1979. A second census as far back as 1969 is used for 61 these cities. In a final 110 cities, only one census is available. For these cities, growth rates are imputed from the cities in that country for which multiple censuses are available.

[13] Available at http://sedac.ciesin.columbia.edu/povmap/

In most countries, all settlements of more than 10,000 are purported to be reported. 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 767 cities have a lower population than their nominal cutoff for any year up to 2008.

Latitudes and longitudes for African cities

These were assigned from three sources: citypopulation.de, the Gridded Rural Urban Mapping Project (CIESIN *et al* 2005), and 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.

Lights and population

Each light in the sampled countries is assigned the population of all cities within 3 kilometers. This reduced the set of city-points from 767 to 656. 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) and CIESIN *et al*. (2005). 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. Of the 2323 lights, in the 18 countries, 541, or 23% contain at least one city for which we have population. However, the others are far less bright and/or extensive lights, consistent with the idea that they correspond to smaller settlements not included in the population data. Thirty-five such lights that contain cities in sampled countries cross a border, at least according to one common set of international boundaries. Of these, seven contain cities on both sides of the border.[14] 111 city points were dropped in this process because they were not near enough to a sampled light. Only one of these has a population over 25,000, and it would have fallen within a light whose centroid fell in another country if such matches were allowed. 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 matches were allowed.

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 stations 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.

[14] 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.

References

Au, Chun-Chung and J. Vernon Henderson (2006), “Are Chinese Cities Too Small?”, *Review of Economic Studies*, Vol. 73, No. 3, pp. 549-576, July 2006

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 Global Estimation of Urban Extents," CIESIN, Columbia University working paper, December 2004.

Becker, Charles M. & Morrison, Andrew R., 1999."Urbanization in transforming economies," in: P. C. Cheshire & E. S. Mills (ed.), *Handbook of Regional and Urban Economics*, edition 1, volume 3, chapter 43, pages 1673-1790 Elsevier.

Benson T with Kaphuka J, Kanyanda S, and Chinula R. (2002). *Malawi: An Atlas of Social Statistics*. Zomba: National Statistical Office, Government of Malawi and International Food Policy Research Institute (IFPRI), pp. 48.

Black, Duncan and J. Vernon Henderson (1999) "The Theory of Urban Growth," *Journal of Political Economy*, 107, 252-284.

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

Bureau of Economic Analysis (BEA), (2006). News Release BEA 06-33. Accessed 15 September 2008 at http://www.bea.gov/bea/newsrelarchive/2006/gdp206a.pdf

Bureau of Economic Analysis (BEA), (2008). Interactive Tables. Accessed 15 September 2008 at http://www.bea.gov/interactive.htm

Center for International Earth Science Information Network (CIESIN), Columbia University, 2005. *Small Area Estimates of Poverty and Inequality (SAEPI) database*. Palisades, NY: CIESIN, Columbia University.

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. Available at http://sedac.ciesin.columbia.edu/gpw.

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

da Mata, D., U. Deichmann, J.V. Henderson, S.V. Lall and H.G. Wang (2007). “Determinants of city growth in Brazil”, *Journal of Urban Economics* 62(2): 252-272 (September 2007).

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

Deaton, Angus and Alessandro Tarozzi (2008) "Using Census and Survey Data to Estimate Poverty and Inequality for Small Areas", *Review of Economics and Statistics* (forthcoming).

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, 15 April 2006

Ebener, Steeve, Christopher Murray, Ajay Tandon and Christopher C 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).

Elbers, Chris, Jean O. Lanjouw and Peter Lanjouw (2003). "Micro-Level Estimation of Poverty and Inequality", *Econometrica*, 71(1): 355-364 (Jan., 2003).

Elvidge, C.D., Baugh, K.E., Kihn, E.A., Kroehl, H.W, Davis, E.R, 1997, "Mapping of city lights using DMSP Operational Linescan System data", *Photogrammetric Engineering and Remote Sensing*, v. 63, p. 727-734.

Elvidge,C.D., Baugh, K.E., Dietz, J.B., Bland, T., Sutton, P.C., Kroehl, H.W. 1999. "Radiance Calibration of DMSP-OLS Low-light Imaging Data of Human Settlements", *Remote Sensing of Environment* 68(1), pp. 77-88.

Elvidge, C.D., V.R. Hobson, I.L. Nelson, J.M. Safran, B.T. Tuttle, K.E. Baugh, and J.B. Dietz, 2002: "Global observation of urban areas based on nocturnal lighting:, *The Land Use and Land Cover Change Newsletter of the LUCC project of the International Geosphere Biosphere Programme and the International Human Dimensions Programme,* December 2002 issue, pp. 10-12.

Elvidge, C.D., Hobson, V.R., Nelson, I.L., Safran, J.M., Tuttle, B.T., Dietz, J.B., Baugh, K.E., 2003, "Overview of DMSP-OLS and scope of applications", in *Remotely Sensed Cities,* Victor Mesev (editor), Taylor and Francis, London, Chapter 13, 281-299.

Elvidge, C.D., K.E. Baugh, J. Safran, B.T. and E.H.Erwin, 2005. "Preliminary Results From Nighttime Lights Change Detection", Proceedings of the ISPRS joint conference 3rd International Symposium Remote Sensing and Data Fusion Over Urban Areas (URBAN 2005)and 5th International Symposium Remote Sensing of Urban Areas (URS 2005) Tempe, AZ, USA, March 14-16 2005. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36(8).

Glaeser, Edward L & Hedi D. Kallal & Jose A. Scheinkman & Andrei Shleifer, 1992. "Growth in Cities," *Journal of Political Economy*, University of Chicago Press, vol. 100(6), pages 1126-52, December.

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

Good, D. F (1994). "The Economic Lag of Central and Eastern Europe: Income Estimates for the Habsburg Successor States, 1870-1910", *The Journal of Economic History*, vol. 54, no. 4, p. 869-891

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

Harris, J. R. and Todaro, M. P. (1970). "Migration, unemployment and development: a two-sector analysis", *American Economic Review*, 60, 126-142.

Heckman, J. J. (1979) Sample Selection Bias as a Specification Error," *Econometrica*, (February 1979), 47(1), 153-161.

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

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

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, May 2006.

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):143-165

Kögel, Tomas and Alexia Prskawetz (2001). "Agricultural Productivity Growth and Escape from the Malthusian Trap", *Journal of Economic Growth* 6(4): 337-357 (December, 2001)

Krugman, Paul (1991). "Increasing returns and economic geography," *Journal of Political Economy*, 99: 483-499.

Kuznets, Simon (1955) "Economic Growth and Economic Inequality," *American Economic Review* 45: 1-28.

Lewis,W. A. (1954). “Economic Development with Unlimited Supplies of Labor”, *The Manchester School*, Vol. 22, pp. 139–191.

Lieske, R.W., 1981. “DMSP primary sensor data acquisition”, Proceedings of the International Telemetering Conference, 17:1013-1020.

Love, T.B., V. Kumar, P. Xie, and W. Thiaw ,2004: “A 20-year daily Africa precipitation climatology using satellite and gauge data”, 2004 AMS Conference on Applied Climatology, American Meteorological Society.

Maccini, Sharon and Dean Yang (2008) “Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall”, *American Economic Review* forthcoming

Miguel, Edward, Shanker Satyanath and Ernest Sergenti (2004) “Economic Shocks and Civil Conflict: An Instrumental Variables Approach”, *Journal of Political Economy*, 112(4): 725-753 (August 2004).

Mistiaen JA, Özler B, Razafimanantena T, and Razafindravonona J. (2002). “Putting Welfare on the Map in Madagascar”. Africa Region Working Paper Series no.34, World Bank, Washington DC., pp. 37.

Secretaria de Planificación y Programación (SEGEPLAN). 2001. "Mapas de Pobreza: Informe Final"

Secretaria de Planificación y Programación (SEGEPLAN). 2005. "Mapas de Pobreza y Desigualdad de Guatemala"

Simler, K and V Nhate. (2005). “Poverty, Inequality, and Geographic Targeting: Evidence from Small-Area Estimates in Mozambique. Food Consumption and Nutrition Division”, Discussion Paper No. 192. International Food Policy Research Institute, Washington, DC, pp. 43.

Small, C., Pozzi, F. and Elvidge, C.D., 2005. “Spatial analysis of global urban extent from DMSP-OLS nighttime lights”< *Remote Sensing of Environment* v. 96: 277-291.

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

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): 509-527, June 2002.

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, J.H. (1826), *von Thunen's Isolated State*. Tr. Peter Hall. London:Pergamon (First German ed. 1826), 1966.

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

World Bank (2005) *Managing Agricultural Production Risk: Innovations in Developing Countries,* Report No. 32727-GLB (June 2005).

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

Digital Number	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

notes:
1) values of 64 and 65 are possible because of intercalibration
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, total GDP, 1992-2003

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	ln(GDP)	ln(GDP)	ln(GDP)	Δln(GDP)	ln(GDP)	ln(GDP)	ln(GDP)
ln(lights/area)	0.3776***	0.2321***	0.2290***		0.3175***	0.1246***	0.2316***
	[0.0193]	[0.0190]	[0.0204]		[0.0301]	[0.0320]	[0.0195]
$\ln(\text{lights/area})^2$			-0.0022				
			[0.0034]				
Δln(lights/area)				0.3005***			
				[0.0412]			
ln(std.dev.(lights))						0.2070***	
						[0.0585]	
gini(lights)							-0.0069
							[0.1006]
Constant	24.0196***	23.9808***	23.9899***	0.2262***	23.5927***	23.6705***	23.9866***
	[0.1562]	[0.0134]	[0.0178]	[0.0183]	[0.0686]	[0.0877]	[0.0874]
Observations	1984	1984	1984	164	588	1984	1984
Number of countries		166	166		49	166	166
(Within) R-squared	0.133	0.686	0.686	0.299	0.739	0.691	0.686
Overall R-sq		0.132	0.133		0.0141	0.159	0.132
Between R-sq		0.131	0.132		0.0106	0.159	0.13
country fixed effects	no	yes	yes	no	yes	yes	yes
time dummies	yes	yes	yes	no	yes	yes	yes
Region	World	World	World	World	Africa	World	World

Robust standard errors in brackets

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

Note: in column 5, long differences are calculated by averaging the first two and last two years

Table 3: Baseline results for the world, per capita GDP, 1992-2003

	(1)	(2)	(3)	(4)	(5)	(6)
	ln(GDPp.c.)	ln(GDPp.c.)	ln(GDPp.c.)	ln(GDPp.c.)	Δln(GDPp.c.)	ln(GDPp.c.)
ln(lights/area)	0.6241***	0.2217***	0.0863***	0.0394		0.3080***
	[0.0080]	[0.0183]	[0.0323]	[0.0561]		[0.0272]
ln(pop. dens.)	-0.4658***	-0.7902***	-0.7294***	-0.9993***		-0.3811***
	[0.0138]	[0.0570]	[0.0581]	[0.1358]		[0.1449]
ln(lights/area)*ln(pop. dens.)			0.0340***	0.0414***		
			[0.0072]	[0.0122]		
ln(lights/area)^2				-0.0137***		
				[0.0048]		
ln(pop. dens.)^2				0.0322**		
				[0.0152]		
Δln(lights/area)					0.2889***	
					[0.0412]	
Δln(pop. density)					-0.8593***	
					[0.1185]	
Constant	10.3519***	11.5197***	11.2100***	11.7470***	0.2082***	9.3838***
	[0.0699]	[0.2238]	[0.2321]	[0.3237]	[0.0250]	[0.4981]
Observations	1984	1984	1984	1984	164	588
Number of countries		166	166	166		49
(Within) R-squared	0.738	0.467	0.477	0.486	0.365	0.399
Overall R-sq		0.0231	0.0415	0.0385		0.636
Between R-sq		0.0218	0.0397	0.0367		0.64
country fixed effects	no	yes	yes	yes	no	yes
time dummies	yes	yes	yes	yes	no	yes
Region	World	World	World	World	World	Africa

Robust standard errors in brackets

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

Note: in column 5, long differences are calculated by averaging the first two and last two years

Table 4: Baseline results for Guatemala, 1994-2002

	(1)	(2)	(5)
	ln(consumption p.c.)	ln(consumption p.c.)	ln(consumption p.c.)
ln(lights/area)	0.2353***	0.2458***	0.0789***
	[0.0188]	[0.0181]	[0.0275]
ln(pop. density)	-0.1930***	-0.1612***	-0.4291***
	[0.0325]	[0.0332]	[0.1129]
Constant	9.1836***	8.9711***	0.1824***
	[0.1479]	[0.1594]	[0.0473]
Observations	293	330	293
R-squared	0.350	0.366	0.093
Satellite	10	15	10,15
Year	1994	2002	1994-2002

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

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

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	ln(light(t)+1)	ln(light(t)+1)	ln(light(t)+1)	ln(light(t)+1)	ln(light(t)+1)	ln(light(t)+1)	ln(light(t)+1)	ln(light(t)+1)	ln(light(t)+1)
rain(t)	0.1518***	0.1590***	0.2012***	0.1486***	0.1211***	0.1654***	0.2507***	0.4917***	0.0246
	[0.0407]	[0.0431]	[0.0496]	[0.0507]	[0.0370]	[0.0497]	[0.0594]	[0.1517]	[0.1023]
rain(t-1)		0.1505***	0.1603***	0.1833***	0.1561***	0.1853***	0.1789***		0.2307**
		[0.0352]	[0.0445]	[0.0583]	[0.0385]	[0.0527]	[0.0517]		[0.0927]
rain(t-2)		0.1448***	0.1555***	0.1640***	0.1730***	0.1255**	0.1319**		0.3068***
		[0.0395]	[0.0420]	[0.0522]	[0.0374]	[0.0503]	[0.0531]		[0.1013]
rain(t-3)			0.0736*	0.0900*	0.1209***	0.1061**	0.0784*		
			[0.0421]	[0.0491]	[0.0373]	[0.0435]	[0.0436]		
rain(t-4)				-0.0513					
				[0.0426]					
rain(t+1)						-0.0126	0.1076**		
						[0.0436]	[0.0498]		
Constant	4.9791***	4.7365***	4.6785***	4.5744***	4.4732***	4.6808***	4.1621***	0.3595***	0.4230***
	[0.0923]	[0.1735]	[0.2487]	[0.3648]	[0.1317]	[0.2324]	[0.3736]	[0.0597]	[0.0617]
Observations	4869	3787	3246	2705	2705	2164	2705	541	541
Cities	541	541	541	541	541	541	541	541	541
(Within) R-sq	0.0460	0.0555	0.0406	0.0483	0.0373	0.0408	0.0434	0.025	0.032
Overall R-sq	0.0018	0.0180	0.0282	0.0282	0.0306	0.0320	0.0293		
Between R-sq	0.0342	0.0355	0.0402	0.0418	0.0430	0.0430	0.0390		
city fixed effects	yes	yes	yes	yes	yes	yes	yes	no	no
time dummies	yes	yes	yes	yes	yes	yes	yes	no	no
error treatment	robust, cluster on city	robust, cluster on city	robust, cluster on city	robust, cluster on city	AR[1]	AR[1]	robust, cluster on city	robust, long differences (first and last years)	robust, long differences (first and last two years)

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

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

	(1)	(2)	(3)	(4)	(5)
	ln(lights(t)+1)	ln(lights(t)+1)	ln(lights(t)+1)	ln(lights(t)+1)	ln(lights(t)+1)
rain(t)	0.1554***	0.1612***	0.1634***	0.1385***	0.1423***
	[0.0420]	[0.0443]	[0.0454]	[0.0316]	[0.0321]
primate*rain(t)	-0.1014**	-0.0764*	-0.0842**	-0.1362	-0.1415
	[0.0460]	[0.0426]	[0.0421]	[0.1573]	[0.1196]
rain(t-1)		0.1522***	0.1538***	0.1102***	0.1131***
		[0.0361]	[0.0367]	[0.0315]	[0.0319]
primate*rain(t-1)		-0.0772*	-0.0790**	-0.1155	-0.119
		[0.0423]	[0.0371]	[0.1654]	[0.1218]
rain(t-2)		0.1475***	0.1475***	0.1081***	0.1101***
		[0.0406]	[0.0414]	[0.0316]	[0.0320]
primate*rain(t-2)		-0.0965**	-0.0621	-0.1013	-0.0936
		[0.0451]	[0.0456]	[0.1527]	[0.1186]
Constant	4.9785***	4.7401***	4.7406***	5.2171***	5.2157***
	[0.0924]	[0.1722]	[0.1720]	[0.0828]	[0.0827]
Observations	4869	3787	3787	3246	3246
Cities	541	541	541	541	541
(Within) R-sq	0.0461	0.0557	0.0558	0.0365	0.0368
Overall R-sq	0.0111	0.0458	0.0587	0.0811	0.117
Between R-sq	0.0815	0.0775	0.0966	0.119	0.166
primate def'n	political	political	pop>200k	political	pop>200k
error structure	robust, cluster on city	robust, cluster on city	robust, cluster on city	AR[1]	AR[1]
city fixed effects	yes	yes	yes	yes	yes
time dummies	yes	yes	yes	yes	yes

standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table A1: Descriptives

Variable	Obs	Mean	std. dev.	Min	Max	Sample	years
ln(GDP p. c.)	1984	8.411	1.124	6.130	10.823	Global countries	1992-2003
ln(lights/area)	1984	-0.002	2.057	-5.750	4.178	Global countries	1992-2003
ln(pop. dens.)	1984	4.018	1.434	0.342	8.856	Global countries	1992-2003
gini(lights)	1984	0.827	0.210	0.045	1.000	Global countries	1992-2003
ln(std. dev.(lights))	1984	1.427	0.943	-1.269	3.085	Global countries	1992-2003
ln(GDP p. c.)	600	7.438	0.836	6.130	9.729	African countries	1992-2003
ln(lights/area)	600	-1.851	1.719	-5.750	2.702	African countries	1992-2003
ln(pop. dens.)	600	3.565	1.278	0.606	6.419	African countries	1992-2003
Δln(GDPp.c.)	164	0.157	0.208	-0.520	0.828	Global countries	1992-2003
Δln(lights/area)	164	0.265	0.362	-1.119	1.525	Global countries	1992-2003
Δln(pop. density)	164	0.149	0.117	-0.140	0.647	Global countries	1992-2003
ln(consumption)	330	8.361	0.398	7.518	9.713	Guatemalan municipios	1994
ln(lights/area)	330	0.856	1.442	-3.281	4.078	Guatemalan municipios	1994
ln(pop. dens.)	330	5.088	1.084	0.528	8.492	Guatemalan municipios	1994
ln(consumption)	293	8.313	0.466	6.838	9.789	Guatemalan municipios	2002
ln(lights/area)	293	0.226	1.694	-4.761	3.870	Guatemalan municipios	2002
ln(pop. dens.)	293	4.786	1.130	0.191	8.172	Guatemalan municipios	2002
Δln(consumption)	293	0.103	0.300	-0.772	1.012	Guatemalan municipios	1994-2002
Δln(lights/area)	293	0.800	0.715	-0.437	3.683	Guatemalan municipios	1994-2002
Δln(pop. dens.)	293	0.332	0.147	0.028	0.939	Guatemalan municipios	1994-2002
ln(pop. dens.)	1999	3.657	1.761	-3.522	9.924	southern African admin. units	various
lit	1999	0.217	0.412	0.000	1.000	southern African admin. units	various
ln(lights/area)	434	-0.179	2.746	-7.600	4.140	lit southern African admin. units	various
ln(pop. dens.)	434	5.331	1.880	-2.407	9.924	lit southern African admin. units	various
ln(lights+1)	4869	5.548	2.126	0.000	11.426	African cities with population data	1995-2003
rain(t)	4869	1.903	0.904	0.007	5.111	African cities with population data	1995-2003
rain(t-1)	4328	1.886	0.893	0.007	5.111	African cities with population data	1996-2003
rain(t-2)	3787	1.896	0.899	0.007	5.111	African cities with population data	1997-2003
rain(t-3)	3246	1.896	0.900	0.007	5.111	African cities with population data	1998-2003
rain(t-4)	2705	1.943	0.921	0.007	5.111	African cities with population data	1999-2003
rain(t+1)	4328	1.893	0.894	0.007	5.111	African cities with population data	1995-2002
primate	4869	0.035	0.184	0.000	1.000	African cities with population data	1995-2003

Table A2. African countries with city population data

country	census 1	census 2	census 3	unit	population cutoff	World Urbanization Prospects 2007cutoff	number of city points	number of city lights	number of 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**

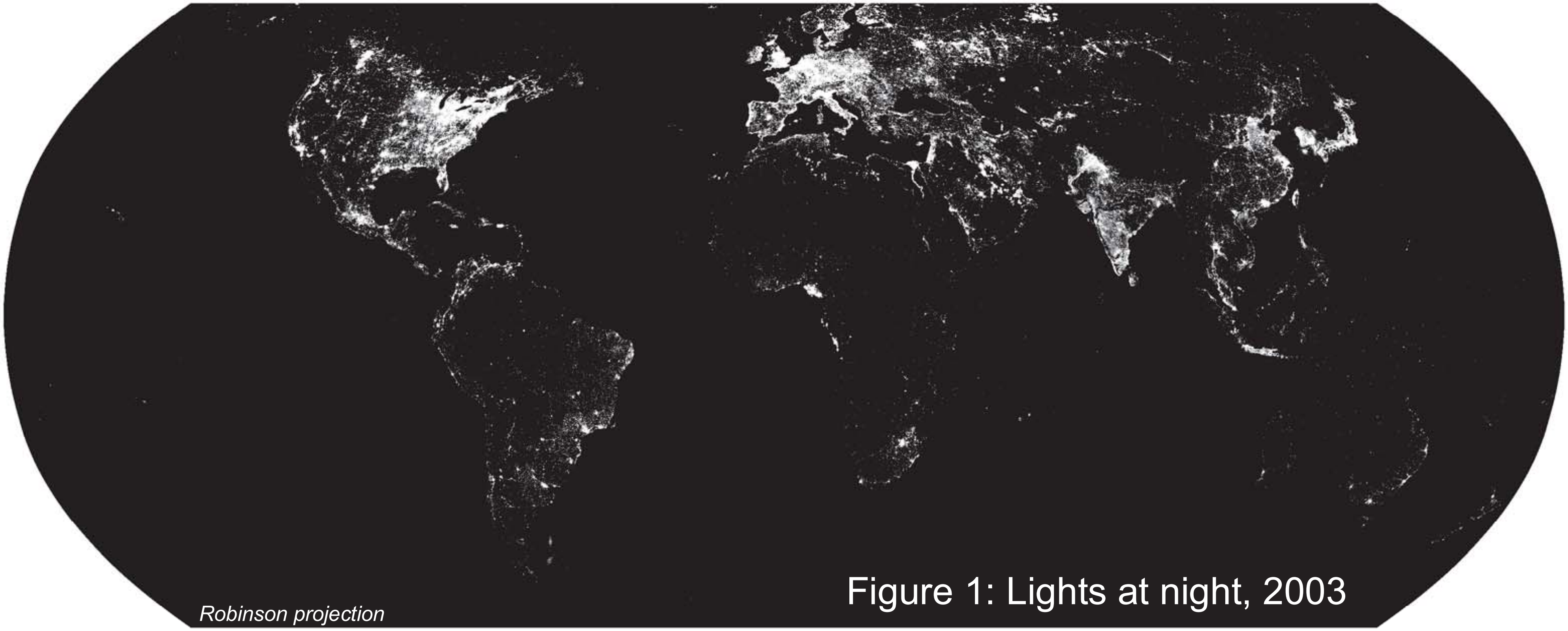


Figure 1: Lights at night, 2003
Robinson projection

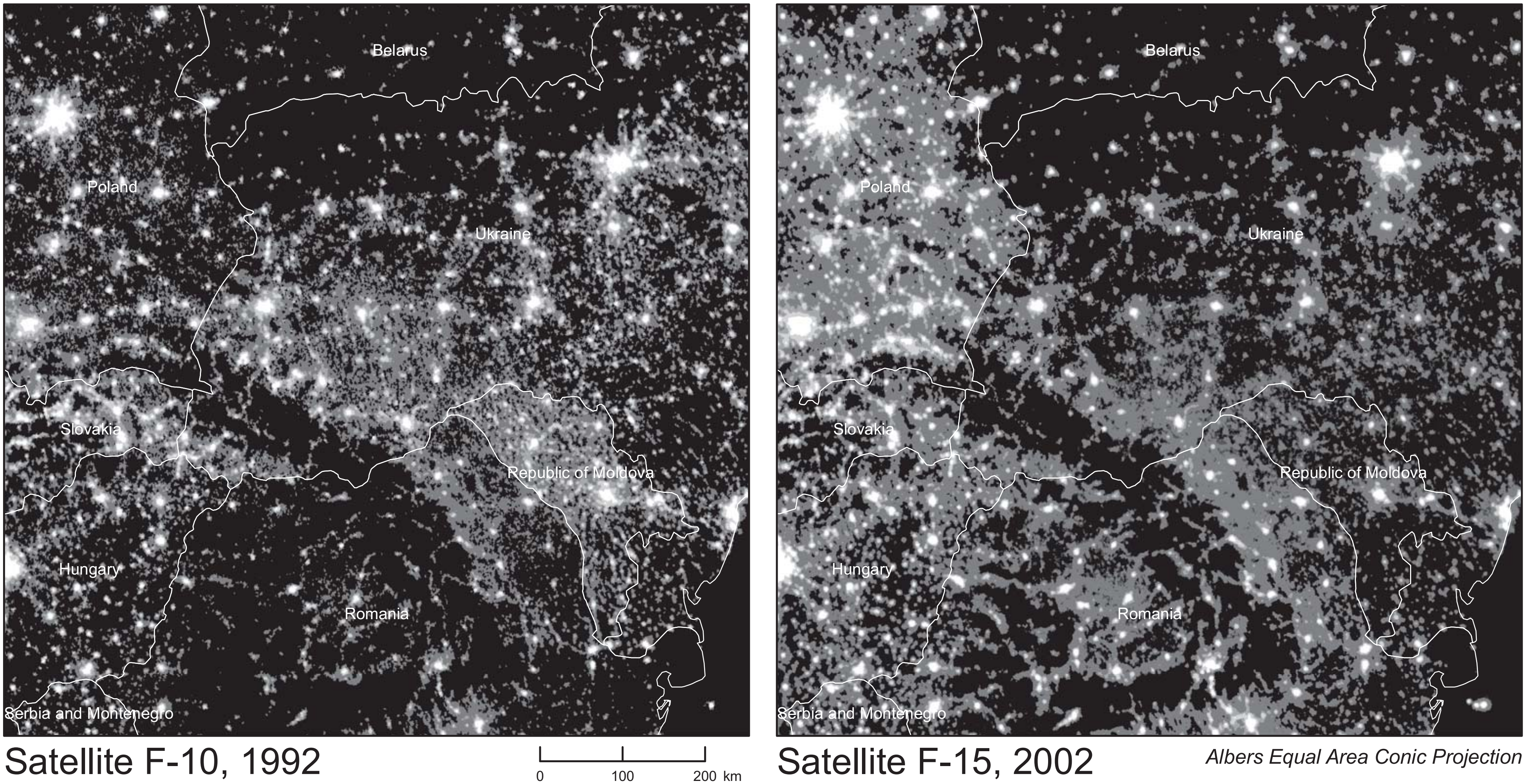


Figure 2: Eastern Europe in Lights

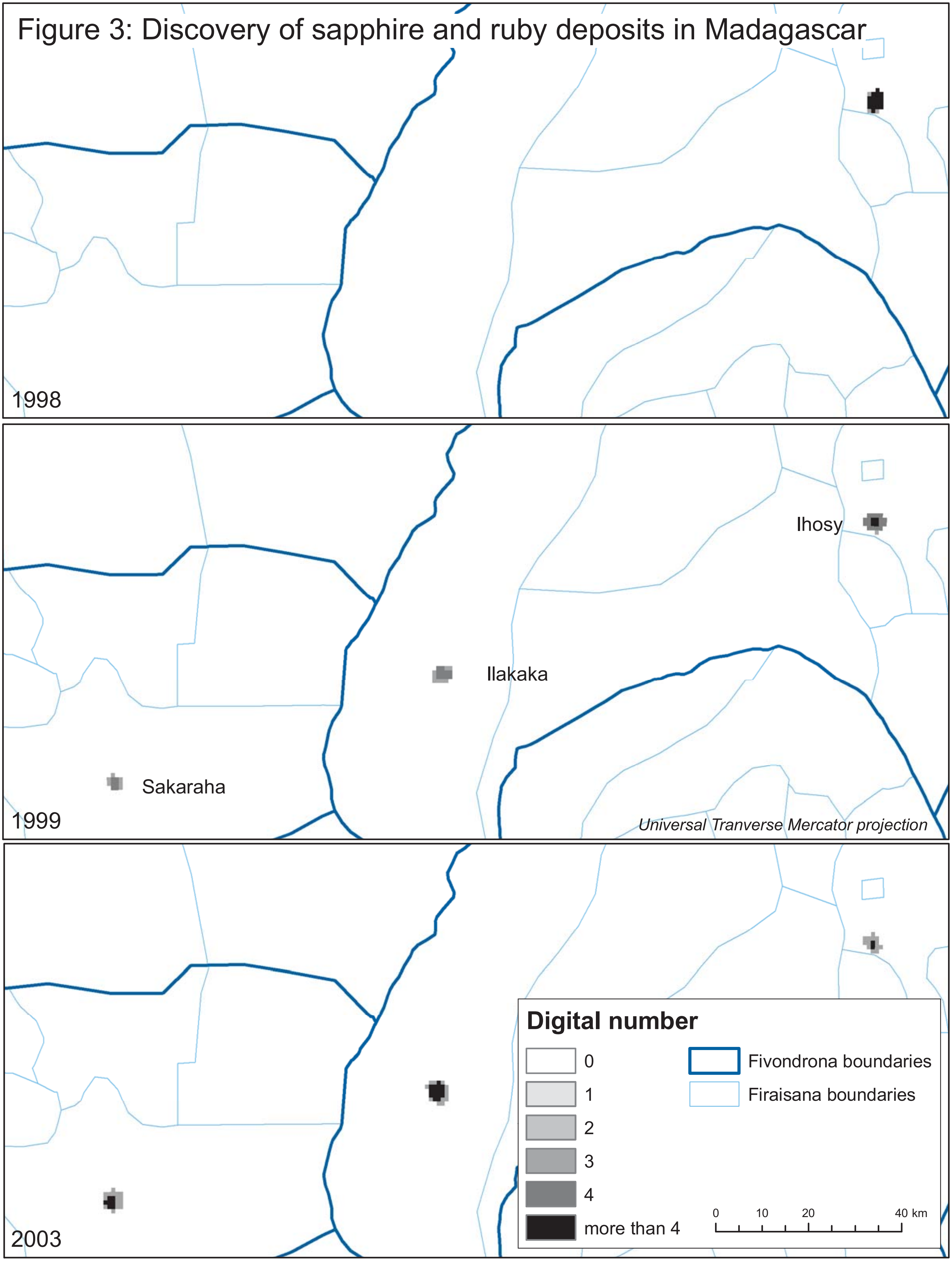
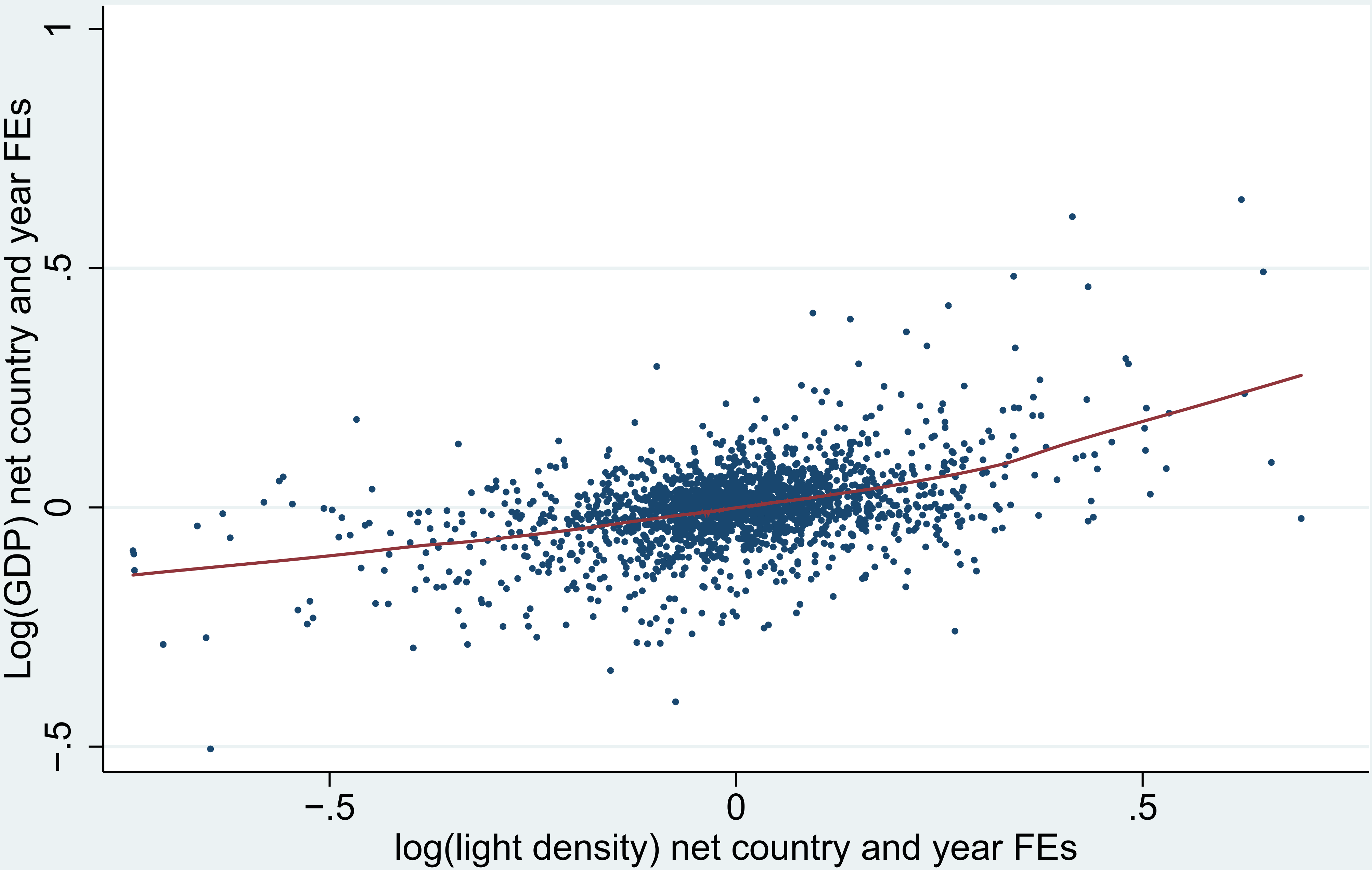


Figure 3: Discovery of sapphire and ruby deposits in Madagascar

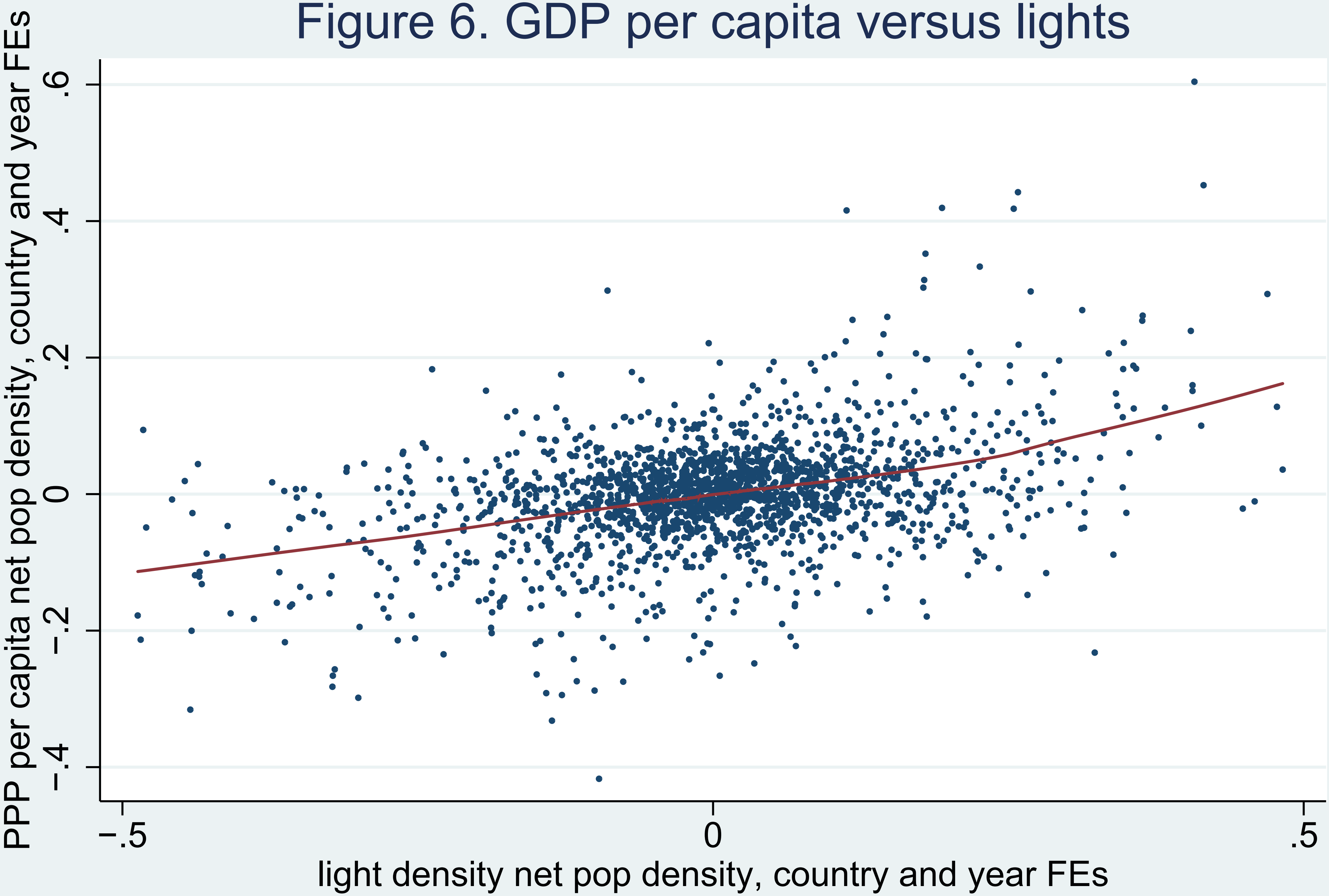
Figure 4. GDP versus lights: panel



note: excludes 3 points to the left and 4 to the right

Figure 5: GDP versus lights: long differences

Figure 6. GDP per capita versus lights



note: excludes 20 points to the left and 16 to the right

Figure 7a: lights in Guatemala

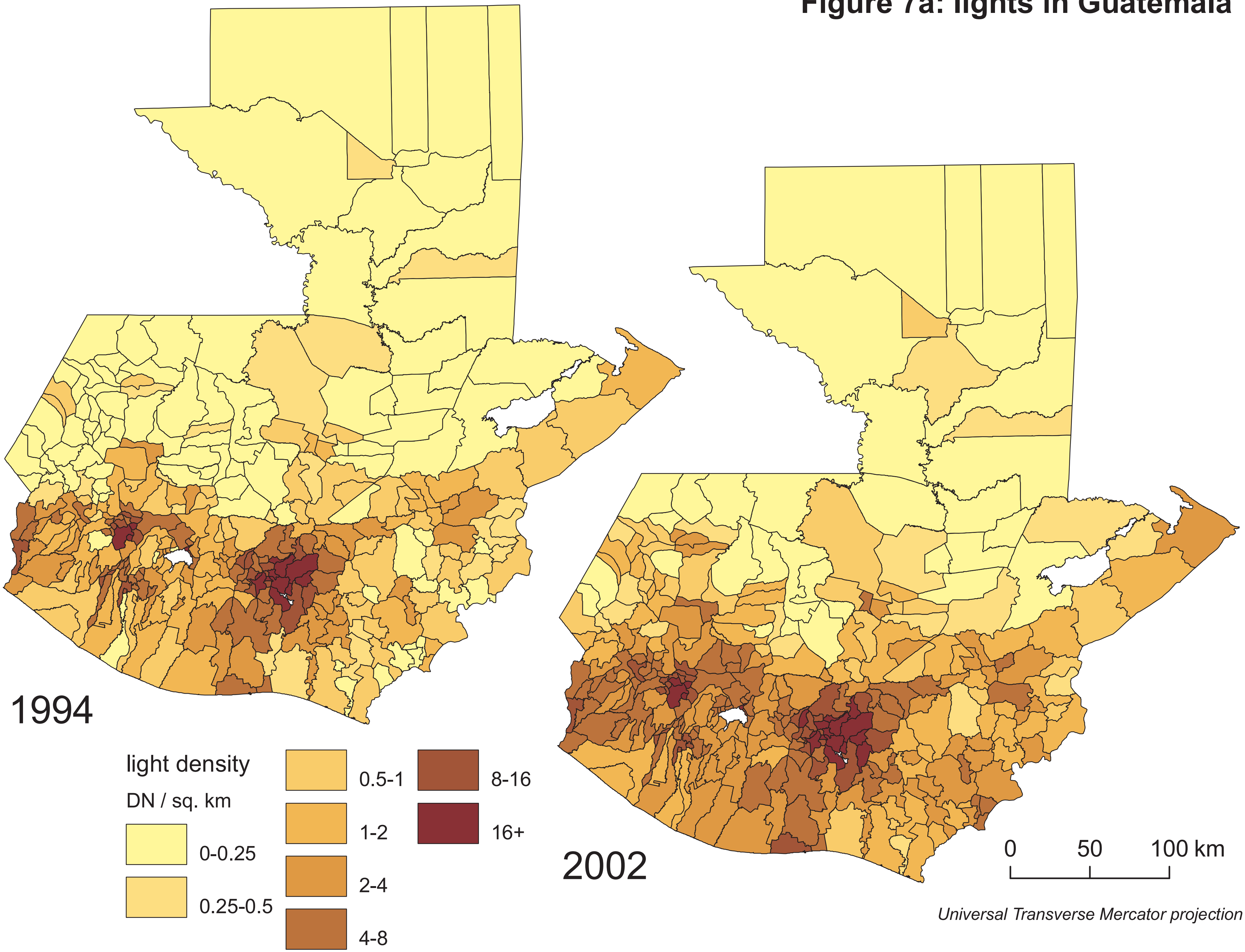


Figure 7b: consumption in Guatemala

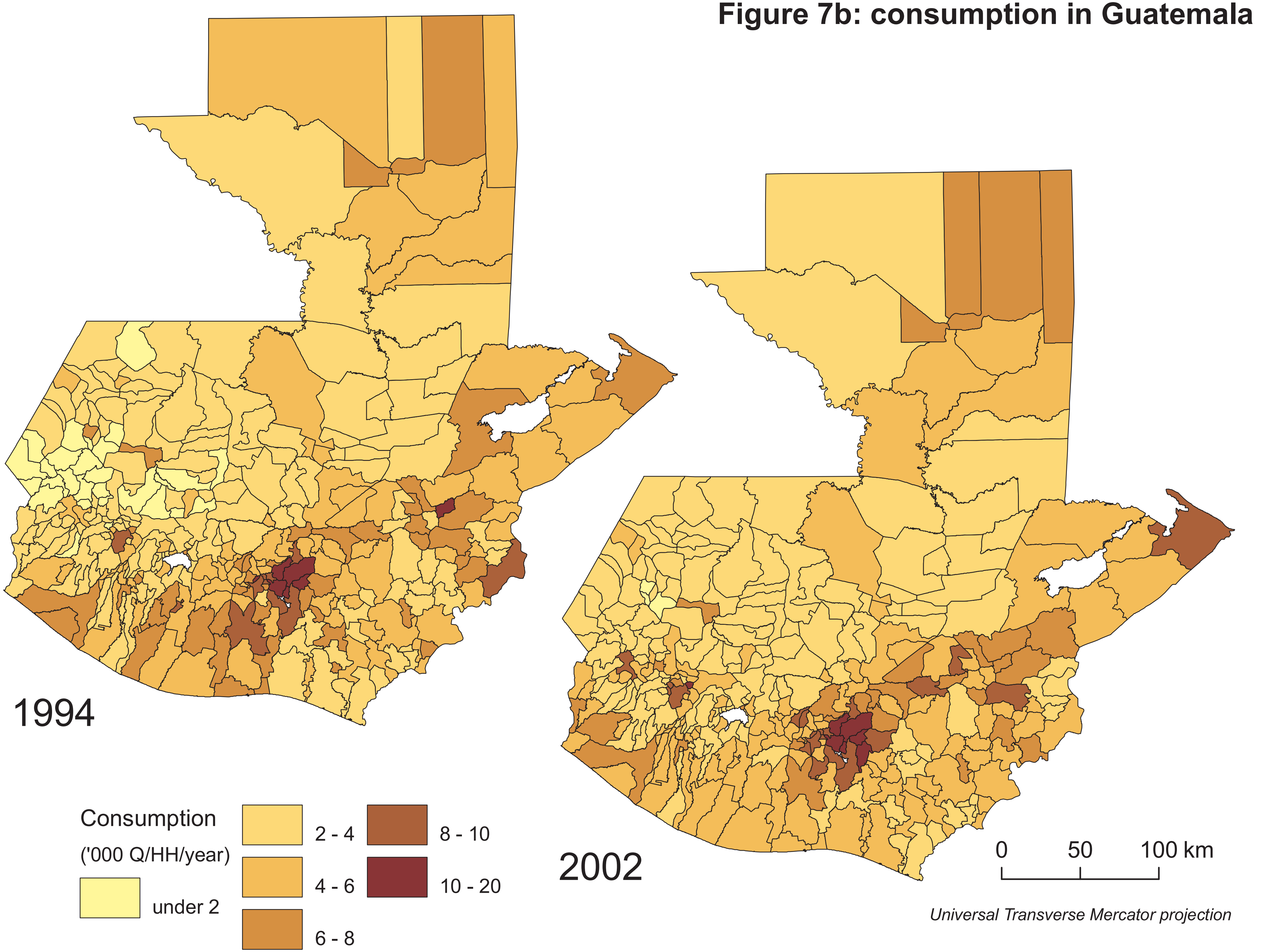


Figure 7c: Population density in Guatemala

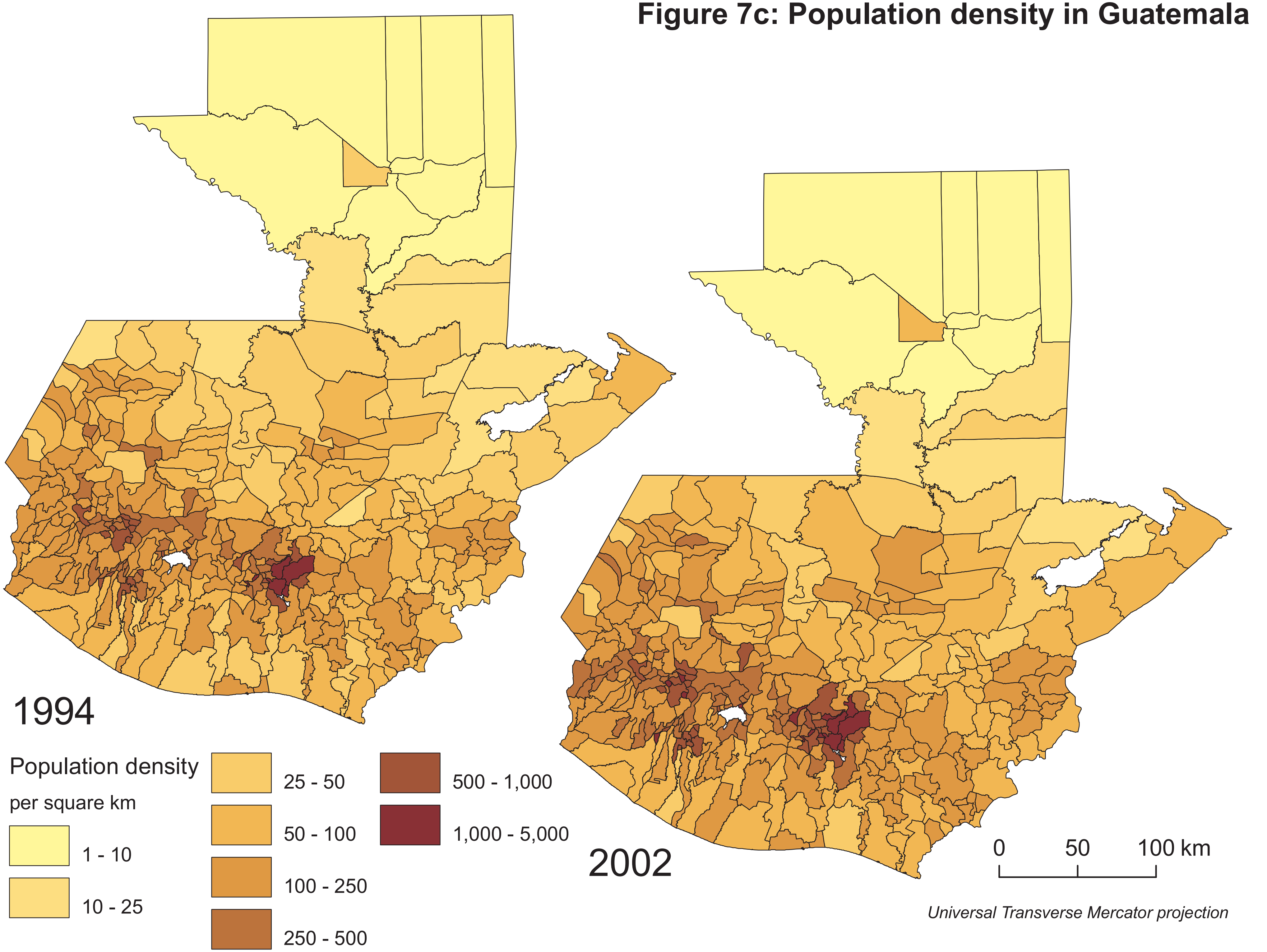


Figure 8a: Outlines of city lights for selected years

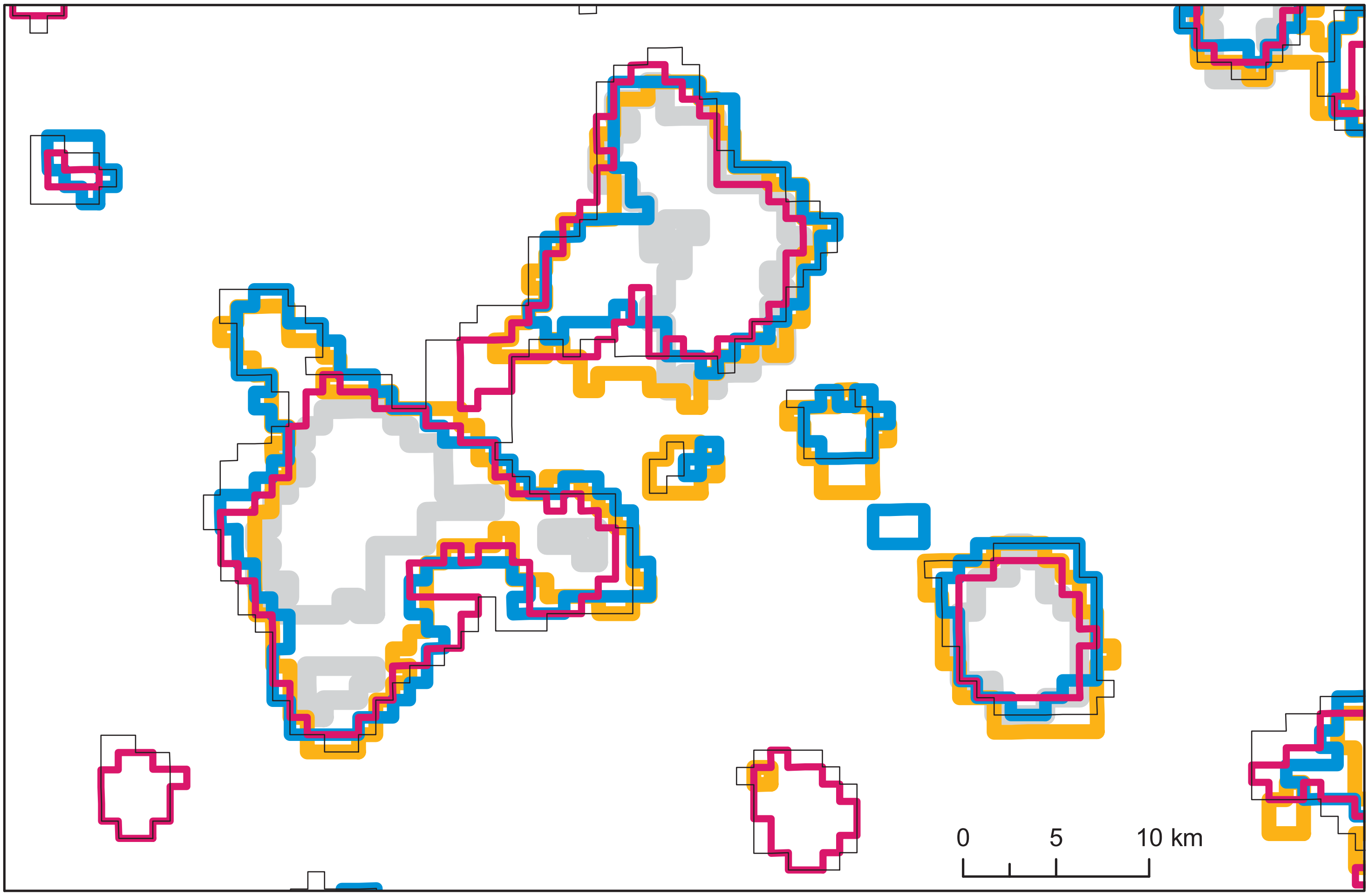


Figure 8b: Light outlines with outer envelope and city point populations