NBER WORKING PAPER SERIES

NEWER NEED NOT BE BETTER: EVALUATING THE PENN WORLD TABLES AND THE WORLD DEVELOPMENT INDICATORS USING NIGHTTIME LIGHTS

Maxim Pinkovskiy
Xavier Sala-i-Martin

Working Paper 22216
http://www.nber.org/papers/w22216

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2016

We are very grateful to Robert Feenstra, Robert Inklaar and Marcel Timmer for very helpful comments. We are very grateful to Rachel Schuh for superb research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research, the Federal Reserve Bank of New York or of the Federal Reserve System.

Newer Need Not be Better: Evaluating the Penn World Tables and the World Development Indicators Using Nighttime Lights
Maxim Pinkovskiy and Xavier Sala-i-Martin
NBER Working Paper No. 22216
May 2016
JEL No. O0,O4,O40

ABSTRACT

Nighttime lights data are a measure of economic activity whose error is plausibly independent of the measurement errors of most conventional indicators. Therefore, we can use nighttime lights as an independent benchmark to assess existing measures of economic activity (Pinkovskiy and Sala-i-Martin (2016)). We employ this insight to find out which vintages of the Penn World Tables and of the World Development Indicators better estimate true income per capita. We find that revisions of the PWT do not necessarily dominate their predecessors in terms of explaining nighttime lights (and thus, predicting unobserved true income). In particular, we find that the PWT 7.1 chain-based GDP series substantially outperforms the constant-price series in both PWT 8.0 and PWT 8.1, the two most recent vintages of the PWT. We additionally find that the World Development Indicators are as good, and often better, measures of unobserved true income as are any recent vintages of the Penn World Tables. Furthermore, we find that each new round of the International Comparisons Programme (ICP) has improved the WDI's ability to predict log unobserved true income. We also find that vintages tend to be good or bad at predicting unobserved true income roughly equally across the sample period, and do not tend to be particularly good at predicting unobserved income in the year of their price survey. We conclude that GDP series based on unadjusted domestic growth rates alone predict growth rates of true income better than series based on PPP adjustments to growth rates.

Maxim Pinkovskiy
Federal Reserve Bank of New York
33 Liberty Street
New York, NY 10045
maxim.pinkovskiy@ny.frb.org

Xavier Sala-i-Martin
Department of Economics
Columbia University
420 West 118th Street, 1005
New York, NY 10027
and NBER
xs23@columbia.edu

1 Introduction

The Penn World Table (PWT) is one of the most frequently used datasets in economics. The original paper introducing it (Summers and Heston 1991) has 4827 cites on Google Scholar as of April 2016. The PWT are frequently updated (they are now in their 8th vintage), and, typically, researchers use the latest version available in their work. However, the different versions of the PWT do not necessarily coincide in the years in which they overlap, and the resulting differences are sufficiently large to influence research outcomes. Ciccone and Jarocinski (2008) find that variables found to be robust determinants of economic growth using one vintage of the PWT no longer remain so when using the following vintage. Johnson et al. (2009) note that countries ranked as slowest growing by one vintage may be ranked as fastest growing by the following one. The problem is that we cannot tell which of the two datasets is wrong: as Johnson et al. (2009) note, the differences across these versions may represent mistakes in the old version, mistakes in the new version or mistakes in both versions. If this is the case, users face a problem in deciding whether to adopt the most recent update of the PWT and whether to stick with the old version.

This problem is magnified further when one considers that the PWT no longer is the only source of PPP-adjusted GDP estimates. The World Bank's World Development Indicators (WDI) have been using a different methodology from the PWT in order to adjust their estimates of national accounts data to purchasing power parity for many years. Given that there are differences in the PWT and WDI estimates of GDP per capita for the same countries and years, it is also not obvious a priori which set of estimates to use. Similarly to the PWT, the WDI are updated every time that the International Comparisons Program (ICP) releases a new worldwide survey of prices. Some of these revisions have attracted controversy. For example, the 2005 PPP revision decreased estimates of Chinese GDP by 40% over the previous version, while this change was largely undone by the subsequent (2011) PPP revision. Hence, whenever researchers wish to use PPP-adjusted estimates of national accounts data, they need to answer several questions. First, should they immediately update to the newest version of the PWT? Second, should they use the PWT or the WDI? Third, if they use the WDI, what version of the WDI should they use?

One way to attempt to answer this is to examine the methodology of construction of the various datasets, and select the one whose methods appear to have the most reasonable assumptions. For example, Johnson et al. (2009) suggest that the WDI growth rate estimates may be subject to fewer sources of error than those of the PWT, and may be preferable in that regard. Often, researchers just use the most recent dataset available on the assumption that progress in dataset construction is linear, and because that dataset has the most up-to-date coverage. However, absent a measure of GDP per capita that is independent of the different PWT and WDI datasets, this methodology of dataset choice cannot really test that it is selecting the right dataset.

In this paper, we propose a different approach. We employ a data-driven way to assess how well different series of GDP estimates measure unobserved true income relative to each other. If we had a measurement of GDP per capita whose measurement error was uncorrelated with the measurement errors of the different vintages of the Penn World Tables, it would be a straightforward matter to see which vintage was better by comparing them both to the independent measure. We follow Pinkovskiy and Sala-i-Martin (2016) and argue that such an independent measurement can be constructed using data on satellite-recorded nighttime lights, which were first studied by Elvidge et al. (1993) and in economics by Henderson, Storeygard and Weil (2012) and Chen and Nordhaus (2011). While errors in different vintages of the PWT and WDI come from errors in the underlying national accounts data (such as faulty assumptions about economic relationships like input-output tables), or from errors in calculating indicators of purchasing power parity between different currencies, errors in the relationship between nighttime lights and economic output come from weather and atmospheric disturbances that affect how light is captured by the orbiting satellites.

It is important to note that our approach is much more reduced-form than traditional analyses of the virtues and defects of different ways of constructing national income statistics. For our empirical tests to yield valid answers, we do not need to know anything about the way that the statistics that we are comparing were constructed, except that we can assume their measurement error to be orthogonal to the measurement error in the nighttime lights. As long as this assumption holds, our method yields impartial assessments on how best to combine the statistics at hand to predict unobserved true income. On the other hand, our method does not shed light on the methodological reasons why one estimator of GDP per capita might be better than another.

Our answer to the first question is that in general, newer versions of the Penn World Tables are not necessarily better than their predecessors. In particular, the chain series index in PWT 7.1 appears to have been more accurate than the closest variables to it in PWT 8.0 and PWT 8.1. This finding is not driven by nighttime lights being a biased indicator of output across industries and is not driven by any subsample of the data (although there is interesting regional heterogeneity in the relative quality of the two PWT vintages). More generally, no version of the PWT appears to have been unequivocally better than its predecessor. On our second question, we find that the current version of the WDI better predicts nighttime lights than any of the three latest versions of the PWT. This finding, which dovetails with the conclusion of Johnson et al. (2009), is partially explained by the fact that WDI 2011 is the first GDP database to use the 2011 ICP in order to implement PPP adjustments. Looking at the different vintages of the World Development Indicators, we answer our third question by noting that each successive PPP revision leads to a version of the WDI that dominates its predecessor. In particular, the 2011 ICP is superior to the 2005 ICP in measuring cross-country income differences (though not necessarily growth rates), while the 2005 ICP is likewise superior to the 1996 ICP.

Johnson et al. (2009) note that the variability in PWT growth rates increases the further one goes away from a given version's benchmark year (the year of the price survey used to compute the PPP adjustments), which is consistent with each price survey being most accurate in its benchmark year and losing accuracy subsequently as prices change. In light of this fact, Johnson et al. (2009) propose that one should interpolate between data in successive PWT versions (or create a chained index of adjacent PWT prices) to measure cross-country differences in any given year. We find only partial evidence for this suggestion. Instead, PWT 7.1 appears to perform well in all years even far away from 2005 (its benchmark year), while other vintages also perform uniformly well or uniformly poorly regardless of which was their benchmark year. On the other hand, the WDI 2011 vintage is unambiguously superior to its predecessors in all years, including long before its benchmark year of 2011.

The rest of the paper is organized as follows. Section 2 describes the data. Section 3 explains the empirical approach, which is very close to Pinkovskiy and Sala-i-Martin (2016). Sections 4, 5 and 6 present the results. Section 7 concludes.

2 Data

2.1 Penn World Tables

We use vintages 6.1 through 8.1 of the Penn World Table. The current version of the Penn World Table (version 8) and its relationship to its predecessors is described in detail in Feenstra, Inklaar and Timmer (2015). Before version 8, the Penn World Table took national accounts data in local currency from different countries, as well as the results of price surveys from the International Comparisons Project (ICP) in a given benchmark year, and then assigned purchasing power parities to countries with missing survey data in the benchmark year by predicting them with a regression.[1] Then, the PWT computed PPP-adjusted GDP away from the benchmark year by weighting together national growth rates of consumption, investment and government spending, with the weights being the current (PPP-adjusted) price component shares.[2] Starting with vintage 8.0, the PWT begun to present measures (such as "rgdpe") that use all available price survey

[1] As of PWT 8.0, this imputation is no longer performed.

[2] This procedure was used to produce the chain-weighted estimate of real GDP per capita (the variable "rgdpch"), which was the real GDP series endorsed by the authors of the PWT (Johnson et al. 2009). The PWT also provides an alternative Laspeyres-index based estimate of real GDP per capita ("rgdpl"), in which consumption, investment and government spending growth rates of GDP are weighted by *benchmark year* PPP-adjusted shares. In PWTs 6.3 through 7.1, the authors also provided an estimate in constant prices that was based on national accounts domestic absorption growth rates alone ("rgdpl2").

While we focus our analysis on "rgdpch" as it is the variable endorsed by the authors of the PWT, we consider the performance of the national growth-rate based measure in PWT8 ("rgdpna"), and we provide appendix tables showing results for "rgdpl" and comparisons of "rgdpl" with "rgdpch."

years to compute GDP in constant prices away from the benchmark year, and that thus rely less on national growth rates. Feenstra, Inklaar and Timmer (2015) argue that these measures are invariant to future price data, and should not change over time unless the underlying national accounts data is revised, a feature that was not present in earlier PWTs.

We use the variable "rgdpch" from Penn World Tables vintages 6.1, 6.2, 6.3, 7.0, and 7.1. This variable is no longer available for PWT version 8; however, the variable correspondence between PWT 7.1 and PWT 8.0 indicates that "rgdpe" is the closest corresponding variable to the GDP concept of "rgdpch." This is intuitive since both "rgdpe" and "rgdpch" are constructed using growth rates that have been modified by price surveys, rather than using national growth rates alone. Feenstra, Inklaar and Timmer (2015) suggest the use of "rgdpe" for "comparing living standards across countries and across years."

PWT 8.0 and 8.1 also includes a version of real GDP that uses only national growth rates to compute GDP away from the 2005 benchmark year, called "rgdpna."[3] Feenstra, Inklaar and Timmer (2015) recommend using this variable to "compare growth of GDP over time in each country," as opposed to comparing living standards across countries, stating that "if the sole object is to compare the growth performance of economies, we would recommend using the "rgdpna" series (and this is closest to earlier versions of PWT)." In our analysis, we will consider the performance of both "rgdpna" and "rgdpe" because we wish to assess whether using national growth rates is worse or better at estimating unobserved true income than is using earlier ICP surveys to compute constant-dollar estimates of true income. However, we will focus on "rgdpch" and "rgdpe" because one of the key efforts of the Penn World Tables is to provide estimates of GDP in constant prices both across space and over time, which is the specific purpose of these variables.

2.2 World Development Indicators

We use the variable "GDP per capita, constant (year of ICP) international dollars" from the 2006, 2013 and 2015 editions of the World Development Indicators. These editions correspond to the latest measurements of GDP per capita using 1996 PPP, 2005 PPP and 2011 PPP respectively. Unlike the PWT, the World Development Indicators do not modify national growth rates with information from the price surveys at all, but simply use them as they appear in the national accounts. In particular, and again unlike the PWT, the WDI take national growth rates directly from the national accounts without any further editing, while the PWT performs some editing of the growth rates before using them (Johnson et al. 2009)

[3] Feenstra, Inklaar and Timmer (2016) note that this variable is conceptually similar to "rgdpl" and "rgdpl2"

2.3 Nighttime Lights

Data on lights at night is collected by the Defense Meteorological Satellite Program Operational Linescan System (DMSP-OLS) satellite program and is maintained and processed by the National Oceanic and Atmospheric Administration (NGDC 2010). Satellites orbit the Earth, sending images of every location between 65 degrees south latitude and 65 degrees north latitude at a resolution of 30 arcseconds (approximately 1 square km at the equator) at 20:30 to 22:00 local time.[4] The images are processed to remove cloud cover, snow and ephemeral lights (such as forest fires) to produce the final product available for download at

http://www.ngdc.noaa.gov/dmsp/downloadV4composites.html

The nighttime lights data is available from 1992 to 2013, and we use the data up to 2010 because most series of GDP per capita represent vintages of the PWT or the WDI that were discontinued before 2010.

Each pixel (1 square kilometer) in the luminosity data is assigned a digital number (DN) representing its luminosity. The DNs are integers that range from 0 to 63. We construct our lights proxy for aggregate income by summing up all the digital numbers across pixels

$$\text{Lights}_{j,t} = \sum_{i=1}^{63} i * (\#\text{ of pixels in country } j \text{ and year } t \text{ with DN} = i\ )$$

This formula has been used to aggregate the nighttime lights maps into lights-based indices for each country and year in nearly the entire literature on nighttime lights in economics, including Henderson, Storeygard and Weil (2012), Chen and Nordhaus (2011) and Michalopoulos and Papaioannou (2013, 2014). For years with multiple satellites available, we average the logarithms of our aggregate luminosity measure, following Henderson, Storeygard and Weil (2012).

It is very well established that lights are very strongly correlated with measures of economic activity, such as national accounts GDP, in levels and growth rates. Henderson, Storeygard and Weil (2012) provide these correlations, dramatic pictures of long-term differences in incomes (North vs. South Korea) as well as short-term fluctuations (the Asian financial crisis of 1997-8) reflected in lights. Michalopoulos and Papaioannou (2013, 2014) present evidence that nighttime light density in a sample of African villages is correlated with development indicators for these villages. Table I of our paper show that there are strong regression relationships between log nighttime lights per capita and all of the PWT and WDI vintages.

[4] There are one or two satellites recording nighttime lights in each year, with an old satellite being retired and a new satellite being launched every few years. The satellites from which data is avaliable are as follows: the satellite F-10 (in orbit 1992-1994), F-12 (1994-1999), F-14 (1997-2003), F-15 (2000-2007), F-16 (2004-2009), F-18 (2010-2013), and the VIIRS satellite (April 2012-present).

Our paper is closest in spirit to Henderson, Storeygard and Weil (2012) and Chen and Nordhaus (2011) in that it also considers the problem of optimally combining measures of economic activity. However, instead of using nighttime lights as a component of such a measure, we use it as an auxiliary variable to help uncover the correlation structure between the measures we do wish to use in our index. We also consider a different type of predictor for true income that do either Henderson, Storeygard and Weil (2012) or Chen and Nordhaus (2011), which allows us to make fewer assumptions on the data generating processes that we consider.

There are also well-known problems with the relation between nighttime lights and economic development, which we need to take into account. Pixels with DN equal to 0 or 63 are top- or bottom-censored. The light data also are affected by overglow and blooming: light tends to travel to pixels outside of those in which it originates, and light tends to be magnified over certain terrain types such as water and snow cover (Doll 2008). Given that we will compute national-level estimates of aggregate lights, it is unlikely that these sources of error will be large enough or sufficiently correlated with important variables that they will confound our analysis. Another problem may be that satellites age in space and are eventually retired. Hence, they might give inconsistent readings from year to year, or new satellites may give fundamentally different readings from old ones. While some evidence of this problem exists, we will show in Section 3 that our calculations are supported by assumptions that allow nighttime lights to have all of the data problems described above, so long as nighttime lights are correlated with true income.

Table I presents univariate regressions of log nighttime lights per capita on each of the 10 measures of GDP per capita that we are considering. All regressions cluster standard errors by country, and all regressions for a given PWT or WDI vintage are estimated on the sample of countries and years for which this vintage is available. Each regression is run in four ways: without fixed effects (Panel I), with year fixed effects (Panel II), with country fixed effects (Panel III) and with country and year fixed effects (Panel IV). We see that the regression coefficient on each GDP measure is statistically significant at 1% and close to unity in magnitude. For example, in row 1 and column 1 of the table, we see the coefficient in a regression of log nighttime lights per capita on log PWT 6.1 GDP per capita and a constant, which is equal to 1.192 (s.e. = 0.053), and which has an R^2 of 0.77. These results suggest a very strong association between log nighttime lights per capita and log GDP per capita. Once we include country fixed effects, the regression coefficients generally decline in magnitude, but remain statistically significant at least at 10%. The exceptions are PWT 8.0 and PWT 8.1, whose coefficients lose statistical significance when both country and year fixed effects are included. The weak correlation of log PWT 8.0 and 8.1 GDP per capita with log lights per capita once country and year fixed effects are controlled for foreshadows the relatively poor quality of these measures as will be uncovered in the later analysis.

2.4 Other Data

We use a number of covariates to test the crucial maintained assumption of our paper; that nighttime lights are correlated with different vintages of the PWT or the WDI only through their joint correlation with true income (see the introduction and Section 3 below). These covariates are log capital formation as percent of GDP, log export share, log import share, log general government expenditure share of GDP, and log consumption share. All these covariates are from the World Development Indicators. The covariates will be discussed at greater length in Section 4.

3 Mathematical Framework

In this paper, we use the methodology of Pinkovskiy and Sala-i-Martin (2016) to determine which datasets do a better job of computing unobserved true income per capita. Therefore, this section largely parallels the exposition of that paper. As we will consider using more than two measurement variables to predict true income per capita, we extend the approach of Pinkovskiy and Sala-i-Martin (2016) to multiple variables. The general approach is closely related to the measurement error literature in econometrics, including Adcock (1878), Griliches (1986), Fuller (1987), and Hausman (2001).

Our goal is to find the best unbiased linear predictor of log true income per capita $y^*_{i,t}$, which is the log total per capita value added in country i and year t that we would compute if national statistical systems could record all income being earned and if PPPs were perfectly computed. We will assume that this true income per capita is generated through some exogenous stochastic process that may not be stationary (if there is economic growth, for example). We cannot observe $y^*_{i,t}$ directly. Instead, we can observe data on log light intensity per capita ($y^L_{i,t}$) and on K different measurements of log GDP per capita ($y^1_{i,t}$, $y^2_{i,t}$,...,$y^K_{i,t}$) for a sample of countries i and years t. These data are related to log true income per capita according to the following system of equations (partialling out constants and other covariates):

$$y^L_{i,t} = \beta^L y^*_{i,t} + \varepsilon^L_{i,t} \tag{1}$$

$$y^1_{i,t} = \beta^1 y^*_{i,t} + \varepsilon^1_{i,t} \tag{2}$$

$$\vdots$$

$$y_{i,t}^{K}=\beta^{K}y_{i,t}^{*}+\varepsilon_{i,t}^{K} \tag{3}$$

We can express this system of K equations in vector form as

$$y_{i,t}=\beta y_{i,t}^{*}+\varepsilon_{i,t} \tag{4}$$

where

$$\begin{aligned} y_{i,t} &= \left(y_{i,t}^{1},y_{i,t}^{2},...y_{i,t}^{K}\right)' \\ \beta &= \left(\beta^{1},\beta^{2},...\beta^{K}\right)' \\ \varepsilon_{i,t} &= \left(\varepsilon_{i,t}^{1},\varepsilon_{i,t}^{2},...,\varepsilon_{i,t}^{K}\right)' \end{aligned}$$

In other words, each of the measured variables is a linear function of log true income per capita, perturbed by some error.[5] This framework is very similar to the one used by Henderson, Storeygard and Weil (2012) and Chen and Nordhaus (2011), except that both of these papers used only one additional measure of log GDP per capita and assumed that $\beta^{1}=1$, which means that log GDP per capita is an unbiased proxy for log true income per capita. Instead, we do not assume that any of our measured proxies are unbiased, and allow them to deviate from log true income per capita along a linear trend. [6]

We assume that the error terms in all three processes are mean independent of true income. That is,

$$E\left(\varepsilon_{i,t}^{L}|y_{i,t}^{*}\right)=E\left(\varepsilon_{i,t}^{1}|y_{i,t}^{*}\right)=\cdots=E\left(\varepsilon_{i,t}^{K}|y_{i,t}^{*}\right)=0 \tag{A1}$$

The critical assumption of this paper is that the error term in the lights equation (1), $\varepsilon_{i,t}^{L}$, is uncorrelated with the error terms from the measurement equation (4) and following, $\varepsilon_{i,t}^{1}$ through $\varepsilon_{i,t}^{K}$ conditional

[5] In fact, we can substantially relax the functional form specification in equation 1 to read

$$y_{i,t}^{L}=f_{i,t}\left(y_{i,t}^{*}\right)+\varepsilon_{i,t}^{L}$$

so long as

$$cov\left(y_{i,t}^{*},f_{i,t}\left(y_{i,t}^{*}\right)\right)\neq 0$$

(The analogous assumption in our framework is $\beta^{L}\neq 0$. We test and confirm both of these assumptions in Section 4).

This is a much more general framework that allows for errors in the lights data such as nonlinearity, top- and bottom-coding, differences in the lights-to-income relationship as satellites age and are replaced, and differences in the lights-to-income relationship across countries because of cultural attitudes to nighttime light and light pollution, as long as this heterogeneity does not dominate the positive relation between nighttime lights and true income within our sample.

[6] Pinkovskiy and Sala-i-Martin (2016) find that when the measurement variables are log GDP per capita and log survey means, it is the case that $\beta^{GDP}>\beta^{Surveys}$, with the difference between the two slope coefficients being significant. In the current paper, we do not typically find that the coefficients on different measurements of log GDP per capita are statistically (or economically) significantly different from each other. Nevertheless, for the sake of generality, we allow the richer model with potentially different coefficients β^{i}, since it nests the model in which all the β^{i}'s are the same.

on true income:

$$E\left(\varepsilon_{i,t}\varepsilon^{L}_{i,t}|y^{*}_{i,t}\right)=0 \tag{A2}$$

Assumption A2 is the key reason for the use of the lights data. This assumption has also been made in Henderson, Storeygard and Weil (2012) and Chen and Nordhaus (2011). This is a plausible assumption because the data generating processes of the lights data and of the different measures of GDPare largely disjoint; lights data is collected by satellites without respect for the statistical assumptions and methods of the underlying countries, whereas measurements of GDP are obtained primarily by using business surveys and the government's accounts, as well as by using price surveys, all of which depend on asking people about their income or about the prices that they face or offer.

We are interested in finding the best unbiased linear predictor of log true income per capita ($y^{*}_{i,t}$) in terms of the vector $y_{i,t}$,

$$z_{i,t}=\gamma' y_{i,t} \tag{5}$$

where γ is a vector of weights on each of the GDP measurements that we are considering:

$$\gamma=\left(\gamma^{1},\gamma^{2},...\gamma^{K}\right)'$$

Hence, we want to compute the vector γ, which minimizes the mean squared error

$$\gamma^{*}=\arg\min_{\gamma}E\left(\left(y^{*}_{i,t}-\gamma' y_{i,t}\right)^{2}\right) \tag{6}$$

subject to the constraint that the proxy be unbiased, that is:

$$E\left(\left(\gamma^{*}\right)' y_{i,t}|y^{*}_{i,t}\right)=y^{*}_{i,t} \tag{7}$$

This constraint can be reformulated as

$$\left(\gamma^{*}\right)'\beta=1 \tag{8}$$

Now, plugging equation (4) into the value function equation (6), we obtain

$$E\left(\left(y^{*}_{i,t}-\gamma' y_{i,t}\right)^{2}\right)=\left(1-\gamma'\beta\right)E\left(\left(y^{*}_{i,t}\right)^{2}\right)+\gamma'\Sigma\gamma \tag{9}$$

where we define

$$\Sigma = var\left(\varepsilon_{i,t}\right) = E\left(\varepsilon_{i,t'}\varepsilon'_{i,t}\right)$$

and applying the constraint equation (8) eliminates the first term.

Therefore, our constrained optimization problem becomes

$$\gamma^* = \arg\min_{\gamma} \gamma'\Sigma\gamma \tag{10}$$

subject to

$$\gamma'\beta = 1 \tag{11}$$

Solving this problem with traditional constrained optimization techniques, we obtain the relation

$$\gamma^* = \frac{1}{\beta'\Sigma\beta}\Sigma^{-1}\beta \tag{12}$$

The population parameters β and Σ on the right hand-side of equation (12) are unknown, and we cannot recover them without further assumptions. However, we now proceed to show that we can use the lights data to recover a vector that is proportional to $\Sigma^{-1}\beta$, and thus to γ^*. To do so, we compute the population regression

$$y^L_{i,t} = b^0 + \ b'y_{i,t} \tag{13}$$

which is the regression of log lights per capita on all our proxies of log GDP per capita as well as on a constant. It will be easy to show that

$$b = \alpha\gamma^* \tag{14}$$

for some unknown constant α. In other words, the vector of regression coefficients on the measurment variables is proportional to the vector of optimal weights of these measurement variables in the best unbiased linear predictor (5). This finding does not allow us to compute γ^*, but permits us to compute ratios of the optimal weights to each other. For example, we can estimate the relative weight of GDP per capita measurement #1, which will be given by

$$\frac{\gamma_1^*}{\gamma_1^* + \gamma_2^* + ... + \gamma_K^*} = \frac{b_1}{b_1 + b_2 + ... + b_K}$$

To prove the equality (14), we note that the formula for the vector of regression coefficients on the log GDP per capita measures, b, is:

$$b^{G} = var\left(y_{i,t}\right)^{-1} cov\left(y_{i,t}^{L}, y_{i,t}\right) \tag{15}$$

Under Assumptions A1 and A2, as well as equations (1) and (4), these variances and covariances can be expressed as follows:[7]

$$var\left(y_{i,t}\right) = \beta\beta' var\left(y_{i,t}^{*}\right) + \Sigma \tag{16}$$

$$cov\left(y_{i,t}^{L}, y_{i,t}\right) = \beta\beta^{L} var\left(y_{i,t}^{*}\right) \tag{17}$$

The expression for the inverse of $var\left(y_{i,t}\right)$ can be simplified using the Sherman-Morrison Formula to

$$var\left(y_{i,t}\right)^{-1} = \left(\beta\beta' var\left(y_{i,t}^{*}\right) + \Sigma\right)^{-1} = \Sigma^{-1} - \frac{1}{1 + \beta'\Sigma\beta var\left(y_{i,t}^{*}\right)} \Sigma^{-1}\beta\beta'\Sigma^{-1} var\left(y_{i,t}^{*}\right) \tag{18}$$

Substituting equations (18) and (17) into equation (15) leads to the formula[8]

$$b = \left(\frac{\beta^{L} var\left(y_{i,t}^{*}\right)}{1 + \beta'\Sigma\beta var\left(y_{i,t}^{*}\right)}\right) \Sigma^{-1}\beta \tag{19}$$

Note that the regression coefficient b is just the vector $\Sigma^{-1}\beta$ multiplied by a scalar, $\left(\frac{var\left(y_{i,t}^{*}\right)}{1+\beta'\Sigma\beta var\left(y_{i,t}^{*}\right)}\right)$. Hence,

$$b = \left(\gamma^{*}\right)\left(\frac{\beta'\Sigma\beta\beta^{L} var\left(y_{i,t}^{*}\right)}{1 + \beta'\Sigma\beta var\left(y_{i,t}^{*}\right)}\right)$$

so the vector of regression coefficients b is equal to the vector of optimal weights γ^{*}, multiplied by some unknown scalar $\alpha = \frac{\beta'\Sigma\beta\beta^{L} var\left(y_{i,t}^{*}\right)}{1+\beta'\Sigma\beta var\left(y_{i,t}^{*}\right)}$.

Therefore, we can estimate the weight that should be placed on any measure of log GDP per capita relative to the total weight on the K proxies for log true income per capita. We cannot estimate each weight individually, or their sum, but we can estimate their ratios.

The core of our analysis in Section 4 will be running regressions similar to equation (13) presenting estimates of the optimal relative weights on the measurements of GDP per capita $\frac{b^{1}}{b^{1}+b^{2}+\ldots+b^{K}}, \frac{b^{2}}{b^{1}+b^{2}+\ldots+b^{K}}, \ldots, \frac{b^{K}}{b^{1}+b^{2}+\ldots+b^{K}}$, when the elementary specifications in equations (1) and (4) – and hence, regression equation (13) – are aug-

[7] In the more general model in which we assume a general functional form for the lights-true income relation (see footnote 8), equation (17) replaces the term $\beta^{L} var\left(y_{i,t}^{*}\right)$ with the term $cov\left(y_{i,t}^{*}, f_{i,t}\left(y_{i,t}^{*}\right)\right)$.

[8] In the more general model in footnote 8, equation (19) should replace the term $\beta^{L} var\left(y_{i,t}^{*}\right)$ with the term $cov\left(y_{i,t}^{*}, f_{i,t}\left(y_{i,t}^{*}\right)\right)$. The rest of the logic proceeds exactly as in the text.

mented by covariates, when they are estimated on different samples, or when when they are estimated using different lights measures.

4 Results 1: Is Newer Really Better?

Table IIA presents the estimates of the coefficients of the bivariate regressions of log nighttime lights per capita on the variables "rgdpch" or "rgdpe" in each vintage of the PWT and on its successor. Again, all regressions cluster standard errors by country and are run on somewhat different samples because of the different availabilities of the vintages. If all newer vintages of the PWT were superior to their predecessors in measuring log true income per capita, and since we have shown in Section 4 that the coefficients of each such regression are proportional to the weights that each PWT and its successor would receive in a best unbiased linear predictor of log true income per capita, we should expect the coefficients on each newer vintage to be statistically significant and large, and the coefficients on each older vintage to be small in magnitude and statistically insignificant. We observe that this does not happen in general. Instead, the coefficients on the older vintages are routinely statistically significant and larger than the coefficients on the newer vintages. Particularly striking is the comparison of PWT 7.1 and PWT 8.0 in Column 2, for which the coefficient on the older (PWT 7.1) vintage is statistically significant and large regardless of which fixed effects are included, while the coefficient on PWT 8.0 is small and insignificant. For example, in row 3 of column 2, in which country fixed effects are included, the coefficient on PWT 7.1 is a precisely measured 0.473 (s.e.= 0.115), while the coefficient on PWT 8.0 is -0.040 (s.e. $= 0.095$), which is essentially zero. Motivated by this finding, we also compare the most recent vintage of the PWT, 8.1, with PWT 7.1 in Column 7. We likewise find that the regression coefficient on PWT 7.1 is statistically significant and large regardless of the fixed effects, while the coefficient on PWT 8.1 is uniformly small.

We can formalize these observations by considering the optimal weights on each pair of successive PWT vintages implied by these bivariate regressions. They are presented in Table IIB. Since these weights are quotients of regression coefficients, their distribution may differ substantially from normality in finite samples, so instead of reporting asymptotic standard errors, we report boostrapped 95% confidence intervals.[9] A natural null hypothesis to consider if newer PWT vintages should indeed be better is that the newer vintage of the PWT should receive unit weight (and the older vintage should receive zero weight). We find that this null hypothesis is rejected for 12 out of the 28 cells of this table (7 comparisons with 4 different types of fixed effects), or for nearly half the tests, which is very unlikely to happen by chance if the null hypothesis

[9] We employ 120 block-bootstrapped samples in which the bootstrap block is the country. We should note that the two weights are functionally dependent on each other since they add up to unity. Therefore, the statistics for the second weight in each regression can be derived from subtracting the statistics for the first weight from unity, and contain no additional information. They are reported merely for the convenience of the reader.

is indeed true. Moreover, we often fail to reject the null hypothesis that the old vintage should receive unit weight. In particular, regardless of which fixed effects are included, we can always reject the null hypothesis that the newer vintage has unit weight, we can always fail to reject the null hypothesis that the old vintage has unit weight, and we can almost always reject the null hypothesis that the weights are equal for the comparisons between PWT 7.1 and PWT 8.0, as well as between PWT 7.1 and PWT 8.1 (columns 2 and 7 respectively). For example, in row 1, column 2, we regress log nighttime lights per capita on log PWT 8.0 GDP per capita and log PWT 7.1 GDP per capita, as well as on a constant. We obtain that the optimal weight on log PWT 8.0 GDP per capita is 0.13, with a confidence interval of $(-0.38, 0.54)$, which excludes unity. On the other hand, the optimal weight on log PWT 7.1 GDP per capita is 0.86, with a confidence interval of $(0.45, 1.3)$, which includes unity. In fact, for this cell, the hypothesis that we should use only the older vintage to calculate unobserved true income per capita would be maintained.

For comparisons between the other vintages, the estimated optimal weights vary much more with whether the regression tries to fit the log or the growth rate of output (whether or not country fixed effects are included). For example, in the comparison between PWT 6.1 and PWT 6.2 (the case study considered in Johnson et. al.), the point estimates suggest placing a weight of about 40% on PWT 6.1 and a weight of about 60% on PWT 6.2 in the optimal proxy if country fixed effects are not included, and weights of 47% and 53% respectively if country fixed effects are included. Similar ambiguity exists in the comparison of PWT 6.3 and PWT 7 (the first vintage to use the price data from ICP 2005 for purchasing power parity adjustments), where the results without country fixed effects suggest an even split of the weight between the two vintages, while the results with country fixed effects suggest placing all the weight on PWT 6.3 (albeit with large confidence intervals on the latter prediction). However, as long as the measurement error in lights is taken as exogenous, even these ambiguous results are radically inconsistent with the hypothesis that each PWT revision represents a clear improvement over the previous vintage, because this hypothesis predicts a weight of unity on each successive new vintage of the PWT in a comparison with its predecessor.

Particularly provocative is the finding that PWT 7.1 appears to outperform PWT 8.1, the latest vintage of the PWT. To illustrate this finding, we plot the partial correlation between nighttime lights and PWT 7.1 in Figure III, and the partial correlation between nighttime lights and PWT 8.1 in Figure IV, both of the partial correlations having been taken from the bivariate regression of log nighttime lights per capita on log GDP per capita from PWT 7.1 and log GDP per capita from PWT 8.1 without further fixed effects. Recall that the slopes of the lines of best fit on these graphs are the regression coefficients.[10] We see that

[10] All figures exclude several countries that generate outliers, namely Zimbabwe, El Salvador, Nigeria, Syria, Yemen, Maldives, Bermuda, Equatorial Guinea, and Singapore. Using this restricted sample, the partial regression coefficients are 0.13 (s.e.=0.19) on PWT 8.1 and 0.92 (s.e.=0.19) on PWT 7.1. These coefficients are slightly more favorable to PWT 8.1 than the coefficients in Panel I, Column 7 of Table IIA.

there is a visible partial correlation between the residuals of log nighttime lights and the residuals of PWT 7.1 log GDP per capita once PWT 8.1 log GDP per capita has been controlled for. On the other hand, once PWT 7.1 log GDP per capita is controlled for, the correlation between the residuals of log nighttime lights and of PWT 8.1 log GDP per capita is virtually nil. This finding is contrasted with the very robust raw correlations of either series with log nighttime lights per capita (Figures I and II).

In tables III and IV we further subject the comparison of PWT 7.1 and PWT 8.1 to greater scrutiny by adding control variables to test our identification assumption, varying the nighttime lights measure, reweighting the specification, and investigating robustness to and heterogeneity across subsamples. A concern in using nighttime lights as a measure of output is that errors in the nighttime lights may be correlated with errors in the PWT through the component-by-component methodology of PWT construction. We recall that the PWT separately constructs and extrapolates PPPs for consumption goods, investment goods and government purchases. Therefore, it is possible (or at least conceivable) that mistakes in the PWT could take the form of systematic underestimation (or overestimation) of either one of these categories. Now, consumption, investment and government spending may have different ratios of light to economic output, for example if investment and government spending are very light-intensive (as it might involve building infrastructure), while consumption is not as light intensive (if it involves getting haircuts or other services). If different PWT vintages systematically understate or overstate some of these components, this could generate a correlation between their measurement error and that of the nighttime lights, violating Assumption A2. To address this concern, in column 2 of Table III we re-run our baseline specifications for the optimal weights on PWT 7.1 and PWT 8.1 while controlling for the log fractions of output constituted by consumption, investment (gross capital formation), government spending, imports and exports. In column 3 of the same table, we additionally include both the levels and the squares of the control variables to allow for nonlinear associations. The optimal weight on PWT 7.1, which is close to unity in the baseline specifications, remains larger than unity when all controls are included without country fixed effects, but declines to a statistically significant 0.73 if country and year fixed effects are included along with the controls. However, if all the controls are included together with their squares, the optimal weight on PWT 7.1 in the country and year fixed effects specification rises back to 0.77. Hence, these regressions provide some evidence that PWT 7.1 remains of value even when compared to the current version of the PWT. The next three columns of Table III (columns 4, 5 and 6) reestimate the baseline specification with different ways of parametrizing nighttime lights: using light density in place of lights per capita, calibrating the function aggregating digital numbers into a single index to best match GDP per capita in the OECD, and using the fraction of the population residing in lit areas (DN > 0) instead of lights per capita. For all of these robustness checks, the optimal weight on PWT 7.1 does not fall below 0.57, and is either statistically significantly different from zero or above unity in magnitude. Furthermore, we may be interested in how the two PWT vintages compare for the more populous countries, rather than treating each country equally. Therefore, in column 7, we reestimate our baseline specifications weighing each observation by population. Regardless of fixed effects, the optimal weight on PWT 7.1 exceeds unity in all specifications.

Feenstra, Inklaar and Timmer (2015) show that estimates of the Balassa-Samuelson effect depend on whether or not the sample of countries and years is restricted to the ones for which contemporaneous price data is available, or the ones for which prices need only to be interpolated rather than extrapolated. We present estimates of the optimal weights on PWT 7.1 and PWT 8.1 using only this sample (the benchmark-interpolated sample) in Column 8 of Table III. Our results for the comparison of PWT 7.1 and PWT 8.1 are unchanged in regressions without fixed effects if the sample is restricted to benchmark or interpolated ICP data. If fixed effects are included, PWT 7.1 still receives a positive weight while PWT 8.1 receives a negative weight, but the estimates are no longer statistically significant.

In Table IV we investigate whether our counterintuitive result that PWT 7.1 is better than the more recent PWT 8.1 is driven by specific regions of the world. Columns 2 through 6 exclude one at a time the developed world, Asia, Latin America, post-Communist countries and Africa from the baseline regression. The optimal weight on PWT 7.1 drops to a statistically significant 0.67 if Asia is excluded in the year fixed effects specification, but most of the estimates of the optimal weight are close to or above unity (including estimates when Asia is excluded but with country fixed effects). Therefore, our finding is not driven by any specific region. However, there is substantial heterogeneity in the optimal weight on PWT 7.1 across regions. Columns 7 through 11 present estimates of our baseline regression on each of the above regions taken separately. Not surprisingly, PWT 7.1 receives a very high weight (224% in the no fixed effects specification) when the regression is ran for Asia (column 8), but it also receives a greater weight than does PWT 8.1 when the regression is run only for Africa (column 11). For the regression in Latin America alone (column 9), PWT 7.1 receives a weight of 47% if country fixed effects are not included but statistically significant weights of 61% and 107% in the country and country and year fixed effects specifications. For the regression in the developed world, no weight estimates are significant and the standard errors explode. This suggests that the PWT 7.1 and the PWT 8.1 largely agree (up to an affine transformation) on the log and growth rate of output in the developed world, a finding that is not surprising given the better statistics of developed countries. Only for the post-Communist countries does PWT 8.1 outperform PWT 7.1, which it does for all four combinations of fixed effects. Therefore, users of the PWT primarily interested in the post-Communist countries did not err if they mechanically updated the Penn World Table after vintage 7.1.

4.1 Alternative Ways of Measuring GDP in PWT 7.1 and PWT 8.1

We have shown in the preceding section is that the variable "rgdpch" from PWT 7.1 is better correlated with nighttime lights than is the variable "rgdpe" from PWT 8.1. However, both PWT 7.1 and PWT 8.1 present alternative measures of GDP to these variables. As we discussed in Section 2, PWT 7.1 also provides a variable, "rgdpl2", which uses edited national accounts growth rates to compute real GDP away from the benchmark year without incorporating any information about prices. PWT 8.1 provides a variable, "rgdpna", which is also based on national accounts growth rates instead of price data. Feenstra, Inklaar and Timmer (2015) describe "rgdpna" as "similar to the series 'rgdpl2' ... except that (i) 'rgdpl2' used the real growth rate of absorption from the national accounts of each country rather than the real growth rate of GDP." Additionally, PWT 8 presented real PPP-adjusted output-based measures of economic activity for the first time within the sequence of PWT datasets. The dataset now includes a new variable "rgdpo," which Feenstra, Inklaar and Timmer (2015) describe as useful for "comparing the productive capacity across countries and across years."

We have focused our analysis on the variables "rgdpch" and "rgdpe" because the "rgdpch" variable has been the one most widely used by the economics literature, while "rgdpe" is the closest variable in concept to "rgdpch" out of the PWT 8 real GDP measures. Moreover, the "rgdpch" and "rgdpe" line of variables in the PWT captures the dataset's distinctive approach of modifying national accounts growth rates by using information from one or more price surveys, while its "rgdpl2" and "rgdpna" line of variables does not, and is conceptually closer to the real GDP measures in the WDI. As we mentioned in Section 2, Feenstra, Inklaar and Timmer (2015) recommend using "rgdpna" over "rgdpe" for pure time-series comparisons of growth rates, stating that "rgdpna" measures are closest to national accounts growth rates. However, Feenstra, Inklaar and Timmer (2015) also suggest that the conceptual foundation for "rgdpe" is stronger than it is for "rgdpna," as the former is expressed in prices that are constant over both time and space, while the latter is calculated only in constant national prices. In this subsection, we explore how well these alternative measures of economic activity explain nighttime lights.

Table V presents estimates of optimal weights on different measures of economic activity in the best unbiased linear estimator on true income, as derived from a regression of nighttime lights on these measures. Column 1 shows the results of the comparison between "rgdpna" from PWT 8.1 and "rgdpch" from PWT 7.1. As in our previous analysis, for the specifications without country fixed effects, PWT 7.1 "rgdpch" receives a much higher and statistically significant optimal weight (0.86), although we cannot reject at 5% the null hypothesis that the weights are equal. However, when we add country fixed effects, PWT 8.1 "rgdpna" now has the much higher optimal weight (1.91), which is statistically significantly different from 0 (as well as from

0.5) with 95% confidence, which is consistent with Feenstra, Inklaar and Timmer (2015)'s recommendation to use this variable to measure growth rates. Column 2 shows the results of a similar comparison between PWT 8.1 "rgdpchna" and PWT 7.1 "rgdpl2," the variable in PWT 7.1 that is conceptually closest to "rgdpna" in using national accounts growth rates alone. The results are almost identical to Column 1; the PWT 7.1 variable receives the higher weight in specifications without country fixed effects, while the PWT 8.1 variable receives the higher weight in specifications with fixed effects. Therefore, while the PWT 8.1 variable "rgdpe" appears to predict nighttime lights worse than PWT 7.1 "rgdpch" does, other variables in PWT 8.1 have improved over PWT 7.1 in predicting growth rates of nighttime lights.[11]

In fact, we can show that there are good reasons to prefer "rgdpna" over "rgdpe" among the PWT 8.1 variables. Column 3 presents optimal weights on "rgdpna" and "rgdpe" in a best unbiased linear predictor of true income. Not surprisingly, "rgdpna" receives the much larger weight whenever country fixed effects are included. However, when country fixed effects are excluded, "rgdpna" still receives the higher weight (around 0.6), although this weight is not significantly different from 0.5 (or from 0) with 95% confidence. Hence, the variable "rgdpna" may have better cross-sectional performance than "rgdpe," and may be better to use for cross-sectional, as well as for pure time-series comparisons of economic activity.

Lastly, we compare the capacity of expenditure-based ("rgdpe") and output-based ("rgdpo") measures of economic activity in PWT 8.1 to explain nighttime lights. Column 4 presents the resulting optimal weights. We see that in all specifications (with and without country fixed effects), the expenditure-based measure "rgdpe" has the larger weight, which is statistically significantly different from 0 in all specifications, and statistically different from 0.5 in all specifications but the country and year fixed effects one. This may not be surprising, as nighttime lights do not measure productivity alone, but rather consumption.

5 Results 2: Is Newer Better for the WDI?

The Penn World Tables are not the only source of comprehensive panel data on PPP-adjusted GDP. The World Development Indicators (WDI), published by the World Bank, also report PPP-adjusted GDP data, with the panel starting a bit later than the Penn World Tables (in the late 1970s and 1980s for most countries). Unlike the PWT, the WDI does not attempt to extrapolate PPPs to nonbenchmark years, but computes all GDPs away from the benchmark year using domestic growth rates (Johnson et al. 2009).

Have the WDI vintages been getting better over time, or have their revisions also produced setbacks,

[11] The superiority of national accounts growth rate based Laspeyres indices (like "rgdpna" or "rgdpl2") over chained indices (like "rgdpch") in predicting nighttime lights growth does not generalize to previous vintages of the PWT. Appendix Table AI presents optimal weights on chained and Laspeyres indices for every vintage of the PWT. We observe that for multiple vintages (though not PWT 8.1), the chained index receives a larger weight than the Laspeyres one does. (The two tend to be very similar in the cross section, and tend to receive statistically indistinguishably different weights in analyses without country fixed effects).

like with the PWT? Given that the WDI, unlike the PWT, performs minimal extrapolation of PPP rates, it is the preferred dataset to use to examine changes in the PPPs induced by different rounds of the ICP over the last two decades. Particularly interesting in this regard is the incorporation of the results from ICP 2005 price survey, which was the first to include China, and which led to a substantial increase in the estimates of developing world prices (and hence, a substantial decrease to developing world PPP-adjusted GDP). One criticism of this price survey round was that it did not use prices from rural China, which are lower than prices in Chinese cities, and which may have contributed to the findings of higher prices. The subsequent price survey round, ICP 2011, used prices from both urban and rural parts of China and largely reversed this price increase. On the other hand, ICP 2005 was the first price survey round to include many developing countries, and may have provided a better picture of the heterogeneity and interrelation of prices in the developing world.

We present estimates of optimal weights and their bootstrapped standard errors for comparisons between different vintages of the WDI and different vintages of the PWT in Table VI. Each cell of the table corresponds to a bivariate regression of log nighttime lights per capita on the logs of the two measures specified, with various combinations of fixed effects as controls. Column 1 compares the WDI with 2011 PPP to the WDI with 2005 PPP. We see that WDI 2011 gets nearly full weight (and we cannot reject a weight of unity) relative to WDI 2005 when we do not use country fixed effects (therefore, looking at cross-country income differences as well as growth rates), but that WDI 2005 gets more weight than WDI 2011 (neither of the weights being statistically different from zero) if country fixed effects are included. The same pattern can be observed in the comparison of WDI 2011 with WDI 1996, and of WDI 2005 with WDI 1996 (in columns 2 and 3 respectively). Each time, the newer WDI vintage gets full or nearly full weight in the regressions without country fixed effects, and the older WDI vintage gets a larger (or a much larger) weight in the regressions with country fixed effects. Since the WDI does not use PPPs to calculate GDP growth rates, it is reasonable to look only at the specifications without country fixed effects if we wish to assess the overall impact of the PPP adjustments. Then, our result is clear: each successive wave of PPP adjustments following ICP price surveys has, on net, been an improvement. However, the later vintages of the WDI are only unambiguously better than their predecessors if we are interested in calculating levels of economic activity, rather than economic growth.

6 Results 3: Should We Use the PWT or the WDI?

Johnson et al. (2009) pose the question of whether the PWT extrapolation procedure generates so much error that it might be worthwhile to use the WDI, and answer it in the affirmative on theoretical grounds.

One of the contributions of our paper is that under our assumption, we can use the nighttime lights data to determine whether it is better to use a given WDI vintage or a given PWT vintage without having to appeal to detailed hypotheses about the likely provenance of the measurement errors. Columns 4 through 9 compare WDI 2011 and WDI 2005 with the three most recent vintages of the PWT: 7.1, 8.0 and 8.1. We selected these vintages for comparison because they are the most recent vintages and because PWT 7.1 appeared to outperform PWT 8.0 and PWT 8.1 in the tests of Table IIB. We observe that in all these regressions, the WDI receives a larger weight than does the PWT. Out of the 24 tests in these columns, only 2 (when comparing WDI 2005 and PWT 7.1 without fixed effects) do not reject a unit weight on the PWT, while none of these tests reject a unit weight on the WDI. The two tests mentioned in the previous sentence are also the only ones in which the point estimates suggest relatively equal weights for the PWT and for the WDI (55% on the WDI and 45% on the PWT). Hence, if one has to make a choice, it is best to use the latest vintage of the WDI rather than the latest vintage of the PWT (or any of the three last vintages of the PWT, for that matter). Since we find that WDI 2011 (which uses the 2011 ICP) outperforms PWT 7.1 (which uses 2005 ICP) in column 4, while WDI 2005 (which also uses 2005 ICP) does not (column 7), it is reasonable to guess that part of the explanation for the advantage of recent WDI vintages over recent PWT vintages lies in the latest vintage of the WDI already incorporating price data from the most recent ICP. However, it is still better to use WDI 2005 than either of the two latest vintages of the PWT, and it is appears to be marginally better to use WDI 2005 than PWT 7.1 (although we cannot reject that using either one is equivalent). This finding roughly confirms the intution of Johnson et al (2009) that the WDI may be superior to the PWT.

In light of our discussion of alternative measures of economic activity in PWT 8.1, it is worth asking whether the WDI outperforms measures of economic activity in the PWT that are based on national accounts growth rates alone. We have seen that PWT 8.1 "rgdpna" improves over PWT 7.1 "rgdpch" in measuring growth rates. Columns 5 and 6 of Table V present comparisons of PWT 8.1 "rgdpchna" to WDI 2011 and WDI 2005, respectively. In all specifications, the WDI variables have a weight higher than unity (although it is typically not statistically significantly different from 0.5, or, when country fixed effects are included, from zero), consistent with our conclusion in the previous paragraph.

6.1 Performance of the PWT and WDI against Market Exchange Rates

The rationale for creating PPP-adjusted measures of economic activity such as the PWT and the WDI is the idea that market exchange rates do not accurately reflect the prices of many goods (in particular, nontradeable goods) that are important in an economy. If nighttime lights are good measures of economic activity, then we should see that regressing log nighttime lights per capita on log GDP per capita at PPP and log GDP per capita at market exchange rates produces a high partial correlation of nighttime lights with the PPP-adjusted measure and a low one with GDP per capita at market exchange rates. However, the methodology of PPP adjustment has been best developed for cross sections, with the problem of how best to combine very incomplete information on prices across time still being open. When comparing PPP-adjusted growth rates with growth rates measured with market exchange rates, the noise coming from PPP adjustment might dominate the bias coming from the use of market exchange rates rather than actual prices.

In Table VII we present optimal weights on various PPP-adjusted measures of log GDP per capita and on log GDP per capita at market exchange rates with and without country and year fixed effects. When country fixed effects are excluded, all PPP-adjusted measures of GDP per capita except the WDI in 1996 receive a weight of greater than 0.5 (and statistically significantly larger than 0.5 in 13 out of the 22 such specifications). However, when country fixed effects are included (so that we are essentially looking at growth rates), the PPP-adjusted measures from the PWT chain series receive a much lower weight than does GDP per capita at market exchange rates. On the other hand, the PPP-adjusted measures from the WDI, and the national accounts-based measure "rgdpna" in PWT 8.1 all receive a much higher weight than does GDP per capita at market exchange rates (which receives a negative weight, usually less than -1, and sometimes statistically significant). These results suggest that the approach to calculating growth rates taken by the WDI improves over using market exchange rates, while the approach taken in all vintages of the PWT has not.[12]

7 Results 4: Should We Use More than One Series?

While head-to-head comparisons of different GDP series are instructive, it is obvious that the optimal way to compute GDP may be to combine several different GDP series into a single proxy. All of the tests that we have done so far have essentially asked how we can best combine two GDP series to form such a proxy. But why should we stop at using only two series?

We are not the first to suggest using multiple GDP series to form an optimal combination to measure GDP. Johnson et al. (2009) innovatively suggest estimating GDP in a particular year using the PWT vintage for which that year is the benchmark year (the year of the price survey underlying the GDP estimates). For years that are not benchmark years, Johnson et al. (2009) suggest chaining the price indices of the preceding and the succeeding benchmark year. The procedure envisioned by Johnson et al. (2009) would therefore use

[12] While these results are for the variables "rgdpch" and "rgdpe" in the PWT, using the Laspeyres measures "rgdpl" and "rgdpl2" produces qualitatively similar results except for PWT 6.1, where "rgdpl" receives a weight of about 0.6

GDP vintages unedited in their benchmark years, while interpolating between them away from benchmark years.

Armed with the nighttime lights data, we can assess the procedure of Johnson et al. (2009). In Table VIII, we look at the optimal combination of the latest created vintages in each benchmark year. There are four such vintages: PWT 6.1 (for 1996), PWT 6.2 (for 2000, a partial update of the 1996 price survey), PWT 7.1 (for 2005) and PWT 8.1 (for the multiple benchmark year approach now used for the PWT). Given the hypothesis of Johnson et al. (2009), the optimal weight should be close to unity for each vintage near its benchmark year, and decline for years that are far away from its benchmark year. (Alternatively, the optimal weight should be unity for PWT 8.1 because it already implements chaining of prices across benchmark years). We estimate optimal weights for the four PWT vintages year-by-year for most years in our sample (including all the benchmark years). We observe that PWT 7.1 receives a fairly large and statistically significant weight in every year that we consider, including those that are far away from its benchmark year, 2005. PWT 6.2 also receives a large weight, but it tends not to be significantly different from zero for many years (including 2000, its benchmark year). PWT 8.1 always receives a statistically insignificant, or a negative weight. Finally, PWT 6.1 always receives a statistically insignificant weight that is equal to a very small 0.17 in its benchmark year, 1996.

The logic of Johnson et al. (2009) would suggest that interpolating between the different vintages of the WDI would also be an improvement (although this is not a suggestion made in their paper). We investigate this hypothesis in Table IX, in which we include all the three WDI series in the regression. Each series corresponds to its benchmark year, or the year of the price survey that was used to construct the estimates. Notably, the estimates from PPP 2011 have already been incorporated into the WDI, so the benchmark years to consider are 1996, 2005 and 2011 (the modification of the 1996 data in 2000 apparently not having been incorporated into a WDI series). We see an even starker result. In all years, including those that are far from 2011, the WDI 2011 series receives a weight that is large (never less than 0.75) and statistically insignificantly different from unity. All the other series receive much smaller weights (never in excess of 0.25), and their confidence intervals always contain zero and never contain unity. Hence, using the latest WDI series appears to strongly dominate using a weighted average of the preceding series.

Therefore, while Johnson et al. (2009) appear to be right in claiming that incorporating historical data may be useful in improving present estimates, it seems that assigning relatively time-constant weights to each series is better than assigning shifting weights that are functions of the benchmark year. In particular, for the WDI, there is strong evidence that using the current vintage is the optimal approach. For the PWT, a weighted average of several vintages (in particular, PWT 7.1 and PWT 6.2) may be better than any individual vintage, but the weights should not vary over time except after PWT 6.2 is discontinued.

8 Conclusion

In this paper, we show that if nighttime lights are an independent measure of output, they can act as an impartial referee between two other measures of output (Pinkovskiy and Sala-i-Martin 2016). We use this result to assess the quality of different ways of computing PPP-adjusted GDP per capita. Our investigation yields several conclusions, both for the methodologist attempting to improve existing measures of GDP, and for the practitioner deciding on which of the existing GDP series to use.

First, we find that the latest vintage of the Penn World Tables (PWT8) is not the best available measure of GDP. Instead, the updating process of the Penn World Tables has generally not involved linear improvement of the dataset, and has in particular produced a deterioration since vintage 7.1 in measuring cross-country differences in levels of economic activity. Moreover, the latest vintage of the World Development Indicators better explains variations in both the level and the growth rate of nighttime lights than does PWT8.

Second, we learn that each successive price survey of the International Comparisons Programme has generally led to better estimates of PPP-adjusted GDP per capita, including the controversial 2005 round. In particular, successive versions of the World Development Indicators have shown steady improvement in measuring differences in income levels across countries (however, it may be the case that older versions of the WDI capture economic growth better than newer versions do).

Third, and more fundamentally, we show that so long as nighttime lights growth rates can be taken as unbiased predictors of the growth rates of true income, the best way to measure growth rates of true income is to use growth rate series based on the national accounts alone, and without the adjustments employed by almost all the versions of the Penn World Tables. The GDP series most successful in explaining nighttime lights growth rates (the WDI and PWT 8.1 "rgdpna") compute growth rates based on national accounts alone, and without incorporating any information about PPPs. On the other hand, the series that goes furthest in attempting to construct a GDP index in prices that were truly constant across space and over time (PWT 8.1 "rgdpe") is less useful in explaining nighttime lights growth than any of the above series.

Finally, our results also provide guidance for researchers seeking a GDP series to use, for example in investigating the determinants of cross-country income differences, or in modeling the macroeconomy. One way to select such a series from among the most recent ones (those based on 2005 PPP or later) is to ask whether there is a unique series that, in every head-to-head comparison against another recent series and for every set of country and year fixed effects, either receives the larger (and statistically significant) weight in the optimal linear combination of the two, or, if it receives the smaller weight, neither of the weights are statistically significantly different from zero. Based on this criterion, we recommend using the WDI 2011 series. When compared in the cross section against PWT 7.1, PWT 8 (the "rgdpe" series as well as the "rgdpna" series), or WDI 2005, it receives a larger and statistically significant weight (which always is above 0.77). When country fixed effects are included in the comparison, WDI 2011 receives the larger weight (which is always greater than unity) except when compared against WDI 2005, but for that comparison, none of the weights are significant. Hence, if one seeks a single GDP series to measure both levels and growth rates of economic activity, the best candidate appears to be WDI 2011.

Author Affiliations

Maxim Pinkovskiy, Federal Reserve Bank of New York

Xavier Sala-i-Martin, J. and M. Grossman Professor of Economics, Columbia University and NBER

References

[1] **Adcock, R. J.** "A Problem in Least Squares." *Analyst.* 5 (1878), pp. 53-4.

[2] **Barro, Robert J.** "Economic Growth in a Cross Section of Countries." *Quarterly Journal of Economics* 106 (1991), no. 2: 407-443.

[3] **Barro, Robert J.** "Human Capital and Growth in Cross-Country Regressions." *Swedish Economic Policy Review* 6 (1999), no. 2: 237-277.

[4] **Barro, Robert J., and Xavier Sala-i-Martin**. "Convergence." *Journal of Political Economy* 100 (1992), no. 2: 223-251.

[5] **Center for International Earth Science Information Network (CIESIN), Columbia University; and Centro Internacional de Agricultura Tropical (CIAT)**. Gridded Population of the World, Version 3 (GPWv3). Palisades, NY: Socioeconomic Data and Applications Center (SEDAC), Columbia University. Available at http://sedac.ciesin.columbia.edu/gpw. 2005.

[6] **Chen, Xi, and William D. Nordhaus**. "Using Luminosity Data as a Proxy for Economic Statistics" Proceedings of the National Academy of Sciences. 2011

[7] **Ciccone, Antonio and Marek Jarocinski**. "Determinants of Economic Growth: Will Data Tell?" *American Economic Journal: Macroeconomics* 2 (2010):4, 222-246.

[8] **Doll, Christopher N.H.** "CIESIN Thematic Guide to Night-time Light Remote Sensing and its Applications." Manuscript. 2008

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

[10] **Elvidge, Christopher D. and Kimberly E. Baugh, Eric A. Kihn, Herbert W. Kroehl, Ethan R. Davis**. "Mapping City Lights With Nighttime Data from the DMSP Operational Linescan System." *Photogrammetric Engineering & Remote Sensing* 63 (1997), no. 6: 727-734

[11] **Elvidge, Christopher D., Kimberly E. Baugh, John B. Dietz, Theodore Bland, Paul C. Sutton and Herbert W. Kroehl.** "Radiance Calibration of DMSP-OLS Low-Light Imaging Data of Human Settlements." *Remote Sensing of Environment*, 68 (1999) no. 1:77-88.

[12] **Elvidge, Christopher D., Kimberly E. Baugh, Sharolyn. J. Anderson, Paul. C. Sutton, and Tilottama Ghosh**. "The Night Light Development Index (NLDI): A Spatially Explicit Measure of Human Development from Satellite Data." *Social Geography* 7 (2012), 23–35.

[13] **Feenstra, Robert C., Robert Inklaar** and **Marcel P. Timmer**, "The Next Generation of the Penn World Table," American Economic Review 105 (2015). vol.10, 3150-3182

[14] **Fuller, Wayne**. *Measurement Error Models.* New York: John Wiley & Sons. 1987.

[15] **Ghosh, Tilottama, Rebecca L. Powell, Christopher D. Elvidge, Kimberly E. Baugh, Paul C. Sutton and Sharolyn Anderson**. "Shedding Light on the Global Distribution of Economic Activity." *The Open Geography Journal* 3 (2010), 148-161.

[16] **Griliches, Zvi**. "Economic Data Issues," in *Handbook of Econometrics, Volume III.* Z. Griliches and M. D. Intriligator, eds. Amsterdam, New York and Oxford: North-Holland, pp. 1465–514. 1986.

[17] **Hausman, Jerry A.** "Mismeasured Variables in Econometric Analysis: Problems from the Right and Problems from the Left." *Journal of Economic Perspectives*, Vol. 15 (2001), No. 4, pp. 57-67.

[18] **Henderson, J. Vernon, Adam Storeygard, and David N. Weil**. "Measuring Economic Growth from Outer Space." *American Economic Review* 102 (2012), no. 2: 994-1028.

[19] **Johnson, Simon, William Larson, Chris Papageorgiou, and Arvind Subramanian**. "Is Newer Better? Penn World Table Revisions and Their Impact on Growth Estimates." *Journal of Monetary Economics* 60 (2013), no. 2: 255-274.

[20] **Baugh, Kimberly E., Christopher D. Elvidge, Tilottama Ghosh and Daniel Ziskin**. "Development of a 2009 Stable Lights Product using DMSP-OLS Data." *Proceedings of the 30th Asia-Pacific Advanced Network Meeting.* 2009.

[21] **Michalopoulos, Stelios, and Elias Papaioannou**. "Pre-colonial Ethnic Institutions and Contemporary African Development." *Econometrica* 81 (2013), no. 1: 113-152.

[22] **Michalopoulos, Stelios, and Elias Papaioannou**. "National Institutions and Sub-national Development in Africa." *Quarterly Journal of Economics* 129 (2014) no. 1:151-213.

[23] **National Geophysical Data Center (NGDC)**. "Version 4 DMSP-OLS Nighttime Lights Time Series." National Oceanic and Atmospheric Administration. http://www.ngdc.noaa.gov/dmsp/downloadV4composites.html. 2010. (accessed September 2013).

[24] **Pinkovskiy, Maxim L.** "Economic Discontinuities at Borders: Estimation Using Satellite Data of Lights at Night." mimeo, Federal Reserve Bank of New York. 2014.

[25] **Pinkovskiy, Maxim L.** and **Xavier Sala-i-Martin**. "Lights, Camera, ... Income! Illuminating the National Accounts-Household Surveys Debate," *Quarterly Journal of Economics*, forthcoming 2016.

[26] **Summers, Robert,** and **Alan Heston**. "The Penn World Table (Mark 5): an expanded set of international comparisons, 1950-1987." NBER Working Paper R1562.1991.

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

[28] **World Bank**. *World Development Indicators.* Washington, DC. 2006, 2014, 2015.

Tables

Table I (I)

Baseline Univariate Regressions — *Dependent Variable is Log Light Intensity per Capita*	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	PWT 6.1	PWT 6.2	PWT 6.3	PWT 7.0	PWT 7.1	PWT 8.0	PWT 8.1	WDI 1996 PPP	WDI 2005 PPP	WDI 2011 PPP
No Fixed Effects										
Log GDP per Capita	1.192***	1.152***	1.129***	1.001***	1.002***	.989***	1.012***	1.135***	1.053***	1.058***
	(.053)	(.051)	(.050)	(.047)	(.044)	(.050)	(.054)	(.058)	(.042)	(.047)
Within R2	.77	.72	.71	.72	.72	.69	.68	.69	.74	.75
Year Fixed Effects										
Log GDP per Capita	1.189***	1.156***	1.135***	1.009***	1.007***	.996***	1.024***	1.139***	1.058***	1.066***
	(.054)	(.052)	(.051)	(.047)	(.045)	(.050)	(.054)	(.058)	(.042)	(.048)
Within R2	.77	.72	.72	.73	.73	.70	.68	.70	.74	.76
Country Fixed Effects										
Log GDP per Capita	.706***	.464***	.330**	.271**	.448***	.327***	.321***	.474***	.582***	.483***
	(.133)	(.168)	(.140)	(.112)	(.129)	(.102)	(.095)	(.124)	(.118)	(.116)
Within R2	.13	.08	.06	.05	.11	.10	.10	.07	.17	.12
Country and Year Fixed Effects										
Log GDP per Capita	.407***	.353**	.335*	.314*	.374**	.170	.170	.536***	.525***	.471***
	(.122)	(.180)	(.191)	(.176)	(.184)	(.113)	(.114)	(.176)	(.177)	(.178)
Within R2	.05	.04	.05	.04	.06	.02	.02	.08	.11	.09
Number of Obs.	1287	2199	2916	3287	3484	3280	3280	2304	3562	3313
Number of Clusters	165	182	184	184	184	164	164	166	179	178

Table I presents estimates for the regressions of log nighttime lights per capita on measures of log GDP per capita at PPP, as described in Section 4. Standard errors in parentheses are clustered by country. Data on nighttime lights from the NOAA, data on measures of log GDP per capita from the Penn World Tables and from the World Development Indicators.

Table IIA

(IIA)

Baseline Regressions: New Vintages of PWT

Dependent Variable is Log Light Intensity per Capita

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	PWT 8.1 vs. PWT 8.0	PWT 8.0 vs. PWT 7.1	PWT 7.1 vs. PWT 7.0	PWT 7.0 vs. PWT 6.3	PWT 6.3 vs. PWT 6.2	PWT 6.2 vs. PWT 6.1	PWT 8.1 vs. PWT 7.1
No Fixed Effects							
New PWT Log GDP per Capita	-2.096***	.129	.716	.585***	.616**	.748*	-.012
	(.758)	(.256)	(.450)	(.224)	(.241)	(.386)	(.274)
Old PWT Log GDP per Capita	3.007***	.859***	.287	.490*	.546**	.443	.991***
	(.717)	(.255)	(.461)	(.252)	(.246)	(.390)	(.263)
Within R2	.71	.73	.72	.73	.73	.78	.73
Year Fixed Effects							
New PWT Log GDP per Capita	-1.998**	.140	.666	.587***	.615**	.734*	.013
	(.835)	(.256)	(.465)	(.224)	(.241)	(.389)	(.276)
Old PWT Log GDP per Capita	2.915***	.854***	.345	.493*	.550**	.455	.973***
	(.788)	(.255)	(.476)	(.252)	(.246)	(.392)	(.264)
Within R2	.71	.73	.73	.74	.73	.78	.73
Country Fixed Effects							
New PWT Log GDP per Capita	.404	-.040	.332	.002	.401*	.345	-.016
	(.352)	(.095)	(.206)	(.216)	(.207)	(.258)	(.091)
Old PWT Log GDP per Capita	-.086	.473***	-.026	.327*	.109	.382***	.449***
	(.389)	(.115)	(.183)	(.195)	(.181)	(.128)	(.134)
Within R2	.10	.12	.06	.06	.09	.13	.12
Country and Year Fixed Effects							
New PWT Log GDP per Capita	.008	-.084	.343	.050	.337	.183	-.066
	(.354)	(.094)	(.215)	(.248)	(.217)	(.202)	(.096)
Old PWT Log GDP per Capita	.161	.434***	.018	.287	.066	.204*	.414**
	(.351)	(.159)	(.186)	(.189)	(.170)	(.113)	(.165)
Within R2	.02	.06	.05	.05	.05	.04	.06
Number of Obs.	3280	3104	3287	2916	2198	1258	3104
Number of Clusters	164	164	184	184	182	163	164

Table IIA presents estimates for the regressions of log nighttime lights per capita on estimates of log GDP per capita at PPP from each vintage of the PWT and from its preceding vintage, as described in Section 4. Standard errors in parentheses are clustered by country. Data on nighttime lights from the NOAA, data on measures of log GDP per capita from the Penn World Tables and from the World Development Indicators.

Table IIB

(IIB)

Optimal Weights: New Vintages of PWT

Dependent Variable is Log Light Intensity per Capita

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	PWT 8.1 vs. PWT 8.0	PWT 8.0 vs. PWT 7.1	PWT 7.1 vs. PWT 7.0	PWT 7.0 vs. PWT 6.3	PWT 6.3 vs. PWT 6.2	PWT 6.2 vs. PWT 6.1	PWT 8.1 vs. PWT 7.1
No Fixed Effects							
New PWT Log GDP per Capita	-2.30***	.13	.71	.54**	.52***	.62*	-.01
	(-4.2,-.63)	(-.38,.54)	(-.21,1.5)	(.16,1.08)	(.07,1.08)	(-.03,1.2)	(-.55,.47)
Old PWT Log GDP per Capita	3.30***	.86***	.28	.45	.47*	.37	1.01***
	(1.6,5.2)	(.45,1.3)	(-.56,1.2)	(-.08,.83)	(-.08,.92)	(-.24,1.03)	(.52,1.5)
Year Fixed Effects							
New PWT Log GDP per Capita	-2.18***	.14	.65	.54**	.52***	.61*	.01
	(-4.3,-.46)	(-.38,.54)	(-.29,1.5)	(.16,1.08)	(.07,1.07)	(-.05,1.2)	(-.54,.49)
Old PWT Log GDP per Capita	3.18***	.85***	.34	.45*	.47*	.38	.98***
	(1.4,5.3)	(.45,1.3)	(-.55,1.2)	(-.08,.83)	(-.07,.92)	(-.23,1.05)	(.50,1.5)
Country Fixed Effects							
New PWT Log GDP per Capita	1.27	-.09	1.08	.00	.78	.47*	-.03
	(-.70,3.5)	(-.54,.34)	(-.68,3.3)	(-2.4,.99)	(-.18,1.9)	(-.04,.86)	(-.48,.35)
Old PWT Log GDP per Capita	-.27	1.09***	-.08	.99**	.21	.52**	1.03***
	(-2.5,1.7)	(.65,1.5)	(-2.3,1.6)	(.00,3.4)	(-.90,1.1)	(.13,1.04)	(.64,1.4)
Country and Year Fixed Effects							
New PWT Log GDP per Capita	.05	-.24	.94	.14	.83	.47	-.19
	(-13.1,4.1)	(-1.3,.28)	(-.63,4.7)	(-4.4,1.5)	(-.79,4.5)	(-.59,1.1)	(-1.1,.32)
Old PWT Log GDP per Capita	.94	1.24***	.05	.85	.16	.52*	1.19***
	(-3.1,14.1)	(.71,2.3)	(-3.7,1.6)	(-.52,5.4)	(-3.5,1.7)	(-.11,1.5)	(.67,2.1)
Number of Obs.	3280	3104	3287	2916	2198	1258	3104
Number of Clusters	164	164	184	184	182	163	164

Table IIB presents estimates for the optimal weights within a best unbiased linear predictor of log unobserved true income using log GDP per capita at PPP from each vintage of the PWT and from its preceding vintage, as described in Section 4. We provide country-block-bootstrapped 95% confidence intervals for each weight in parentheses below the estimate. Data on nighttime lights from the NOAA, data on measures of log GDP per capita from the Penn World Tables and from the World Development Indicators.

Table III

(III)

PWT 8.1 vs. PWT 7.1: Robustness Checks

Dependent Variable is Log Light Intensity per Capita unless otherwise noted

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Baseline	Additional Covariates		Different Dep. Var.				
		All Controls	Nonlinear Controls	Light Density	Calibrated Lights	Fraction Pop. Lit.	Weighted	Bench Mark
No Fixed Effects								
Log GDP per Capita PWT 8.1	-.01	-.02	-.00	.11	-.05	-.01	-.28	-.05
	(-.55,.47)	(-.79,.46)	(-.71,.43)	(-.34,.61)	(-.85,.59)	(-.55,.50)	(-.56,.67)	(-1.01,.41)
Log GDP per Capita PWT 7.1	1.01***	1.02***	1.00***	.88***	1.05***	1.01***	1.28***	1.05***
	(.52,1.5)	(.53,1.7)	(.56,1.7)	(.38,1.3)	(.40,1.8)	(.49,1.5)	(.32,1.5)	(.58,2.0)
Year Fixed Effects								
Log GDP per Capita PWT 8.1	.01	.00	.02	.11	-.10	-.08	-.25	-.01
	(-.54,.49)	(-.79,.49)	(-.66,.46)	(-.35,.62)	(-.98,.57)	(-.67,.47)	(-.52,.66)	(-.95,.43)
Log GDP per Capita PWT 7.1	.98***	.99***	.97***	.88***	1.10***	1.08***	1.25***	1.01***
	(.50,1.5)	(.50,1.7)	(.53,1.6)	(.37,1.3)	(.42,1.9)	(.52,1.6)	(.33,1.5)	(.56,1.9)
Country Fixed Effects								
Log GDP per Capita PWT 8.1	-.03	.26*	.22	.26**	-.04	.42***	-.54	-8.05
	(-.48,.35)	(-.01,.63)	(-.03,.61)	(.01,.57)	(-.26,.23)	(.16,1.01)	(-.88,1.06)	(-23.7,22.9)
Log GDP per Capita PWT 7.1	1.03***	.73***	.77***	.73***	1.04***	.57*	1.54*	9.05
	(.64,1.4)	(.36,1.01)	(.38,1.03)	(.42,.98)	(.76,1.2)	(-.01,.83)	(-.06,1.8)	(-21.9,24.7)
Country and Year Fixed Effects								
Log GDP per Capita PWT 8.1	-.19	.18	.16	.09	-.44	-.09	-.44	.05
	(-1.1,.32)	(-.40,.70)	(-.24,.61)	(-.51,.60)	(-2.4,.12)	(-11.7,6.7)	(-.83,.67)	(-.68,1.7)
Log GDP per Capita PWT 7.1	1.19***	.81***	.83***	.90**	1.44***	1.09	1.44**	.94
	(.67,2.1)	(.29,1.4)	(.38,1.2)	(.39,1.5)	(.87,3.4)	(-5.7,12.7)	(.32,1.8)	(-.79,1.6)
Number of Obs.	3104	2901	2901	3104	3096	775	3104	1049
Number of Clusters	164	159	159	164	164	158	164	158

Table III presents estimates for the optimal weights within a best unbiased linear predictor of log unobserved true income using log GDP per capita at PPP from PWT 8.1 and the same variable from PWT 7.1, as described in Section 4. We provide country-block-bootstrapped 95% confidence intervals for each weight in parentheses below the estimate. Columns 2 and 3 include log fractions of WDI GDP constituted by consumption, gross capital formation, government spending, exports and imports as controls in the regression equation (13), which is used to construct the optimal weights. Columns 4, 5 and 6 use log nighttime light density per square kilometer, a calibrated-parameter aggregator of nighttime lights and the fraction of the population of a country residing in lit areas as dependent variables in the regression equation (13). Column 7 weights each country-year by its population. Column 8 estimates the baseline specification over countries and years with benchmark or interpolated price data. Data on population from the Gridded Population of the World database.

Table IV (IV)

Weights in the Optimal Proxy: Subsamples *Dependent Variable is Log Light Intensity per Capita*											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Baseline	Excluding Regions					Only Regions				
		No Rich World	No Asia	No Latin America	No Post Comm.	No Africa	Rich World	Asia	Latin America	Post Comm.	Africa
No Fixed Effects											
PWT 8.1	-.01	.04	.30	-.03	-.12	-.32	-.52	-1.24***	.53	1.10***	.40
	(-.55,.47)	(-.61,.43)	(-.13,.63)	(-1.03,.55)	(-1.07,.33)	(-1.2,.39)	(-6.1,18.5)	(-2.9,-.24)	(-3.1,4.0)	(.42,1.6)	(-.67,.82)
PWT 7.1	1.01***	.95***	.69***	1.03***	1.12***	1.32***	1.52	2.24***	.46	-.10	.59***
	(.52,1.5)	(.56,1.6)	(.36,1.1)	(.44,2.0)	(.66,2.0)	(.60,2.2)	(-17.5,7.1)	(1.2,3.9)	(-3.0,4.1)	(-.69,.57)	(.17,1.6)
Year Fixed Effects											
PWT 8.1	.01	.07	.32	.00	-.09	-.30	-.55	-1.31**	.52	1.06**	.42
	(-.54,.49)	(-.57,.45)	(-.12,.64)	(-1.01,.58)	(-1.08,.36)	(-1.2,.39)	(-24.4,7.3)	(-3.2,-.17)	(-5.3,7.1)	(.36,1.6)	(-.61,.82)
PWT 7.1	.98***	.92***	.67***	.99***	1.09***	1.30***	1.55	2.31***	.47	-.06	.57***
	(.50,1.5)	(.54,1.5)	(.35,1.1)	(.41,2.0)	(.63,2.0)	(.60,2.2)	(-6.3,25.4)	(1.1,4.2)	(-6.1,6.3)	(-.63,.63)	(.17,1.6)
Country Fixed Effects											
PWT 8.1	-.03	-.04	.09	-.05	-.24	.12	-50.29	-.17	.38*	1.52***	-.66
	(-.48,.35)	(-.58,.46)	(-.82,.82)	(-.53,.42)	(-.98,.14)	(-.19,.47)	(-56.1,126.4)	(-.59,.15)	(-.03,.87)	(1.00,2.7)	(-23.2,6.0)
PWT 7.1	1.03***	1.04***	.90**	1.05***	1.24***	.87***	51.29	1.17***	.61***	-.52**	1.66
	(.64,1.4)	(.53,1.5)	(.17,1.8)	(.57,1.5)	(.85,1.9)	(.52,1.1)	(-125.4,57.1)	(.84,1.5)	(.12,1.03)	(-1.7,-.00)	(-5.0,24.2)
Country and Year Fixed Effects											
PWT 8.1	-.19	-.18	-.13	-.17	-.66	.08	-1.51	-.19	-.07	.96***	-1.48
	(-1.1,.32)	(-2.2,.43)	(-2.3,1.08)	(-3.3,.52)	(-9.3,3.6)	(-.28,.35)	(-83.1,13.8)	(-2.4,.24)	(-1.2,.61)	(.46,2.6)	(-9.9,8.1)
PWT 7.1	1.19***	1.18***	1.13*	1.17**	1.66*	.91***	2.51	1.19**	1.07**	.03	2.48
	(.67,2.1)	(.56,3.2)	(-.08,3.3)	(.47,4.3)	(-2.6,10.3)	(.64,1.2)	(-12.8,84.1)	(.75,3.4)	(.38,2.2)	(-1.6,.53)	(-7.1,10.9)
Number of Obs.	3104	2591	2439	2534	2584	2268	513	665	570	520	836
Number of Clusters	164	137	129	134	136	120	27	35	30	28	44

Table IV presents estimates for the optimal weights within a best unbiased linear predictor of log unobserved true income using log GDP per capita at PPP from PWT 8.1 and the same variable from PWT 7.1, as described in Section 4. We provide country-block-bootstrapped 95% confidence intervals for each weight in parentheses below the estimate. Columns 2 through 6 perform estimation excluding the listed region. Columns 7 through 11 perform estimation only on the listed region. Data on nighttime lights from the NOAA, data on measures of log GDP per capita from the Penn World Tables and from the World Development Indicators

Table V (V)

Optimal Weights: Other GDP Variables in PWT8.1 *Dependent Variable is Log Light Intensity per Capita*						
	(1)	(2)	(3)	(4)	(5)	(6)
	PWT 8.1 NA vs. PWT 7.1 RGDPCH	PWT 8.1 NA vs. PWT 7.1 Laspeyres	PWT 8.1 NA vs. PWT 8.1 RGDPE	PWT 8.1 RGDPE vs. PWT 8.1 RGDPO	PWT 8.1 NA vs. WDI 2011	PWT 8.1 NA vs. WDI 2005
No Fixed Effects						
PWT 8.1 NA (or RGDPE)	.13	.12	.62*	1.65***	-.00	-.14
	(-.47,.56)	(-.50,.56)	(-.02,1.1)	(.70,2.3)	(-.43,.44)	(-1.5,.50)
Other GDP Series	.86***	.87***	.37	-.65	1.00***	1.14***
	(.43,1.4)	(.43,1.5)	(-.12,1.02)	(-1.3,.29)	(.55,1.4)	(.49,2.5)
Year Fixed Effects						
PWT 8.1 NA (or RGDPE)	.14	.13	.59*	1.64***	.00	-.13
	(-.45,.57)	(-.49,.57)	(-.04,1.09)	(.60,2.3)	(-.39,.45)	(-1.5,.54)
Other GDP Series	.85***	.86***	.40	-.64	.99***	1.13***
	(.42,1.4)	(.42,1.4)	(-.09,1.04)	(-1.3,.39)	(.54,1.3)	(.45,2.5)
Country Fixed Effects						
PWT 8.1 NA (or RGDPE)	1.91***	1.79**	1.17***	1.16***	-.32	-.57
	(.65,2.8)	(.49,2.7)	(.73,1.5)	(.55,1.8)	(-1.8,1.3)	(-1.8,1.03)
Other GDP Series	-.91	-.79	-.17	-.16	1.32	1.57*
	(-1.8,.34)	(-1.7,.50)	(-.50,.26)	(-.87,.44)	(-.38,2.8)	(-.03,2.8)
Country and Year Fixed Effects						
PWT 8.1 NA (or RGDPE)	1.91***	1.79***	1.26***	1.31**	-.46	-.89
	(.73,3.3)	(.49,3.2)	(.78,1.7)	(.26,3.1)	(-3.3,1.4)	(-4.6,.77)
Other GDP Series	-.91	-.79	-.26	-.31	1.46	1.89**
	(-2.3,.26)	(-2.2,.50)	(-.78,.21)	(-2.1,.73)	(-.42,4.3)	(.22,5.6)
Number of Obs.	3104	3104	3280	3280	2987	3225
Number of Clusters	164	164	164	164	160	162

Table V presents estimates for the optimal weights within a best unbiased linear predictor of log unobserved true income using various measures of log GDP per capita at PPP from PWT 8.1, PWT 7.1 and the World Development Indicators, as described in Section 4.1. We provide country-block-bootstrapped 95% confidence intervals for each weight in parentheses below the estimate.

Table VI (VI)

Optimal Weights: Comparisons of PWT and WDI *Dependent Variable is Log Light Intensity per Capita*									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	WDI, 2011 vs. WDI, 2005	WDI, 2011 vs. WDI, 1996	WDI, 2005 vs. WDI, 1996	WDI, 2011 vs. PWT 7.1	WDI, 2011 vs. PWT 8.0	WDI, 2011 vs. PWT 8.1	WDI, 2005 vs. PWT 7.1	WDI, 2005 vs. PWT 8.0	WDI, 2005 vs. PWT 8.1
No Fixed Effects									
Newer / WDI Measure	.94***	1.06***	1.18***	.85***	.77***	.92***	.54	.93***	1.14***
	(.43,1.3)	(.77,1.3)	(.87,1.7)	(.49,1.2)	(.49,1.02)	(.61,1.1)	(-.18,1.3)	(.48,1.4)	(.70,1.7)
Older / PWT Measure	.05	-.06	-.18	.14	.22*	.07	.45	.06	-.14
	(-.35,.56)	(-.38,.22)	(-.73,.12)	(-.20,.50)	(-.02,.50)	(-.15,.38)	(-.30,1.1)	(-.42,.51)	(-.71,.29)
Year Fixed Effects									
Newer / WDI Measure	.95***	1.05***	1.18***	.86***	.77***	.90***	.54	.92***	1.11***
	(.44,1.3)	(.77,1.3)	(.86,1.7)	(.50,1.2)	(.48,1.02)	(.59,1.1)	(-.17,1.2)	(.47,1.4)	(.67,1.6)
Older / PWT Measure	.04	-.05	-.18	.13	.22*	.09	.45	.07	-.11
	(-.36,.55)	(-.38,.22)	(-.72,.13)	(-.20,.49)	(-.02,.51)	(-.12,.40)	(-.29,1.1)	(-.41,.52)	(-.69,.32)
Country Fixed Effects									
Newer / WDI Measure	.21	-.89	-1.32	1.80***	1.17***	1.13***	1.84***	1.17***	1.15***
	(-1.5,2.0)	(-2.8,1.2)	(-2.8,.93)	(.97,3.1)	(.83,1.8)	(.79,1.7)	(.88,3.1)	(.84,1.6)	(.82,1.5)
Older / PWT Measure	.78	1.89*	2.32**	-.80	-.17	-.13	-.84*	-.17	-.15
	(-1.06,2.5)	(-.23,3.8)	(.06,3.8)	(-2.1,.02)	(-.88,.16)	(-.73,.20)	(-2.1,.11)	(-.65,.15)	(-.59,.17)
Country and Year Fixed Effects									
Newer / WDI Measure	.30	-.81	-1.12	1.93***	1.24***	1.19***	2.00***	1.26***	1.24***
	(-1.9,3.0)	(-2.6,.99)	(-2.9,.93)	(.94,4.2)	(.86,2.5)	(.81,2.4)	(.89,5.1)	(.87,2.1)	(.84,2.1)
Older / PWT Measure	.69	1.81**	2.12**	-.93	-.24	-.19	-1.00*	-.26	-.24
	(-2.0,2.9)	(.00,3.6)	(.06,3.9)	(-3.2,.05)	(-1.5,.13)	(-1.4,.18)	(-4.1,.10)	(-1.1,.12)	(-1.1,.15)
Controls									
Number of Obs.	3230	2239	2290	3301	2987	2987	3380	3225	3225
Number of Clusters	174	163	165	178	160	160	179	162	162

Table VI presents estimates for the optimal weights within a best unbiased linear predictor of log unobserved true income using two measures of PPP-adjusted log GDP per capita, as described in Sections 5 and 6. The newer / WB measure is the one listed first in the column heading, while the older / PWT measure is the one that is listed second. We provide country-block-bootstrapped 95% confidence intervals for each weight in parentheses below the estimate.

Table VII

(VII)

Optimal Weights: Market Exchange Rates vs. PPP-Adjustment

Dependent Variable is Log Light Intensity per Capita

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	PWT 6.1	PWT 6.2	PWT 6.3	PWT 7.0	PWT 7.1	PWT 8.0	PWT 8.1	PWT 8.1 NA	WDI 1996 PPP	WDI 2005 PPP	WDI 2011 PPP
No Fixed Effects											
PPP-Adjusted Dataset	.80***	.86***	.82***	1.04***	1.10***	.77***	.66***	.72***	.39	1.38***	1.17***
	(.51,1.08)	(.61,1.1)	(.43,1.1)	(.67,1.4)	(.72,1.4)	(.48,1.1)	(.35,.98)	(.30,1.2)	(-.08,.78)	(1.04,1.6)	(.95,1.3)
Market Exchange Rates	.19	.13	.17	-.04	-.10	.22	.33**	.27	.60***	-.38**	-.17
	(-.08,.48)	(-.12,.38)	(-.16,.56)	(-.45,.32)	(-.43,.27)	(-.10,.51)	(.01,.64)	(-.22,.69)	(.21,1.08)	(-.63,-.04)	(-.38,.04)
Year Fixed Effects											
PPP-Adjusted Dataset	.79***	.85***	.83***	1.06***	1.11***	.79***	.70***	.74***	.41	1.40***	1.19***
	(.50,1.07)	(.60,1.1)	(.44,1.1)	(.71,1.4)	(.74,1.4)	(.50,1.1)	(.38,1.01)	(.31,1.2)	(-.06,.79)	(1.06,1.6)	(.97,1.3)
Market Exchange Rates	.20	.14	.16	-.06	-.11	.20	.29*	.25	.58***	-.40**	-.19*
	(-.07,.49)	(-.12,.39)	(-.16,.55)	(-.48,.28)	(-.44,.25)	(-.11,.49)	(-.01,.61)	(-.24,.68)	(.20,1.06)	(-.64,-.06)	(-.39,.02)
Country Fixed Effects											
PPP-Adjusted Dataset	.59	.25	.02	-.53	-.14	-.05	.00	1.80*	2.09***	2.30**	3.25**
	(-.17,1.08)	(-.41,1.02)	(-1.2,1.4)	(-4.8,.83)	(-1.2,.78)	(-.72,.27)	(-.55,.31)	(-.25,3.6)	(.16,2.8)	(.35,3.8)	(.45,4.4)
Market Exchange Rates	.40	.74*	.97	1.53**	1.14**	1.05***	.99***	-.80	-1.09	-1.30	-2.25
	(-.08,1.1)	(-.02,1.4)	(-.47,2.2)	(.16,5.8)	(.21,2.2)	(.72,1.7)	(.68,1.5)	(-2.6,1.2)	(-1.8,.83)	(-2.8,.64)	(-3.4,.54)
Country and Year Fixed Effects											
PPP-Adjusted Dataset	.76	.23	.06	-.20	-.35	-.19	-.14	1.73*	2.06***	2.41**	3.32***
	(-.22,1.5)	(-.69,2.0)	(-2.1,1.9)	(-4.3,2.6)	(-3.9,.94)	(-1.6,.20)	(-1.7,.21)	(-.14,4.0)	(.30,2.9)	(.23,4.5)	(.48,5.0)
Market Exchange Rates	.23	.76	.93	1.20*	1.35**	1.19***	1.14***	-.73	-1.06	-1.41	-2.32
	(-.56,1.2)	(-1.09,1.6)	(-.97,3.1)	(-1.6,5.3)	(.05,4.9)	(.79,2.6)	(.78,2.7)	(-3.0,1.1)	(-1.9,.69)	(-3.5,.76)	(-4.0,.51)
Number of Obs.	1259	2125	2813	3178	3369	3203	3203	3203	2262	3463	3281
Number of Clusters	162	180	183	183	183	164	164	164	166	179	178

Table VI presents estimates for the optimal weights within a best unbiased linear predictor of log unobserved true income using a measure of PPP-adjusted log GDP per capita and log GDP per capita at market exchange rates, as described in Section 6.1. We provide country-block-bootstrapped 95% confidence intervals for each weight in parentheses below the estimate. Data on GDP at market exchange rates from the World Development Indicators.

Table VIII

(VIII)

Weights in the Optimal Proxy Each Year: PWT — *Dependent Variable is Log Light Intensity per Capita*	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	1992	1994	1995	1996	1999	2000	2001	2004	2005	2006	2008	2010
PWT 8.1 (multiple)	-.45*	-.46**	-.40	-.34	-.40*	-.28	-.36	.11	.27	.18	.12	.28
	(-.98,.05)	(-.87,-.03)	(-.83,.03)	(-.72,.14)	(-1.01,.01)	(-1.04,.14)	(-1.3,.26)	(-.99,.68)	(-.62,.76)	(-.64,.70)	(-.72,.62)	(-.55,.70)
PWT 7.1 (2005)	.66***	.44***	.42**	.57***	.62***	.62***	.84***	.88***	.72***	.81***	.87***	.71***
	(.23,1.1)	(.15,.85)	(.10,.86)	(.21,1.07)	(.27,1.1)	(.30,1.2)	(.33,1.8)	(.31,1.9)	(.23,1.6)	(.29,1.6)	(.37,1.7)	(.29,1.5)
PWT 6.2 (2000)	.38	.85**	.71	.58	.77**	.50	.51**	0***	0***	0***	0***	0***
	(-.26,1.3)	(.03,1.4)	(-.40,1.5)	(-.37,1.5)	(.26,1.3)	(-.18,1.3)	(.13,.95)					
PWT 6.1 (1996)	.41	.15	.26	.17	-.00	.15	0***	0***	0***	0***	0***	0***
	(-.67,1.1)	(-.69,.95)	(-.90,1.2)	(-1.1,1.2)	(-.82,.59)	(-.93,.81)						
Number of Obs.	126	135	139	155	132	128	161	164	164	164	164	164

Table VIII presents estimates for the optimal weights within a best unbiased linear predictor of log unobserved true income using the latest measures of PPP-adjusted log GDP per capita for each benchmark year from the PWT, as described in Section 6. Each regression is run restricting the sample to the year in the column heading. We provide country-block-bootstrapped 95% confidence intervals for each weight in parentheses below the estimate. Data on nighttime lights from the NOAA, data on measures of log GDP per capita from the Penn World Tables.

Table IX

(IX)

Weights in the Optimal Proxy Each Year: WDI — *Dependent Variable is Log Light Intensity per Capita*	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	1992	1994	1995	1996	1999	2000	2001	2004	2005	2006	2008	2010
WDI 2011	1.20***	1.20***	.88***	.84***	1.00***	1.01***	1.07***	1.10***	1.19***	1.09***	.75**	.77**
	(.67,1.7)	(.44,1.9)	(.33,1.3)	(.26,1.3)	(.40,1.5)	(.47,1.4)	(.59,1.4)	(.69,1.5)	(.67,1.5)	(.67,1.6)	(.25,1.3)	(.27,1.1)
WDI 2005	-.08	-.05	.17	.25	.12	.05	-.07	-.19	-.15	-.09	.24	.22
	(-.61,.51)	(-.68,.76)	(-.35,.82)	(-.32,.96)	(-.39,.84)	(-.42,.72)	(-.61,.45)	(-.75,.42)	(-.67,.46)	(-.61,.32)	(-.35,.74)	(-.17,.72)
WDI 1996	-.12	-.14	-.05	-.10	-.12	-.06	.00	.08	-.04	0***	0***	0***
	(-.43,.19)	(-.51,.20)	(-.50,.22)	(-.59,.19)	(-.53,.16)	(-.42,.21)	(-.38,.29)	(-.36,.50)	(-.36,.25)			
Number of Obs.	156	158	160	160	161	161	161	159	151	173	173	171

Table IX presents estimates for the optimal weights within a best unbiased linear predictor of log unobserved true income using the latest measures of PPP-adjusted log GDP per capita for each ICP survey year from the World Development Indicators, as described in Section 6. Each regression is run restricting the sample to the year in the column heading. We provide country-block-bootstrapped 95% confidence intervals for each weight in parentheses below the estimate. Data on nighttime lights from the NOAA, data on measures of log GDP per capita from the World Development Indicators.

Figures[13]

Figure I (I)

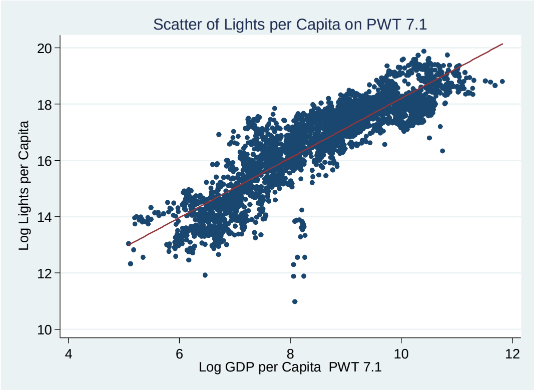
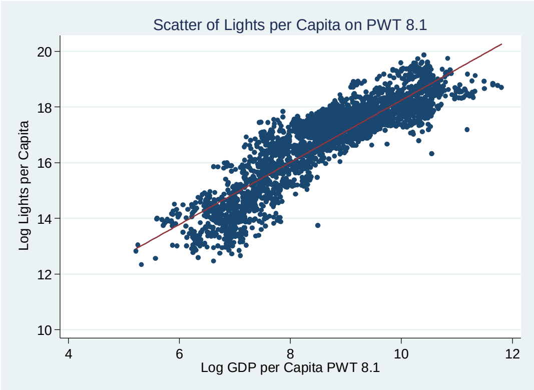


Figure II (II)



[13] All figures exclude Zimbabwe, El Salvador, Nigeria, Syria, Yemen, Maldives, Bermuda, Equatorial Guinea and Singapore because these countries generate outliers that interfere with the visual presentation of the data. Excluding or including these countries does not affect the estimated weights. See footnote 8 for a more detailed description.

Figure III (III)

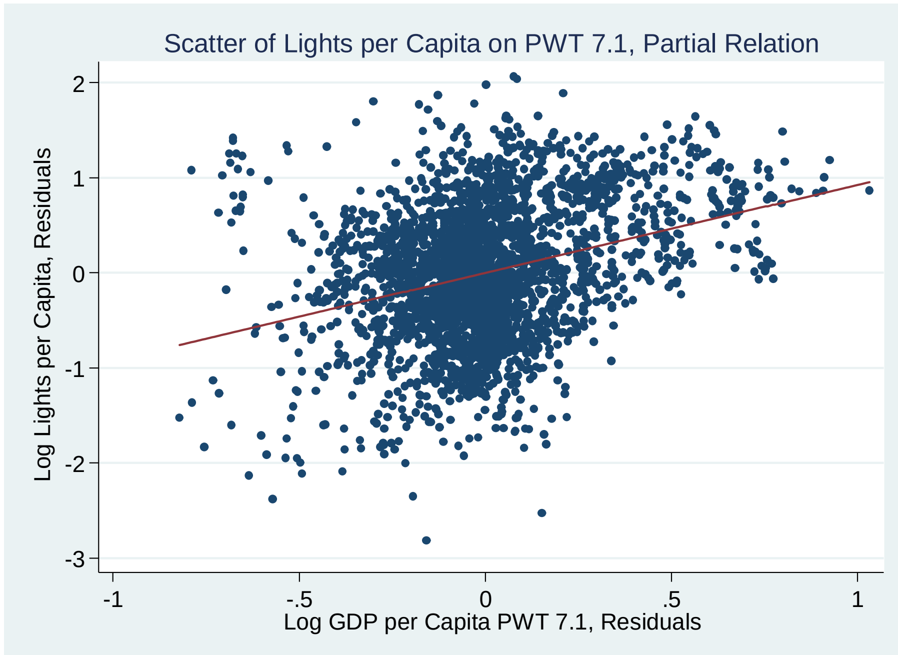
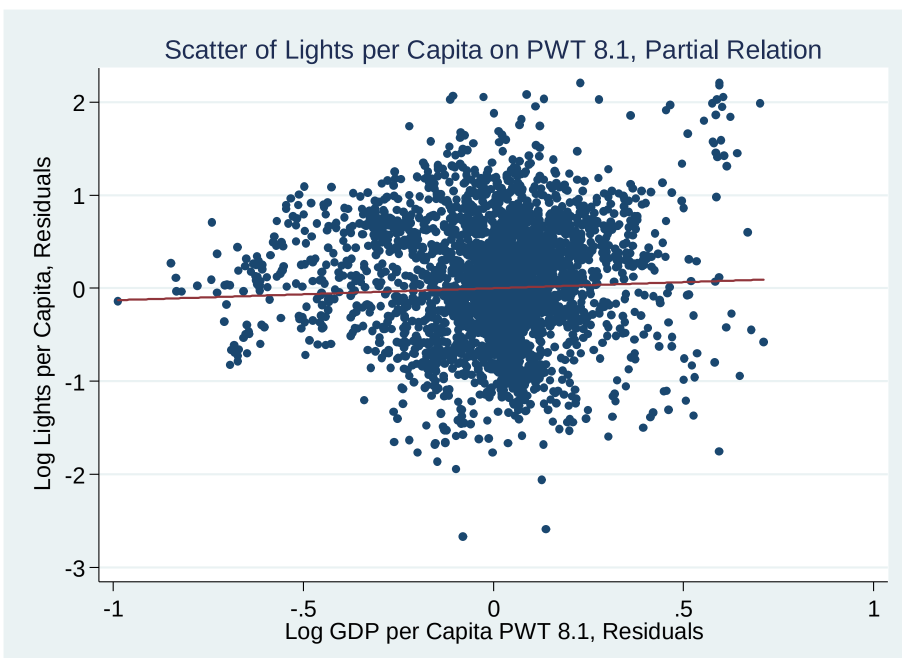


Figure IV (IV)



Appendix Table AI

(AI)

Optimal Weights: Chain Series vs. Laspeyres *Dependent Variable is Log Light Intensity per Capita*										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	PWT 6.1 vs. L1	PWT 6.2 vs. L1	PWT 6.3 vs. L1	PWT 7.0 vs. L1	PWT 7.1 vs. L1	PWT 6.3 vs. L2	PWT 7.0 vs. L2	PWT 7.1 vs. L2	PWT 8.0 vs. L2	PWT 8.1 vs. L2
No Fixed Effects										
Chain Series	-15.09	12.93	.77	5.98	2.77	.83	3.04	-.20	.21	.37
	(-81.3,97.7)	(-17.2,44.1)	(-4.5,18.8)	(-30.7,25.6)	(-35.6,34.2)	(-.58,2.9)	(-3.8,4.7)	(-4.9,3.2)	(-.28,.84)	(-.12,1.02)
Laspeyres	16.09	-11.93	.22	-4.98	-1.77	.16	-2.04	1.20	.78***	.62**
	(-96.7,82.3)	(-43.1,18.2)	(-17.8,5.5)	(-24.6,31.7)	(-33.2,36.6)	(-1.9,1.5)	(-3.7,4.8)	(-2.2,5.9)	(.15,1.2)	(-.02,1.1)
Year Fixed Effects										
Chain Series	-17.33	11.20	.66	5.83	3.62	.78	2.93	-.29	.22	.40
	(-85.6,101.7)	(-20.4,42.1)	(-5.3,18.3)	(-27.4,31.3)	(-34.7,37.2)	(-.67,2.9)	(-3.5,4.7)	(-4.8,3.2)	(-.27,.83)	(-.09,1.04)
Laspeyres	18.33	-10.20	.33	-4.83	-2.62	.21	-1.93	1.29	.77***	.59**
	(-100.7,86.6)	(-41.1,21.4)	(-17.3,6.3)	(-30.3,28.4)	(-36.2,35.7)	(-1.9,1.6)	(-3.7,4.5)	(-2.2,5.8)	(.16,1.2)	(-.04,1.09)
Country Fixed Effects										
Chain Series	17.21	35.76*	1.26	2.82	3.06	.80	.65	.23	-.20	-.17
	(-17.3,52.8)	(5.7,83.7)	(-10.6,16.4)	(-9.0,61.6)	(-32.9,24.4)	(-.84,5.5)	(-5.1,5.0)	(-2.8,3.2)	(-.55,.22)	(-.50,.26)
Laspeyres	-16.21	-34.76*	-.26	-1.82	-2.06	.19	.34	.76	1.20***	1.17***
	(-51.8,18.3)	(-82.7,-4.7)	(-15.4,11.6)	(-60.6,10.0)	(-23.4,33.9)	(-4.5,1.8)	(-4.0,6.1)	(-2.2,3.8)	(.77,1.5)	(.73,1.5)
Country and Year Fixed Effects										
Chain Series	34.52	36.86	.90	1.71	-1.44	.51	.45	.07	-.29	-.26
	(-30.9,113.9)	(-42.5,107.3)	(-11.6,33.8)	(-47.3,48.3)	(-67.7,27.3)	(-.93,7.6)	(-5.3,9.5)	(-4.6,4.4)	(-.88,.15)	(-.78,.21)
Laspeyres	-33.52	-35.86	.09	-.71	2.44	.48	.54	.92	1.29***	1.26***
	(-112.9,31.9)	(-106.3,43.5)	(-32.8,12.6)	(-47.3,48.3)	(-26.3,68.7)	(-6.6,1.9)	(-8.5,6.3)	(-3.4,5.6)	(.84,1.8)	(.78,1.7)
Number of Obs.	1287	2199	2916	3287	3484	2916	3287	3484	3280	3280
Number of Clusters	165	182	184	184	184	184	184	184	164	164

Table VI presents estimates for the optimal weights within a best unbiased linear predictor of log unobserved true income using two measures of PPP-adjusted log GDP per capita, as described in Section 5. The newer / WB measure is the one listed first in the column heading, while the older / PWT measure is the one that is listed second. We provide country-block-bootstrapped 95% confidence intervals for each weight in parentheses below the estimate. Data on nighttime lights from the NOAA, data on measures of log GDP per capita from the Penn World Tables and from the World Development Indicators.