NBER WORKING PAPER SERIES

SHINING A LIGHT ON PURCHASING POWER PARITIES

Maxim Pinkovskiy
Xavier Sala-i-Martin

Working Paper 24419
http://www.nber.org/papers/w24419

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2018

We are very grateful to Matthew Bosch, Andrew Haughwout, Robert Feenstra, Nicole Gorton, Robert Inklaar, Michelle Jiang, Chris Papageorgiou, Marcel Timmer, Wilbert van der Klaauw and seminar participants at the Federal Reserve Bank of New York and at the IMF for very helpful comments. We are very grateful to Noah Kwicklis and Rachel Schuh for superb research assistance. Any views expressed in this paper are the authors' and do not necessarily reflect those of the Federal Reserve Bank of New York, the Federal Reserve System, or the National Bureau of Economic Research. All errors are our own.

Shining a Light on Purchasing Power Parities
Maxim Pinkovskiy and Xavier Sala-i-Martin
NBER Working Paper No. 24419
March 2018
JEL No. A1,E01,F00

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 generate three findings in the study of PPP-adjusted estimates of GDP around the world between 1992 and 2010. First, we find that while market exchange rates described poor economies better than did PPP-adjusted estimates in the late 1990s (Dowrick and Akmal 2008; Almas 2012), this pattern has disappeared by the 2010s. Second, we also find that estimates of PPPs have been steadily improving from one price survey round to the next, including during the controversial 2005 and 2011 rounds. Third, we leverage this fact to assess whether it is optimal to measure relative prices as close as possible to the year of interest or to use the latest available relative price data and discard the rest, and provide a theoretical framework in which the latter may be optimal. Using data from the Penn World Tables, we find that, indeed, it is optimal to only use the latest price data, and hence, to revise existing PPP-adjusted estimates whenever a new price survey is released.

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

One of the defining achievements of economic measurement is the computation of purchasing power parities (PPPs) to take into account differences in prices between countries. The first estimates of GDP by country at PPP were constructed as part of the celebrated Penn World Table project (Kravis, Heston and Summers 1978, Summers and Heston 1991). Currently, PPP adjustment is employed by the leading sources of cross-country data, including the World Development Indicators as well as the Penn World Tables. The creation of what has been considered to be reliable PPP adjustment has fueled the cross-country growth literature (Barro 1991 and the vast literature following) as well as the measurement of world distribution of income (Chen and Ravallion 2001, Sala-i-Martin 2006).

However, there exist important concerns about PPP adjustment as it stands now. First, while PPP adjustment attempts to improve on market exchange rates by capturing the prices of nontraded goods, it introduces a variety of other biases. For example, price surveys used to compute PPPs are fundamentally limited in how well they can measure the underlying quality of transacted goods because this has to be assessed by the surveyors themselves (which is known as quality measurement bias), while market exchange rates incorporate quality differentials between traded goods via the revealed preference of millions of consumers. If the quality differences between tradeables and nontradeables are very small, then market exchange rates may better represent economic differentials than PPPs do. Moreover, as was realized by Kravis, Heston and Summers (1982), "valuation at other than own prices tends to inflate the aggregate value of the bundle of goods because no allowance is made for the substitutions in quantities toward the goods that are relatively cheap," which is known as substitution bias. Dowrick and Akmal (2008) argue that these biases warrant an alternative method of comparing the economies of different countries, and develop such metrics by combining purchasing power parities and market exchange rate data. Almas (2012), on the other hand, uses survey-based data on food consumption to compare PPP-adjusted estimates of GDP with the income levels that would rationalize food shares in typical Engel curves for food. She finds that while rich countries are well described by PPP-adjusted data, poor countries may be better described by estimates of GDP based on market exchange rates. While data limitations preclude Almas (2012) from looking at more than 32 countries (and only at 10 countries using her preferred methodology), her results point to the importance of assessing the heterogeneity in the performance of PPP adjustment across different types of countries, and further raise concerns about the quality of current PPP-adjustment procedures. Hence, a first question is: does adjusting for PPP provide better estimates of the underlying economy than using market exchange rates, especially for poor countries?

Second, there has been considerable volatility in PPPs. Purchasing power parities are computed by the International Comparisons Project (ICP) using international price surveys. These have largely focused on the developed world until 1996. For countries without survey data, prices were extrapolated using a regression-based procedure and not based on direct data from these countries. The state of affairs changed dramatically following the 2005 round of surveys, which was the first to include China and several other large developing countries. The new estimates suggested that prices in developing countries were higher (hence, the economies of these countries were smaller) than was previously thought. However, this conclusion was largely undone after the subsequent round of surveys in 2011. The instability of the PPP estimates does not on its own give us a clue whether the changes implemented in the 2005 or the 2011 surveys led to progress or regress. As Deaton and Aten (2017) write in their analysis of the difference between the two price survey rounds, "if ICP 2011 is inconsistent with ICP 2005, the problems could come from either ICP 2005 or ICP 2011, or both, or, in the case of ICP 2005, could have been inherited from 1993/1995." Deaton and Aten (2017) go further to state that "those who know the institutional history of the project believe that, through a host of methodological improvements and overall control, 2011 was a clear improvement over the two previous rounds," and suggest that the process of comparing prices across regions in the 2005 ICP may have artificially depressed developing world prices. However, they go on to say that "our results provide a plausible story for the conflict between the 2005 and 2011 ICP benchmarks...but the evidence is far from conclusive. There are many other changes from ICP 2005 to ICP 2011 that [our] hypothesis does not explain." While methodological considerations can suggest plausible hypotheses for which price estimates from a particular survey may be preferable to those from another, it is difficult to reach firm conclusions absent an independent benchmark for what the measurement goal may be. So a second question is: have successive rounds of price surveys generated improvements or deteriorations in our estimates of PPP?

In addition, the volatility of the successive price surveys raises a third question: should we use or discard past prices data when new data becomes available? The standard practice has been to revise both current and all past estimates of GDP once a new ICP price survey is conducted, essentially throwing out the price data from the previous rounds. This appears to be suboptimal. Johnson et al. (2013) propose an alternative approach, which assigns prices from each ICP survey to the benchmark year of the survey, and interpolates between prices in the benchmark years. A version of this approach has now been incorporated into the Penn World Table (as of version 8.0, see Feenstra, Inklaar and Timmer 2015) alongside the more conventional estimates based on the most recent PPP data. While Feenstra et al. (2015) caution that national accounts growth rates may be more accurate than the "PPP-adjusted" growth rates used in the multiple-benchmark approach, it remains unclear whether the multiple-benchmark PPPs might outperform the most recent PPPs in a cross section.

Our contribution in this paper is to answer these three questions using satellite-recorded nighttime lights as an independent benchmark for unobserved true income, following the methodology of a previous paper of ours, Pinkovskiy and Sala-i-Martin (2016). If we had a measurement of GDP whose error was uncorrelated with the measurement errors of the different PPP-adjusted GDPs, it would be a straightforward matter to see which set of PPPs was better by comparing them both to the independent measure. In Pinkovskiy and Sala-i-Martin (2016) we argued that such an independent measurement can be constructed using data on satellite-recorded nighttime lights, which were first studied by Elvidge et al. (1997, 1999, 2012) and in economics by Henderson, Storeygard and Weil (2012) and Chen and Nordhaus (2011). While errors in different versions of GDP 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.[1]

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. However, by looking at the patterns of which estimators appear to outperform others and in which samples, we can exclude hypotheses that would predict counterfactual patterns of estimator quality.

Our nighttime lights analysis gives clear-cut answers to the three questions above in ways that reconciles current practice with the concerns about PPP measurement discussed above. First, to test whether PPP-adjusted GDP is a better measure than GDP at market exchange rates, we run bivariate regressions of log nighttime lights on log GDP at PPP and log GDP at market exchange rates, and compare the coefficients on the right hand-side variables, which should be proportional to the optimal weights of these variables in the best unbiased linear estimator of log unobserved true income. We clearly conclude that, at least for the most recent (2011) PPP measures, PPP-adjusted GDP is a better measure of unobserved true income (correlates better with nighttime lights) than GDP at market exchange rates. However, and in accordance with Almas (2012), we observe that PPPs based on the 1996 price survey are a worse measure of unobserved true income

[1] Almas (2012) also compares PPP-adjusted GDP and GDP at market exchange-rates to a third series, namely the share of food in consumption.

for poor countries than GDP at market exchange rates is. After the 2005 and 2011 price survey rounds, PPP-adjusted GDP became a better measure of unobserved true income than GDP at market exchange rates both for rich and for poor countries alike.

To answer the second question, we regress log nighttime lights on log GDP at PPP and log GDP at a previous PPP, and compare the coefficients. We find that PPPs have been improving over time, at least for the last two rounds of ICP in 2005 and 2011. On the other hand, we find that the underlying national accounts data that undergoes PPP-adjustment has not been generally improving over time, and may even have deteriorated in some time periods on average. This is notwithstanding revisions to the national accounts data over time, some of which have been large, such as the recent Nigerian rebasing of its GDP (The Economist, 04/12/2014). While we undoubtedly have better measures of economic activity than we had fifteen years ago, most of the improvement has come from better estimates of PPPs.

Lastly, we answer the third question by regressing log nighttime lights on log Penn World Tables GDP at 2011 PPP and log Penn World Tables GDP at the synthetic PPP used in the multiple benchmark series of the PWT. We find that GDP at the latest PPP is a much better estimator of log unobserved true income than GDP at the multiple benchmark year PPP. This conclusion appears to be counterintuitive because we are finding that it is optimal to ignore data. However, it can be understood by the fact that the quality of the price data is improving over time, and the rate of the improvement may be fast enough to outweigh the magnitude of the fluctuations in prices. A straightforward model of these processes (in Section 3.2) also suggests that it is optimal to use the latest price data to measure prices in the earliest years because the measurement error problem is most significant for them, while contemporaneous price data in later years is optimal because its measurement error is fairly low, and less important to reduce relative to matching the annual variation in price data. This is exactly the pattern that we find in the data.

Our conclusion is somewhat dispiriting because, intuitively, one would have preferred a GDP series that would change continuously with additional data, rather than requiring revisions to long-ago observations whenever an update is made. For now, it appears that price survey methodology is continuously improving, and improving so rapidly that current estimates of prices now (or methods of their aggregation) may be superior estimates of prices in the past than were estimates of those past prices made in the past. Once our methods of estimating prices reach a steady state, it may become preferable to move to the approach of continuous variation of GDP estimates embodied in the multiple-benchmark series.

The rest of the paper is organized as follows. Section 2 describes the data. Section 3 presents the empirical approach, which is an extension of Pinkovskiy and Sala-i-Martin (2016), as well as a stylized model that explains our main empirical finding that using latest price data is superior to using contemporaneous price data. Sections 4, 5 and 6 present the results. Section 7 concludes.

2 Data

2.1 World Development Indicators

We use data on GDP at constant PPP-adjusted international dollars, at constant US dollars and at constant local currency units from every vintage in the archives of the World Development Indicators from April 2005 (the first vintage available in the archives) to May 2017. The PPP-adjusted international dollars are constructed using prices from the 1996 ICP until November 2007, using 2005 ICP between April 2008 and April 2014, and using the 2011 ICP since May 2014. Since November 2007, the World Development Indicators PPP-adjust GDP by multiplying the constant local currency GDP data in every year by a country-specific time-constant PPP; before November 2007, the PPP factor was allowed to vary over time.

2.2 Penn World Tables

We use vintages 8.0 and 9.0 of the Penn World Table. The current version of the Penn World Table (version 9) and its relationship to its predecessors is described in detail in Feenstra, Inklaar and Timmer (2015). Each of these vintages includes a measure of GDP similar to the WDI measures in that time-constant PPPs from the latest price survey are used to multiply series of national accounts data expressed in constant local currency (this variable is called "rgdpna" in the PWT, and will be referred to as the latest price series), as well as a measure of GDP for which the PPP adjustment factor is not time-constant, but instead is an interpolation of the price data available from all the ICP price surveys (this variable is called "rgdpe" in the PWT, and will be referred to as the "multiple-benchmark" series). The intent of the multiple-benchmark series was to take into account the recommendations of Johnson et al. (2013) and have a measure that is invariant to future price data, and would not change over time unless the underlying national accounts data is revised, a feature that was not present in earlier PWTs.

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.[2] The images are processed to remove cloud

[2] 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).

cover, snow and ephemeral lights (such as forest fires) to produce the final product.[3] The nighttime lights data is available from 1992 to 2013, and we use the data up to 2010 because many series of GDP per capita represent vintages of the WDI that were archived 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

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

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. Figure 1 of our paper show that there are very strong correlations between log nighttime lights and log GDP at market exchange rates, or in PPP-adjusted international dollars in the WDI.

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

[3] The nighttime lights data are publicly available for download as spatial geotiffs at http://www.ngdc.noaa.gov/dmsp/downloadV4composites.html

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.

The first two panels of Figure 1 present scatterplots of log aggregate nighttime lights computed using the formula (1) against log GDP in constant 2005 US dollars and constant 2011 PPP-adjusted international dollars, respectively, from the May 2014 vintage of the WDI.[4] The data covers the years 1992-2010 as discussed above. We see that the scatterplots lie very close to the regression lines drawn through them, with very few outliers. Therefore, we are confident that nighttime lights are strongly related to economic activity in a typical country.

2.4 Covariates

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 as follows:

GDP component shares

- Log Consumption share of GDP
- Log Capital formation as percent of GDP
- Log General government expenditure share of GDP
- Log shares of GDP in exports and imports

Population Shares

- Log total population
- Log rural population share
- Log urban population share

[4] The scatterplots are constructed over the baseline sample in the paper, which is described in Section 2.5

Area

- Log area (time-constant)

- Log arable area (time-constant)

Sectoral Shares

- Log agriculture share of GDP

- Log manufacturing share of GDP

- Log services share of GDP

All these covariates come from the World Development Indicators. In addition, we also use two variables from Dell, Jones and Olken (2014) to capture variation in climate, which may affect the perception of nighttime lights data by the satellites. These are

- Average temperature, population-weighted by 1-degree grid cell

- Average precipitation, population-weighted by 1-degree grid cell.

Dell, Jones and Olken (2014) compute a balanced panel of these data for all the countries in the world up to 2006; we extend their series by running regressions of each variable against year for each country, and generating regression forecasts for the years 2007-2010. We do not believe this procedure to entail a substantial loss in information as the cross-sectional variation in this variable is the main component that we wish to capture, since PPP adjustments are largely cross-sectional and all of our regressions will exclude country fixed effects.

2.5 Sample

We look at a panel of countries from 1992 to 2010 for which covariate data is available at least for some years, linearly interpolating and extrapolating any missing data.[5] The sample grows over time as GDP data for new countries are added to the World Development Indicators. By May 2014, we have 139 countries in the sample.

[5] The covariates include all the variables in Section 2.4, as well as log GDP at 2011 PPP from PWT 9.0 for the purpose of dividing countries into below- and above-median GDP per capita categories.

3 Mathematical Framework

3.1 Using Nighttime Lights to Recover Optimal Weights on GDP Vintages

Although the main contribution of this paper is applied and not theoretical, for the benefit of the reader, we give another exposition of the methodology of Pinkovskiy and Sala-i-Martin (2016) to determine which datasets do a better job of computing unobserved true income by using an independent benchmark. We also provide an alternative and more intuitive proof of why this methodology is valid. 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 unobserved true income $y^*_{i,t}$, which is the log total 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 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 ($y^L_{i,t}$) and on two different measurements of log GDP ($y^1_{i,t}$, $y^2_{i,t}$) for a sample of countries i and years t. For example, $y^1_{i,t}$ may be log GDP at PPP, while $y^2_{i,t}$ could be log GDP at market exchange rates. These data are related to log unobserved true income according to the following system of equations (partialling out constants and other covariates):

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

$$y^1_{i,t} = \beta^1 y^*_{i,t} + \varepsilon^1_{i,t} \tag{3}$$

$$y^2_{i,t} = \beta^2 y^*_{i,t} + \varepsilon^2_{i,t} \tag{4}$$

In other words, each of the measured variables is a linear function of log unobserved true income, perturbed by some error.[6] This framework is similar to the one used by Henderson, Storeygard and Weil

[6] In fact, we can substantially relax the functional form specification in equation 2 to read

$$y^L_{i,t} = f_{i,t}\left(y^*_{i,t}\right) + \varepsilon^L_{i,t}$$

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

(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. [7]

We assume that the error terms in all three processes are uncorrelated with true income. That is,

$$cov\left(y_{i,t}^{*}, \varepsilon_{i,t}^{L}\right) = cov\left(y_{i,t}^{*}, \varepsilon_{i,t}^{1}\right) = cov\left(y_{i,t}^{*}, \varepsilon_{i,t}^{2}\right) = 0 \tag{A1}$$

In a model where the lights equation is linear in $y_{i,t}^{*}$, this assumption always holds trivially because it is always possible to set the β^k's to the regression coefficients of each proxy on unobserved true income, and the $\varepsilon_{i,t}^{k}$'s to the empirical residuals:[8]

$$\begin{aligned} \beta^k &= \frac{cov\left(y_{i,t}^{k}, y_{i,t}^{*}\right)}{var\left(y_{i,t}^{*}\right)} \\ \varepsilon_{i,t}^{k} &= y_{i,t}^{k} - \beta^k y_{i,t}^{*}, \text{ for } k = 1, 2, L \end{aligned}$$

The second and critical assumption of this paper is that the error term in the lights equation (2), $\varepsilon_{i,t}^{L}$, is uncorrelated with the error terms from the measurement equations (3) and (4) conditional on true income:

$$E\left(\varepsilon_{i,t}^{1}\varepsilon_{i,t}^{L}|y_{i,t}^{*}\right) = E\left(\varepsilon_{i,t}^{2}\varepsilon_{i,t}^{L}|y_{i,t}^{*}\right) = 0 \tag{A2}$$

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 GDP are 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.

does not dominate the positive relation between nighttime lights and true income within our sample.

[7] 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.

[8] In a model where $y_{i,t}^{L} = f_{i,t}\left(y_{i,t}^{*}\right) + \varepsilon_{i,t}^{L}$, straightforward uncorrelatedness needs to be strengthened to mean independence:

$$E\left(\varepsilon_{i,t}^{L}|y_{i,t}^{*}\right) = E\left(\varepsilon_{i,t}^{1}|y_{i,t}^{*}\right) = E\left(\varepsilon_{i,t}^{2}|y_{i,t}^{*}\right) = 0 \ (A1')$$

and this assumption has content. However, given that it is always possible to construct errors that are uncorrelated with unobserved true income, this assumption should not be difficult to satisfy.

We are interested in finding the best unbiased linear predictor of log true income per capita ($y^*_{i,t}$) in terms of $y^1_{i,t}$ and $y^2_{i,t}$:

$$z_{i,t} = \gamma_1 y^1_{i,t} + \gamma_2 y^2_{i,t} \tag{5}$$

where γ_1 and γ_2 are the weights on each of the GDP measurements that we are considering:

Hence, we want to compute the vector (γ_1, γ_2), which minimizes the mean squared error

$$(\gamma^*_1, \gamma^*_2) = \arg \min_{(\gamma_1,\gamma_2)} E\left(\left(y^*_{i,t} - \gamma_1 y^1_{i,t} - \gamma_2 y^2_{i,t}\right)^2\right) \tag{6}$$

The minimand of this problem is the minimand of an ordinary least squares regression of unobserved true income $y^*_{i,t}$ on the proxies $y^1_{i,t}$ and $y^2_{i,t}$. This is intuitive because as a consequence of the Gauss-Markov theorem, ordinary least squares produces the best predictor of the dependent variable that is a linear function of the right hand-side variables. Since we do not have data on $y^*_{i,t}$, we cannot run this regression. However, if we run instead the regression

$$y^L_{i,t} = \ b_1 y^1_{i,t} + b_2 y^2_{i,t} \tag{7}$$

which is the regression of log lights per capita on all our proxies of log GDP per capita (and partialling out the constant as well as the same other covariates that may be part of equations 2, 3 and 4), we can compute the vector (γ^*_1, γ^*_2) up to a proportionality constant.

The regression coefficients from equation (7) solve the minimization problem:

$$(b^*_1, b^*_2) = \arg \min_{(b_1,b_2)} E\left(\left(y^L_{i,t} - b_2 y^1_{i,t} - b_2 y^2_{i,t}\right)^2\right) \tag{8}$$

Substituting in for $y^L_{i,t}$, we can restate this problem in terms of $y^*_{i,t}$. First, we obtain that

$$(b^*_1, b^*_2) = \arg \min_{(b_1,b_2)} E\left(\left(\beta^L y^*_{i,t} + \varepsilon^L_{i,t} - b_2 y^1_{i,t} - b_2 y^2_{i,t}\right)^2\right)$$

Then, we open the square to express the minimand as a function of three terms:[9]

[9] If we assume that

$$y^L_{i,t} = f_{i,t}\left(y^*_{i,t}\right) + \varepsilon^L_{i,t}$$

and make the mean independence assumption $A1'$ instead of assumption A1, we can get the same result by defining

$$\beta^L_{OLS} = \frac{cov\left(f_{i,t}\left(y^*_{i,t}\right), y^*_{i,t}\right)}{var\left(y^*_{i,t}\right)}$$

$$(b_1^*, b_2^*) = \arg\min_{(b_1,b_2)} \left[\left(\beta^L\right)^2 E\left(\left(y_{i,t}^* - \frac{b_2}{\beta^L} y_{i,t}^1 - \frac{b_2}{\beta^L} y_{i,t}^2\right)^2\right) + E\left(\left(\varepsilon_{i,t}^L\right)^2\right) + 2\beta^L E\left(\varepsilon_{i,t}^L \left(y_{i,t}^* - \frac{b_2}{\beta^L} y_{i,t}^1 - \frac{b_2}{\beta^L} y_{i,t}^2\right)\right)\right]$$

The second term in this minimand is independent of the variables over which minimization is conducted and may be omitted. The third term is equal to zero because

$$E\left(\varepsilon_{i,t}^L \left(y_{i,t}^* - \frac{b_2}{\beta^L} y_{i,t}^1 - \frac{b_2}{\beta^L} y_{i,t}^2\right)\right) = E\left(\varepsilon_{i,t}^L \left(1 - \frac{b_2\beta^1}{\beta^L} - \frac{b_2\beta^1}{\beta^L}\right) y_{i,t}^* - \varepsilon_{i,t}^L \left(\frac{b_2}{\beta^L}\varepsilon_{i,t}^1 + \frac{b_2}{\beta^L}\varepsilon_{i,t}^2\right)\right) = 0$$

by assumptions A1 and A2. Lastly, the coefficient $\left(\beta^L\right)^2$ in front of the first term is irrelevant for optimization and may be omitted. Therefore, the optimization problem becomes:

$$\begin{aligned} (b_1^*, b_2^*) &= \arg\min_{(b_1,b_2)} E\left(\left(y_{i,t}^* - \frac{b_2}{\beta^L} y_{i,t}^1 - \frac{b_2}{\beta^L} y_{i,t}^2\right)^2\right) \\ &= \beta^L \left(\gamma_1^*, \gamma_2^*\right) \end{aligned}$$

the last line following straightforwardly by identifying $b_1^*/\beta^L = \gamma_1^*$ and $b_2^*/\beta^L = \gamma_2^*$ from the maximization problem (6).

Since we do not know the proportionality constant β^L without further assumptions, γ_1^* and γ_2^* remain unidentified. However, we can identify the ratio of the optimal weights to their sum, which turns out to be

and noting that

$$\begin{aligned} & E\left(y_{i,t}^L - b_1 y_{i,t}^1 - b_2 y_{i,t}^2\right)^2 \\ = \; & E\left(f_{i,t}\left(y_{i,t}^*\right) - \beta_{OLS}^L y_{i,t}^* + \beta_{OLS}^L y_{i,t}^* + \varepsilon_{i,t}^L - b_1 y_{i,t}^1 - b_2 y_{i,t}^2\right)^2 \\ = \; & E\left(f_{i,t}\left(y_{i,t}^*\right) - \beta_{OLS}^L y_{i,t}^*\right)^2 \\ & +2E\left(f_{i,t}\left(y_{i,t}^*\right) - \beta_{OLS}^L y_{i,t}^*\right)\left(\beta_{OLS}^L y_{i,t}^* + \varepsilon_{i,t}^L - b_1 y_{i,t}^1 - b_2 y_{i,t}^2\right) \\ & +E\left(\beta_{OLS}^L y_{i,t}^* + \varepsilon_{i,t}^L - b_1 y_{i,t}^1 - b_2 y_{i,t}^2\right)^2 \end{aligned}$$

The first term of the minimand in the last line of the optimization does not depend on b_1 and b_2 and therefore is a constant that may be dropped. The second term is equal to zero because

$$\begin{aligned} & E\left(\left(f_{i,t}\left(y_{i,t}^*\right) - \beta_{OLS}^L y_{i,t}^*\right) y_{i,t}^*\right) \\ = \; & E\left(\left(f_{i,t}\left(y_{i,t}^*\right) - \beta_{OLS}^L y_{i,t}^*\right) \varepsilon_{i,t}^k\right) \\ = \; & 0 \end{aligned}$$

for $k = 1, 2, L$ by the definition of β_{OLS}^L and by the mean independence assumption $A1'$. Note that mean independence rather than uncorrelatedness has to be used to argue that

$$E\left(f_{i,t}\left(y_{i,t}^*\right) \varepsilon_{i,t}^k\right) = 0$$

The last term is the minimand considered in the text, and the proof proceeds analogously.

equal to the ratio of the regression coefficients to their sum:

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

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 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 (7) presenting estimates of the optimal relative weights on the measurements of GDP per capita $\frac{b^1}{b^1+b^2}$, $\frac{b^2}{b^1+b^2}$ when the elementary specifications in equations (2), (3) and (4) – and hence, regression equation (7) – are augmented by covariates, or when they are estimated on different samples.

3.2 Theoretical Considerations on Vintage Selection

Before beginning our empirical analysis, it is useful to consider conceptually in what situations an analyst might prefer to use current data in order to estimate prices in the past. We suppose that we desire to measure p_t, the true PPP in year t, for some location. We can think of the available price data from the various ICP surveys as representing a series

$$\hat{p}_t = p_t + \varepsilon_t$$

where p_t is the true PPP in year t and ε_t is measurement error in prices associated with the survey carried out at time t. For simplicity, we will assume that this measurement error is orthogonal to the realizations of the process of the true PPPs $\{p_\tau\}_{\tau=1}^{T}$.

When the analyst considers whether to use current or historical prices to estimate PPP in year t, she wishes to solve the problem

$$\min_{1 \leq \tau \leq T} E\left(p_t - \hat{p}_\tau\right)^2$$

The objective function of this problem can then be described by

$$E\left(p_t - \hat{p}_\tau\right)^2 = E\left(\varepsilon_\tau^2\right) + E\left(p_t - p_\tau\right)^2$$

in light of the independence assumption presented above. We further assume that the variance of the measurement error in prices is declining over time, which is equivalent in our setup to the price survey data

$\hat{p}_t$ improving in MSE over time. Hence,

$$E\left(\varepsilon_\tau^2\right) := \sigma_\varepsilon^2(\tau)\,,\ \frac{\partial}{\partial\tau}\sigma_\varepsilon^2(\tau) < 0$$

We also assume that the true PPP process satisfies

$$E\left(p_t - p_\tau\right)^2 = \sigma_p^2 \cdot 1\left(t \neq \tau\right) \tag{9}$$

The intuition for this formula is that using current price measurements to approximate historical prices is prima facie suboptimal because prices change over time. If the variance of the measurement error term $\sigma_\varepsilon^2(\tau)$ weren't shrinking over time, it would be optimal to use contemporaneous prices surveys to measure historical prices. A PPP process satisfying equation (9) is $p_t = \bar{p} + \pi_t$, were π_t is i.i.d.

In the stylized example outlined above, it is easy to see that when one is choosing between using the price survey conducted in year t and the price survey conducted in the last year (T), one should choose the latter whenever

$$\sigma_\varepsilon^2(T) < \sigma_\varepsilon^2(t) - \sigma_p^2$$

In particular, as long as $\sigma_\varepsilon^2(T) < \sigma_\varepsilon^2(1) - \sigma_p^2$, it is optimal to use $\hat{p}_T$ to measure p_t for all t less than a cutoff T^*, and to use $\hat{p}_t$ for all t greater or equal to that cutoff. So (as long as static MSE reduction is the objective) it should be optimal to use the latest available price data for the *earliest* years for which one wishes to compute PPPs, and use the current price data for the later years for which one is seeking to compute prices.[10]

4 Results 1: Comparison of PPPs and Market Exchange Rates

In this section we address our first question, which is whether PPP-adjusted GDP is a better predictor of economic activity than is GDP at market exchange rates. We recall that market exchange rates do not compare prices for nontraded goods, which comprise a large amount of economic activity. However, PPP-adjustment is often flawed because it does not often capture relevant quality differences of both traded and nontraded goods, and frequently involves pricing the consumption basket of one country with the prices of

[10] It is trivial to relax equation (9) to allow $E\left(\left(p_t - p_\tau\right)^2\right)$ to be $\sigma_p^2\phi\left(|t-\tau|\right)$, an increasing function of $|t-\tau|$ that is bounded above by σ_p^2 and that is equal to zero when $t=\tau$. The same condition as in the main text; $\sigma_\varepsilon^2(T) < \sigma_\varepsilon^2(1) - \sigma_p^2$ ensures that using $\hat{p}_T$ instead of $\hat{p}_t$ to measure p_t is optimal for all t at least up to a cutoff T^* that is weakly greater than the cutoff in the case that $\phi\left(|t-\tau|\right)$ takes the form of an indicator function. It also may become possible that the set of t such that using $\hat{p}_T$ instead of $\hat{p}_t$ is optimal may consist of several intervals in the space $t = \{1, 2, .., T\}$.

another country, which leads to well-known biases when comparing the two countries (quality and substitution bias). Dowrick and Akmal (2008) and Almas (2012) provide evidence that GDP at market exchange rates may be a good supplement to, or superior to PPP-adjusted data, at least for large subsets of countries.

4.1 Regression Coefficients

To understand whether we should use GDP at PPP or GDP at market exchange rates, our model in Section 3 suggests that we should look at partial correlations between these variables and log nighttime lights. Figure 1 presents scatterplots illustrating the raw and partial correlations between log aggregate nighttime lights (lights for short) and log GDP at 2011 PPP (2011 PPP for short), or log GDP at market exchange rates (MAR for short). The two log GDP variables are taken from the May 2014 World Development Indicators vintage, which is the first WDI vintage to use 2011 PPP. As we discussed in Section 2, it is obvious from graphs 1) and 2) of the figure that both 2011 PPP and MAR are very tightly correlated with lights. Now, graphs 3) and 4) present the partial correlations between lights and 2011 PPP or MAR, respectively, holding the remaining variable fixed. We now observe that there remains a strong partial correlation between lights and 2011 PPP, even holding MAR fixed. Hence, even conditional on the information contained in MAR, the value of 2011 PPP is helpful for predicting lights, and, by extension, log unobserved true income. On the other hand, once we condition on 2011 PPP, the lack of partial correlation between lights and MAR shows that, conditional on 2011 PPP, MAR is just noise that is unrelated to log unobserved true income. Therefore, when one is predicting log unobserved true income, one would not wish to include MAR whenever one can condition on 2011 PPP. We recall that the slopes of the scatterplots in the partial correlation graphs are just the regression coefficients in the multiple regression of lights on 2011 PPP and MAR.

More formally, in Table IA, we run the bivariate regression (7) of log nighttime lights on log GDP at PPP and log GDP at market exchange rates for PPPs using the 2011 PPP, the 1996 PPP and the 1996 PPP as modified by the Penn World Tables, 6.1. (We include the latter PPP because we wish to compare our results to Almas (2012), who uses data from PWT 6.1 to find substantial PPP bias for poor countries.) We run each regression three ways. The first panel just runs the regression without covariates (analogous to the scatterplots in Figure 1). Then, in the second panel, we present regressions with year fixed effects, which may capture changes in satellite sensitivity to light across years. In the third and last panel, we present regressions with both year fixed effects and the covariates described in Section 2.4.

Column (1) of the first panel of Table IA presents the regression corresponding to the scatterplots in Figure 1. We see that the regression coefficient on log GDP at 2011 PPP is 0.96 and statistically significantly different from zero at the 1% significance level, while the regression coefficient on log GDP at market exchange rates is 0.06 and statistically indistinguishable from zero at normal significance levels. The upper bound of the confidence interval for the coefficient on log GDP at market exchange rates is 0.28, suggesting at most a modest role for log GDP at market exchange rates in explaining log nighttime lights once log GDP at 2011 PPP has been controlled for. Including year fixed effects in Panel II changes little. While the satellites recording nighttime lights tend to be considerably more or less sensitive in some years rather than others, creating a putative role for year fixed effects, the cross-sectional variation engendered by cross-country income differences drives most of the findings and dwarfs any over-time errors due to satellite sensitivity.

As discussed in Sections 2.4 and 3.1, our analysis may be confounded if errors in nighttime lights are correlated with errors in GDP at PPP or market exchange rates. Particularly concerning sources of this correlation may be if output produced by different industries (such as agriculture vs. manufacturing) and settings (urban vs. rural), or allocated to different uses (such as consumption vs. investment) has different light intensity, and also differential measurement error in the national accounts. For example, the 2005 ICP price survey collected data on Chinese prices in the urban centers only, omitting the rural areas. If urban output is more light-intensive than rural output is, and if urban prices are higher than rural prices are, then the error in lights may be correlated with the error in the PPP-adjusted GDP data. Therefore, Panel III includes the fourteen controls described in Section 2.4, which extensively focus on sectoral composition, investment and government shares, the rural-urban population breakdown and climatic variables that may affect the perception of nighttime lights by satellites. Our inclusion of controls such as the log urban population share and the log agriculture share of GDP should preclude the omitted variables problem discussed above, at least on average and to first order. We see that including these covariates, if anything, increases the coefficient on log GDP at 2011 PPP, and turns the coefficient on log GDP at market exchange rates negative.

The picture changes when we look at GDP at 1996 PPP instead of 2011 PPP. Column (2) of Panel I shows that regressing log nighttime lights on log GDP at 1996 PPP (1996 PPP for short) and log GDP at market exchange rates (MAR for short; all data is from the November 2007 version of the World Development Indicators) generates coefficients on the two measures of log GDP that are both statistically significant and roughly equal (the coefficient on 1996 PPP is 0.53 and the coefficient on MAR is 0.47). We get similar results when we include year fixed effects (column 2, panel II) and when we include both year fixed effects and covariates, the coefficient on 1996 PPP becomes a statistically insignificant 0.16, while the coefficient on MAR becomes a statistically significant 0.53.

Almas (2012) uses data on Engel curves, market exchange rate data and PWT 6.1 PPP-adjusted GDP to argue for substantial PPP bias, especially for poor countries. We will return to the question of differential PPP bias over time for rich and poor countries in Section 4.3, but to make our main results comparable to those of Almas (2012), we also present estimates in which the source of the 1996 PPP-adjusted GDP data is from PWT 6.1. If we run the regression of lights on log GDP from PWT 6.1 and log GDP at market exchange rates (column 3; the market exchange rate variable is from the April 2005 vintage of the World Development Indicators), we get coefficients of 0.66 on log GDP from PWT 6.1 and 0.37 on log GDP at market exchange rates, with similar magnitudes if we add year fixed effects or covariates (though with covariates, the coefficient on log GDP at market exchange rates becomes statistically insignificant). Hence, it appears that when 1996 PPP is used instead of 2011 PPP, the advantage of PPP adjustment over market exchange rates is considerably less clear-cut. This finding dovetails with the results of Dowrick and Akmal (2008) and Almas (2012), who use 1996 PPP data to find that PPP-adjusted estimate do not necessarily dominate market exchange rates.

4.2 Optimal Weights

Table IB presents the same results as does Table IA, but in the form of optimal relative weights, rather than regression coefficients. Recall that the optimal relative weight of each variable is equal to its coefficient divided by the sum of the coefficients. Since the regression coefficients are normally distributed, their quotients may have nonstandard distributions, so we bootstrap the regressions that we run, and present the 2.5th and 97.5th percentiles of the distributions of the optimal relative weights in lieu of standard errors. We have 200 bootstrap replications. Since each regression has only two right-hand-side variables, the weights and confidence intervals on the second variable (in this case, log GDP at market exchange rates) are just $1-$ the weights and confidence intervals on the first variable, but are presented for clarity.

Upon shifting our focus to optimal weights, it is worth examining several null hypotheses. First, we wish to consider the null that a series should receive zero weight in the best unbiased linear predictor, which obtains whenever the confidence interval includes zero. Second, we may test the null hypothesis that a series should receive unit weight in the best unbiased linear predictor, which is equivalent to the other right hand-side variable that is included in the predictor receiving zero weight.

We see from column 1 that when we wish to form the best unbiased linear estimator from 2011 PPP and MAR, then we reject the null hypothesis that the optimal weight on 2011 PPP is equal to zero regardless of whether year fixed effects or controls are added. We fail to reject the null hypothesis that this weight is equal to unity in the first two panels, while in the third panel, we reject this null hypothesis because the entire confidence interval for this weight is above unity. The point estimates suggest that the weight on 2011 PPP should be approximately unity, while the weight on MAR should be zero. On the other hand, when we look at 1996 PPP in the World Development Indicators or in PWT 6.1 (columns 2 and 3), we obtain that its point estimates are far away from unity (0.52 and 0.63) and this 95% confidence intervals lies strictly in the unit interval unless covariates are added to the regression. If covariates are added, we fail to reject that the optimal weight on 1996 PPP in the WDI is equal to zero, although we do reject this null for 1996 PPP in PWT 6.1.

4.3 Differences in Results for Rich and Poor Countries

Almas (2012) suggests that GDP data adjusted by 1996 PPP may be closer to true economic activity for richer countries, while GDP data adjusted by contemporaneous market exchange rates better reflects economic activity in poorer countries. Following this finding, we present the analog of Table IB separately for countries above and below median GDP per capita in 1992, which we call rich and poor countries, respectively,[11] in Tables IIA and IIB. We see in Table IIA that for rich countries, the optimal weight on log GDP at PPP is above 0.9, regardless of specification and both for 2011 and 1996 PPP. All these estimates are statistically indistinguishable from unity at 95% confidence, and statistically distinguishable from zero with the same confidence, except in the specification that uses data at 1996 PPP from the WDI and contains covariates.[12] Unlike the optimal weight estimates in Table IB, which varied from constancy at around 100% of the weight for log GDP at PPP when 2011 PPP was used to close to 50% of the weight for log GDP at PPP and 50% of the weight for market exchange rates when 1996 PPP was used, the optimal weight estimates in Table IIA are stable at around 100% or more of the weight for log GDP at PPP regardless of the PPP vintage. There must then be considerable heterogeneity across countries in the optimal weights that should be given to log GDP at PPP relative to log GDP at market exchange rates, and this heterogeneity must be driven by the poor countries.

In accordance with this intuition, Table IIB presents the results for poor countries. We observe that when 2011 PPP is used (column 1), the estimated optimal weight on log GDP at PPP is above 0.93, statistically indistinguishable from unity, and statistically distinguishable from zero with 95% confidence or more, regardless of whether covariates are employed. On the other hand, in column 2, where the November 2007 WDI vintage, which is at 1996 PPP, is used, the optimal weight on log GDP at PPP is quantitatively small, statistically indistinguishable from zero, and statistically different from unity with 99% confidence. Correspondingly, the optimal weight on log GDP at market exchange rates is large, statistically different from zero at 99% confidence, and statistically insignificantly different from unity. In column 3, where PWT 6.1 log GDP is used (which also is at 1996 PPP but with some idiosyncratic adjustments), the optimal weight on log GDP at PPP is somewhat larger (at about $0.19 - 0.25$) but still considerably and statistically

[11] We measure GDP per capita in 1992 with the national accounts-based series of PWT 9.0

[12] We also reject the null hypothesis that the estimate is unity for the specification with 2011 PPP and covariates because the lower confidence bound is above unity.

significantly below unity, and statistically indistinguishable from zero.

These results suggest that the arguments of Dowrick and Akmal (2008) and Almas (2012) that PPP-adjusted estimates may be less accurate than GDP at market exchange rates for poor countries were valid when they were formulated (and for the data that was used in these papers) but that improvements of the PPP-adjusted data over time (and, specifically, by the use of 2011 PPP) has solved the problem that these papers highlighted. The poorer performance of 1996 PPP relative to market exchange rates that was shown in Table IB is entirely driven by its performance for poor countries, which, as we see from Table IIB, has been remedied by 2011 PPP.

4.4 Optimal Weights by WDI Vintage

While in Tables IA through IIB we considered WDI data from May 2017 (the first vintage using 2011 PPP) and November 2007 (the last vintage using 1996 PPP), it is interesting to examine the optimal weight on log GDP at PPP relative to log GDP at market exchange rates for all available vintages of the WDI. Figure 2 presents the plot of the optimal relative weight on log GDP at PPP for all available vintages starting with April 2005. All estimates are from the baseline specification without year fixed effects and controls. Vertical lines denote the last vintages using 1996 PPP (released in November 2007) and 2005 PPP (April 2014). We see at once that for all vintages before the introduction of 2005 PPP – not just November 2007 – the optimal relative weight on log GDP at PPP was marginally above 0.5 and statistically significantly different from unity with 95% confidence. However, as soon as 2005 PPP is introduced in April 2008, the optimal relative weight on log GDP at PPP shoots up to unity (or slightly above). While the optimal relative weight declines discretely upon the adoption of 2011 PPP, the decline is not large and the confidence intervals reject the null that it returns to its original level of 0.5, while at the same time failing to reject the null that it remains at unity. Hence, we observe further evidence that the introduction of 2005 PPP improved the estimation of GDP relative to market exchange rates over the price data that existed before.

Figure 3 illustrates that this improvement took place disproportionately for poor countries, in parallel with the results of Table IIB. In the left panel of Figure 3 we observe optimal weights for rich countries. We see that their optimal relative weight of log GDP at PPP is always above unity, and statistically significantly so for all vintages with 2005 PPP. However, in the right panel, where we present optimal relative weights on log GDP at PPP for poor countries, this weight is reasonably precisely estimated to be zero for all vintages using 1996 PPP, with values above 0.5 excluded at 95% confidence. Once 2005 and later pricing data is used to compute PPP-adjusted estimates, the optimal relative weight on log GDP at PPP rises sharply to unity, although the confidence intervals on many estimates tend to be wide, with the lower bound skirting zero. These results are consistent with our conclusion from Table IIB that while earlier PPP vintages had considerable bias problems, later versions appear to have largely remedied them.

4.5 Did Extrapolation Matter?

In this section, we found that PPPs are more descriptive of economic activity than market exchange rates are in 2011 but much less so in 1996. One possible explanation for this result may be that for many countries in the 1996 ICP, prices were not directly obtained through surveys, but rather extrapolated from other countries' survey results via regression. On the other hand, in 2011, virtually all countries participated in price surveys. If extrapolated price estimates were inherently worse than direct survey evidence, then it is not surprising that PPPs better describe economic activity in 2011 than in 1996. Alternatively, the quality of all PPP estimates, including the ones for countries that provided direct survey evidence in both 1996 and 2011, may have improved over these fifteen years. One way to check which of these explanations might be correct is by conducting our regressions from Table IB only on the sample of countries that had price surveys in 1996. If we see that the results are the same as in Table IB, then we can conclude that the replacement of extrapolated price estimates with direct survey data didn't do much to improve PPPs relative to market exchange rates. This is exactly what we do in Table III: we reproduce Table IB for the sample of countries that participated in the 1996 price survey. First, we see that the estimates for data at 2011 PPP on the new sample (column 1 of Table III) are very similar to the estimates for data at 2011 PPP for the full sample (column 1 of Table IB). Using the 2011 PPPs, it is always optimal to give the PPPs 100% or more of the weight relative to market exchange rate based estimates. Second, we see that the point estimates of the optimal weights on log GDP at 1996 PPP relative to log GDP at market exchange rates (columns 2 and 3) are always much larger than the corresponding estimates in Table IB. While in Table IB, we found it optimal to assign roughly 50% of the weight to market exchange rates (with the weight being significantly greater than zero), on the restricted sample in Table III we find that it is optimal to assign close to zero weight to market exchange rates, and these weights are never significantly greater than zero. We can always reject the null hypothesis that the optimal weight on log GDP at 1996 PPP is equal to zero except in column 2, panel 3 (the estimates using World Bank data on GDP at PPP as well as including covariates). While the sample of countries that participated in the 1996 ICP may also have other properties that make its PPPs particularly reflective of unobserved true income, these results are consistent with PPPs providing better estimates of economic size than market exchange rates do as long as the PPPs are actually measured, rather than extrapolated.

5 Results 2: Comparison of Successive Vintages of PPPs

Our second question was whether PPPs have been improving over time. To answer it, we compare the partial correlations of WDI vintages containing different PPPs with nighttime lights, following the theory in Section 3. One important confounder that was not present in Section 4 will be changes to the underlying national accounts data across WDI vintages. Without adequately controlling for these changes, we run the risk of attributing the effects of differences in the national accounts data to differences in PPPs. We will alleviate this concern by showing that differences in the PPPs dominate the different correlations between various WDI vintages and nighttime lights.

5.1 Graphical Results

First, we consider comparisons between successive vintages of the WDI that involve a change in the PPP. Panels 1) and 2) of Figure 4 present partial scatterplots of log nighttime lights against log GDP at 2011 PPP from the May 2014 vintage (2011 PPP for short) and log GDP at 2005 PPP from the April 2014 vintage (2005 PPP for short), respectively. We observe in Panel 1 that there is a strong positive partial correlation between lights and 2011 PPP, which indicates that conditional on knowing 2005 PPP, there is still considerable information to be gained about the nighttime lights from data on log GDP at the more recent PPP. On the other hand, Panel 2 shows that the partial correlation between lights and 2005 PPP, once 2011 PPP is conditioned on, is zero. Hence, a best unbiased linear estimator of log nighttime lights (and, by our assumptions, of log unobserved true income) that is based on the two log GDP series should place most of its weight on the newer 2011 PPP series rather than on the older 2005 PPP series.

Panels 3) and 4) of Figure 4 present the same analysis for log GDP at 2005 PPP (measured in April 2008) and log GDP at 1996 PPP (measured in November 2007). We again refer to them as 2005 PPP and 1996 PPP for short (although 2005 PPP is measured with a different dataset than in the previous paragraph). Once again, we see that there is a strong, positive partial correlation between lights and log GDP at the newer PPP. However, for the older PPP, the partial correlation is not just zero, but unambiguously negative. The implication to draw is that, conditional on 2005 PPP, places with high 1996 PPP actually have lower log nighttime lights (and hence, log unobserved true income) than places with low 1996 PPP. This can be understood intuitively as that places where log GDP was revised upward by the transition from 1996 PPP to 2005 PPP should have their log GDP revised further upward still, and vice versa for places where log GDP was revised downward. In contrast to the case in which both vintages enter with positive weights, which implies that each vintage introduces some idiosyncratic error that would be efficient to average out, the case with a positive and a negative weight suggests that not only is one of the vintages unambiguously better than the other, but that we should revise the data further in the direction of the better vintage, and away from the worse vintage.

5.2 Optimal Weights

Table IV presents optimal relative weights for log GDP at various PPPs. Column 1 presents the optimal weights derived from the regression of log nighttime lights on log GDP at 2005 PPP (measured in April 2008) and log GDP at 1996 PPP (measured in November 2007), which is the regression generating the scatterplots in Panels 3 and 4 of Figure 4. As in Table IB, the first panel presents estimates from just the baseline regression without any additional controls, the second panel adds year fixed effects, and the third panel adds both year fixed effects and the covariates discussed in Section 2.4. We see that 2005 PPP should receive a weight of 1.43, with a confidence interval between 1.2 and 1.6, while 1996 PPP should receive a weight of -0.43, with a confidence interval that includes only negative values. We conclude that the best estimator of unobserved true income that is based on 2005 PPP and 1996 PPP should place unit weight on 2005 PPP, and place a further 0.43 of the weight on the difference between 2005 PPP and 1996 PPP. Hence, while 1996 PPP can be helpful in estimating unobserved true income, it is helpful only in pinpointing which observations should be revised even more strongly in the direction of 2005 PPP. We interpret these estimates as indicating that 2005 PPP is not only a better measure of log unobserved true income than is 1996 PPP, but that further revisions should go further in revising the prediction of unobserved true income in the direction of 2005 PPP and away from 1996 PPP. Adding year fixed effects (column 1, panel 2) does not change our results, while adding controls brings the optimal relative weight on 2005 PPP closer to unity, which still indicates consistently that 2005 PPP is the better measure.

As we have mentioned at the beginning of this section, one might be concerned that the WDI vintages in November 2007 and April 2008 differ not only through the revision of the PPP, but also because of revisions to the underlying national accounts data in local currency units (LCU). If national accounts are improving over time, then such change would naturally favor the later vintage over the earlier vintage. While we will show that, in fact, the national accounts are not consistently improving over time, in column 2 of Table IV we address this concern by also including estimates of log GDP at constant LCU for both the November 2007 and April 2008 vintages. These variables should capture any changes in the underlying national accounts data, if any, and leave the regression coefficients on the PPP-adjusted measures be driven solely by the differences in PPPs. We see that the estimates of the optimal relative weights on 2005 PPP remain virtually unchanged with the inclusion of the LCU controls. For every regression involving the comparison of 1996 PPP and 2005 PPP, we conclusively reject the null hypothesis that the older PPP should receive a weight of unity, and moreover, we reject the null hypothesis that it should receive any positive weight above 0.01. Hence, the data expresses a clear preference for using the newer PPP over the older PPP.

Columns 3 and 4 compute optimal weights for 2011 PPP relative to 1996 PPP in the same manner, and we observe the same results in favor of the newer PPP. All of the regressions considered reject the null hypothesis that the older PPP vintage (1996 PPP) should receive a weight of more than 0.16, and all the point estimates of the optimal relative weight on the older vintage are negative. In contrast, the regressions typically fail to reject the null hypothesis that the newer vintage (2011 PPP) should receive unit weight, or, if they reject it, the confidence interval lies above, rather than below, unity.

Columns 5 and 6 compute optimal weights for 2011 PPP relative to 2005 PPP. Once again, we can always reject the null hypothesis that the weight on 2005 PPP should be unity while the weight on 2011 PPP should be zero. We always fail to reject the null hypothesis that the 2011 PPP vintage should get the whole weight, while the 2005 PPP vintage should get no weight, which is also what the point estimates suggest.

It is interesting to examine whether the apparent improvement of PPPs over time may stem from the growth of the number of countries participating in the price surveys (and thus, from the substitution of data for imputed estimates) or from improvement in either the data quality or the methodology for countries already participating in the ICP. If indeed the improvement of the PPPs comes exclusively from countries that conduct price surveys for the first time, then estimating our regressions from Table IV over countries that are already conducting price surveys in 1996 should yield no particularly higher weight for the newer PPP. Table V presents estimates for the baseline specification from Table IV, run over the sample of countries participating in the 1996 ICP. The results are qualitatively unchanged from Table IV; in particular, we always fail to reject the null hypothesis that the newer vintage should receive a weight of unity, and always reject the null hypothesis that the older vintage should receive a weight of unity. Therefore, we conclude that improvements in data quality and methodology for countries where price data was already collected have been important to the improvement of PPP data over time.

5.3 Changes to the National Accounts Data over Time

In the previous section, we alleviated our earlier concern that changes in GDP estimates across consecutive vintages may be driven by changes in the underlying national accounts data rather than by changes to the PPP by including log GDP at local currency from both vintages as controls. The fact that we find our estimates to remain essentially stable suggests that any changes in the national accounts data are not driving our results. However, it is still interesting to ask whether the national accounts data may be improving over time. To that end, Table VI compares the performance of GDP estimates in the earliest and latest available

WDI vintages that share the same PPP. For example, to look at the evolution of national accounts estimates that rely on 1996 PPP, we include log GDP at 1996 PPP measured in April 2005 (the earliest available WDI vintage) and log GDP at 1996 PPP measured in November 2007 (the latest available WDI vintage that uses 1996 PPP) on the right hand-side of our regression specification (column 1). We observe that, first, the estimates of the optimal weights are very noisy, and the weight on neither WDI vintage is statistically significantly different from zero. Second, if anything, the weight on the older WDI vintage appears to be larger than the weight on the newer WDI vintage. Since the PPP used in both vintages is the same, this suggests that the older national accounts data was more predictive of nighttime lights than the newer national accounts data is. Adding controls or weighting observations by country population does not alter our conclusions. Surprisingly enough, when we compare the performance of the earliest and latest vintage that use 2005 PPP (column 2), and especially the earliest and latest vintage that use 2011 PPP (column 3), we see, if anything, a stronger tendency for the data to suggest evidence in favor of the older rather than the newer vintage of the WDI, and hence, of the national accounts data. For example, in column 2, where we compare log GDP from the April 2014 WDI vintage with log GDP from the April 2008 WDI vintage, regardless of the inclusion of controls we can reject the null hypothesis that the newer vintage should receive all of the weight and that the older vintage should receive zero weight, while failing to reject the converse hypothesis that the newer vintage should receive zero weight and the older vintage should receive all of the weight. We reach the same conclusion for all specifications in column 3, where we compare the latest archived WDI vintage (May 2017) with the earliest WDI vintage that uses 2011 PPP (May 2014). Therefore, if anything, while updates of the PPPs appear to increase the predictive power of GDP data in the WDI for nighttime lights (and hence, for unobserved true income), updates of the underlying national accounts data seem to lower it.

We attempt to visualize the comparison between changes in the underlying national accounts data and changes in the PPP in Figure 5. In this figure, we run regressions of log nighttime lights on log GDP at PPP from each vintage of the WDI and its directly preceding vintage. We then compute a "t-statistic" for the optimal relative weight on the successor vintage for each regression by dividing the weight by one-fourth of the width of its 95% confidence interval, and plot these t-statistics against the vintage date. We see that most of these t-statistics lie in the range between -2 and 2, but the t-statistics corresponding to vintages in which the PPP is changed stand out; the t-statistic for April 2008 (first vintage with 2005 PPP) is 8.5, while the t-statistic for May 2014 (first vintage with 2011 PPP) is 3.5. The largest t-statistic for a vintage without a PPP change is for July 2013, and is equal to 2.77. Therefore, it appears that PPP revisions dominate the landscape of shocks to GDP quality across WDI vintages.

5.4 How does 2005 PPP Perform For Estimating GDP in 1996?

Johnson et al. (2013) have argued for PPP-adjusting GDP data using the closest price survey available to the year for which GDP is estimated. While this argument makes intuitive sense, Section 3.2 outlined an elementary model of a data generating process of price data for which this suggestion would not be optimal; namely, newer data may be so much more precisely estimated relative to the over-time variation in prices that price measures taken today are better measures of prices a long time ago than were the price measures a long time ago. Our first pass at trying to determine whether the theory of Johnson et al. (2013) or the elementary model of Section 3.2 better describes the data is to look at the performance of different PPP vintages in predicting nighttime lights (and hence, unobserved true income) in different years. According to the argument of Johnson et al. (2013), 1996 PPP should be relatively better at predicting log nighttime lights around 1996, while 2005 PPP should be better at predicting log nighttime lights around 2005. On the other hand, according to the argument in Section 3.2, 2005 PPP might be more predictive of log nighttime lights in any year relative to 1996 PPP if it is more predictive of nighttime lights on average.

Panel 1 of Figure 6 presents the results of the regression of lights on 1996 PPP and 2005 PPP separately for each year. We see that in all years – including 1996, the year of the price survey underlying the older PPP – the optimal weight on 2005 PPP is estimated to be greater than unity, with confidence intervals that comfortably exclude zero (or, for that matter, unity itself). This result parallels our finding that log nighttime lights are more strongly correlated with 2005 PPP than they are with 1996 PPP (Table IV, columns 1 and 2), but the fact that this statement holds individually for every year, rather than predominantly in years closer to 2005 than to 1996, suggests that there are no special advantages to using contemporaneously measured prices to measure GDP.

Panel 2 of Figure 6 similarly presents the results of regressing year-by-year log nighttime lights on 2005 PPP and 2011 PPP. Once again, the optimal weight on 2011 PPP is close to unity and statistically significantly different from zero uniformly across years (with the only exception being the year 2008). We conclude that when one PPP seems to be more correlated with nighttime lights than another PPP, this tends to be uniform across years, rather than particularly strong for the year in which the PPP was constructed and weaker in other years.

6 Results 3: Should We Use Contemporaneous Prices or Latest Prices?

We have just seen in Section 5.4 that later PPPs may be better measures of underlying economic activity in a given year than PPPs constructed in that year. In particular, if data at 2005 PPP is better than data at 1996 PPP at measuring economic activity in 1996, it seems intuitive that one might wish to discard the data at 1996 PPP for the purposes of estimating unobserved true income and use only the data at 2005 PPP. In this section, we explicitly test this idea using alternative series from the Penn World Tables (PWT). As of vintage 8.0, the PWT reports two distinct GDP series: one using the latest available PPP (an extrapolated version of the 2005 PPP for PWT 8.0 and the 2011 PPP for PWT 9.0), and the other using an interpolation of all the previous PPPs including the current one. We will refer to the latter series as the multiple benchmark series. Since the evidence of the previous section, especially the fact that later PPP measures are more predictive of nighttime lights than earlier measures in the years for which the earlier measures were computed, is consistent with relative prices changing slowly, it is reasonable to believe that the multiple benchmark series is a good approximaton to a GDP series using price data that vary over time and that are computed using the best methodology available to date.

6.1 Graphical Evidence

We can test whether and when using the latest price data offers a better estimate of nighttime lights, and hence, unobserved true income, than using the best contemporaneous price data available by a simple linear regression of log nighttime lights on log PWT GDP estimated with the multiple benchmark methodology and log PWT GDP estimated by using the latest price survey. Figure 7 presents the partial scatters of log nighttime lights on log PWT 9.0 GDP at 2011 PPP (latest PPP for short) holding log PWT 9.0 multiple benchmark GDP (multiple benchmark series, or contemporaneous series, for short) fixed and vice versa. While the scatters are not as linear as the plots in Figures 1 and 4, there is still an unmistakeable positive partial correlation between lights and latest PPP, and no partial correlation between lights and the multiple benchmark series, suggesting that analysts should put more weight on GDP estimates with latest prices than GDP estimates with contemporaneous prices.

6.2 Optimal Weights

Table VII presents the optimal relative weights on the multiple benchmark series and log GDP at latest PPP. Once again, interesting null hypotheses to test are that the weight on log GDP at latest PPP is zero (which would suggest that log multiple benchmark GDP is a better predictor of log nighttime lights than log GDP at latest PPP) and that the weight on log GDP at latest PPP is unity (which would suggest the reverse conclusion). For each PWT vintage in 8.0 and 9.0 we present results without controls, with year fixed effects, and with year fixed effects and the covariates described in Section 2.4 as usual. We also present results on a sample of countries and years for which either direct price data exists from the PPP surveys (these would be observations for 1996 and 2005 for the countries participating in the price surveys) or which lie in between two price surveys for the same country, allowing prices to be interpolated within country and across time. The multiple benchmark methodology would be expected to produce more accurate estimates for the interpolated prices sample than for the non-interpolated sample. Finally, motivated by the potential for outliers in Figure 7, we also present specifications in which we estimate the optimal relative weights using median regression instead of ordinary least squares. Median regression minimizes least absolute deviations, instead of least squares, which makes it more resistant to outliers than ordinary least squares is.

We observe that for each set of estimates in Table VII we can reject the null hypothesis that the optimal weight on the log GDP at latest PPP series is zero with 99% confidence, and for each set of estimates but one we fail to reject that the optimal weight on the log multiple benchmark GDP series is zero at traditional significance levels (the sole exception, which looks at PWT 8.0 data on the benchmark sample, and includes year fixed effects and controls, involves a rejection in favor of the alternative that the optimal weight on log multiple benchmark GDP is negative). In particular, the first set of estimates (column 1, panel 1), which corresponds to the ratio of the partial effects depicted in Figure 7, suggests that all the weight should be placed on the latest PPP and no weight should be placed on the multiple benchmark series. For the analysis with PWT 9.0, point estimates of the optimal relative weight on log GDP at latest (2011) PPP are all between 0.93 and 1.06, while for regressions with PWT 8.0 the point estimates in regressions without covariates on the main sample can be as low as 0.76, but return to a value of around unity when covariates are included. Hence, on average, it appears that using the latest price data is a much more effective way of explaining the nighttime lights data (and hence, is likely to be a much better measure of unobserved true income) than using all available price data.

Once again, it is relevant to ask whether the difference in predictive power between GDP at the latest PPP and GDP at the multiple benchmark PPP is driven by more countries participating in price surveys for later rounds of the ICP. Table VIII presents estimates of the specifications in Table VII run over the sample of countries that participated in the 1996 ICP. We see that the entry of new countries into the ICP is not driving our results; in each specification we reject the null hypothesis that the weight on log GDP at the multiple benchmark PPP is unity and fail to reject the corresponding hypothesis for log GDP at the latest PPP.

6.3 Optimal Weights over Time

We conclude our analysis by considering the pattern of the optimal weights on the latest PPP and the multiple benchmark series over time. As we recall from Section 3.2, a straightforward model of the tradeoffs between using more precise (hence, later) data and using data that better corresponds to the prices that one wishes to measure (the contemporaneous data) suggests that the analyst should use the latest price data to estimate prices in the distant past, while she should use the contemporaneous price data to estimate prices closer to the period of the latest price data. Hence, in our bivariate regression of log nighttime lights on log GDP at latest prices (2011 PPP) and log GDP at contemporaneous prices (multiple benchmark series), the optimal weight on the log GDP at latest prices should be high for early years in the sample, but then should decline as one approaches 2011. Figure 8 presents estimates of this optimal relative weight on the latest PPP in PWT 9.0 and PWT 8.0 for each year in our sample, separately, and we see that the implication of our model in Section 3.2 is confirmed. In particular, the optimal relative weight on log PWT 9.0 GDP at 2011 PPP is precisely estimated to be unity until around 2005, after which year it is no longer significantly different from zero at 5% significance, and by 2009 the point estimate of this weight is negative. The intuition for this pattern (and the result of the model) is that contemporaneous prices a long time ago are measured very imprecisely relative to the latest price data, and since prices change slowly, the latest price data is preferable. However, contemporaneous prices a few years before the latest price data may be measured reasonably precisely, and it may be worth putting some weight on them as they should reflect contemporaneous prices differences quite well.

7 Conclusion

Our results are informative to researchers wishing to use PPPs in validating and affirming current practices in the literature, such as using PPPs directly rather than combining them with market exchange rate data, and revising PPPs in faraway prior years after a new round of price surveys is completed. There may be reason to revisit the literatures of the 1990s and 2000s that used 1996 PPP, and follow the recommendations of Dowrick and Akmal (2008) and Almas (2012) to address the PPP bias, possibly by computing weighted averages of PPPs and market exchange rates. However, papers that use more recent PPP data should largely be free of these concerns, and market exchange rate data currently has little to contribute to estimating GDP once PPPs are incorporated. This improvement in price data has occurred, obviously, because direct price data now replaces extrapolation for many countries, but also because survey methodology and data analysis have improved for countries that had direct price data to begin with. We have also learned, reassuringly, that

PPPs have improved considerably over time (in contrast to the underlying national accounts data, which, if anything, may have deteriorated). We also learn, perhaps counterintuitively, that we cannot do much better than revising previous PPP estimates each time we get a new price survey, because the price surveys appear to be improving rapidly, and there is almost no benefit to be obtained from older price data. While using multiple rounds of price data may appear appealing in principle, it will become efficient in practice only for the years when the measurement error of price surveys becomes sufficiently small compared to the rate of change of relative prices over time.

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] **Almås, Ingvild.** "International Income Inequality: Measuring PPP bias by estimating Engel curves for food." American Economic Review 102, no. 2 (2012): 1093-1117.

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

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

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

[6] **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.

[7] **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.

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

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

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

[11] **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

[12] **Dowrick, Steve, and Muhammad Akmal.** "Contradictory trends in global income inequality: A tale of two biases." Review of Income and Wealth 51, no. 2 (2005): 201-229.

[13] The Economist. "Step Change: Revised Figures Show Nigeria Africa's Largest Economy," https://www.economist.com/news/finance-and-economics/21600734-revised-figures-show-nigeria-africas-largest-economy-step-change. Date Created: 04/12/2-14. Date Accessed: 02/21/2018.

[14] **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

[15] **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.

[16] **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.

[17] **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

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

[19] **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.

[20] **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.

[21] **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.

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

[23] **Kravis, Irving B., Alan W. Heston, and Robert Summers**. "Real GDP per capita for more than one hundred countries." The Economic Journal 88, no. 350 (1978): 215-242.

[24] **Kravis, Irving B., Alan Heston, and Robert Summers.** World product and income: International comparisons of real gross product. The World Bank, 1982.

[25] **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.

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

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

[28] **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).

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

[30] **Pinkovskiy, Maxim L.** and **Xavier Sala-i-Martin**. "Lights, Camera, ... Income! Illuminating the National Accounts-Household Surveys Debate," *Quarterly Journal of Economics*, 131, vol. 2: 579-631.

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

[32] **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

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

Tables

Table IA (IA)

Comparing PPP and Market Exchange Rates			
Lights Regression Coefficients			
Dep. Var. is Log Light Intensity			
	(1)	(2)	(3)
PPP Vintage	2011 PPP	1996 PPP	1996 PPP
PPP Data Source	WDI May 2014	WDI Nov 2007	PWT 6.1
Mrkt Exchange Rates Data Source	WDI May 2014	WDI Nov 2007	WDI Apr 2005
No Fixed Effects			
Log GDP at PPP	.96***	.53***	.66***
	(.11)	(.12)	(.12)
Log GDP at Market Exchange Rates	.06	.47***	.37***
	(.11)	(.12)	(.11)
Number of Obs.	2634	2079	1113
Number of Clusters	139	140	138
Year Fixed Effects			
Log GDP at PPP	.97***	.54***	.65***
	(.11)	(.12)	(.12)
Log GDP at Market Exchange Rates	.06	.47***	.37***
	(.11)	(.12)	(.11)
Number of Obs.	2634	2079	1113
Number of Clusters	139	140	138
Year Fixed Effects + Controls			
Log GDP at PPP	1.24***	.16	.61***
	(.15)	(.19)	(.20)
Log GDP at Market Exchange Rates	-.32***	.53***	.23
	(.11)	(.15)	(.14)
Number of Obs.	2634	2079	1113
Number of Clusters	139	140	138

Table IA presents estimates for the regressions of log nighttime lights on log GDP at PPP and log GDP at market exchange rates 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 at PPP and at market exchange rates are from the World Development Indicators. The covariates are described in Section 2.4. The sample comprises of all countries and years with full information on covariates.

Table IB (IB)

Comparing PPP and Market Exchange Rates Lights Regression-Based Optimal Weights *Dependent Variable is Log Light Intensity*			
	(1)	(2)	(3)
PPP Vintage	2011 PPP	1996 PPP	1996 PPP
PPP Data Source	WDI May 2014	WDI Nov 2007	PWT 6.1
Mrkt Exchange Rates Data Source	WDI May 2014	WDI Nov 2007	WDI Apr 2005
No Fixed Effects			
Log GDP at PPP	.93***	.52***	.63***
	(.67,1.1)	(.26,.78)	(.39,.86)
Log GDP at Market Exchange Rates	.06	.47***	.36***
	(-.12,.32)	(.21,.73)	(.13,.60)
Number of Obs.	2634	2079	1113
Number of Clusters	139	140	138
Year Fixed Effects			
Log GDP at PPP	.93***	.53***	.63***
	(.67,1.1)	(.27,.78)	(.39,.85)
Log GDP at Market Exchange Rates	.06	.46***	.36***
	(-.12,.32)	(.21,.72)	(.14,.60)
Number of Obs.	2634	2079	1113
Number of Clusters	139	140	138
Year Fixed Effects + Controls			
Log GDP at PPP	1.35***	.23	.72***
	(1.08,1.6)	(-.54,.82)	(.32,1.1)
Log GDP at Market Exchange Rates	-.35**	.76***	.27
	(-.67,-.08)	(.17,1.5)	(-.12,.67)
Number of Obs.	2634	2079	1113
Number of Clusters	139	140	138

Table IB presents estimates of the optimal weights from the regressions of log nighttime lights on log GDP at PPP and log GDP at market exchange rates as described in Section 4. Country block-bootstrapped 95% confidence intervals for the weights are in parentheses. Data on nighttime lights from the NOAA, data on measures of log GDP at PPP and at market exchange rates are from the World Development Indicators. The covariates are described in Section 2.4. The sample comprises of all countries and years with full information on covariates.

Table IIA (IIA)

Comparing PPP and Market Exchange Rates, > Median Income Countries **Lights Regression-Based Optimal Weights** *Dependent Variable is Log Light Intensity*			
	(1)	(2)	(3)
PPP Vintage	2011 PPP	1996 PPP	1996 PPP
PPP Data Source	WDI May 2014	WDI Nov 2007	PWT 6.1
Mrkt Exchange Rates Data Source	WDI May 2014	WDI Nov 2007	WDI Apr 2005
No Fixed Effects			
Log GDP at PPP	1.23***	1.31***	1.21***
	(.97,1.4)	(1.03,1.5)	(.99,1.4)
Log GDP at Market Exchange Rates	-.23*	-.31**	-.21*
	(-.44,.02)	(-.58,-.03)	(-.46,.00)
Number of Obs.	1212	959	524
Number of Clusters	64	65	63
Year Fixed Effects			
Log GDP at PPP	1.23***	1.31***	1.21***
	(.97,1.4)	(1.03,1.5)	(.99,1.4)
Log GDP at Market Exchange Rates	-.23*	-.31**	-.21*
	(-.45,.02)	(-.57,-.03)	(-.45,.00)
Number of Obs.	1212	959	524
Number of Clusters	64	65	63
Year Fixed Effects + Controls			
Log GDP at PPP	1.00**	.49	1.09*
	(.28,1.4)	(-1.8,1.6)	(-1.3,1.6)
Log GDP at Market Exchange Rates	-.00	.50	-.09
	(-.46,.71)	(-.60,2.8)	(-.61,2.3)
Number of Obs.	1212	959	524
Number of Clusters	64	65	63

Table IIA presents estimates of the optimal weights from the regressions of log nighttime lights on log GDP at PPP and log GDP at market exchange rates as described in Section 4. Country block-bootstrapped 95% confidence intervals for the weights are in parentheses. Data on nighttime lights from the NOAA, data on measures of log GDP at PPP and at market exchange rates are from the World Development Indicators. The covariates are described in Section 2.4. The sample comprises of all countries and years with full information on covariates, and with the country's PWT 9.0 GDP per capita at 2011 PPP above the median of its distribution.

Table IIB (IIB)

Comparing PPP and Market Exchange Rates, $<$ Median Income Countries **Lights Regression-Based Optimal Weights** *Dependent Variable is Log Light Intensity*			
	(1)	(2)	(3)
PPP Vintage	2011 PPP	1996 PPP	1996 PPP
PPP Data Source	WDI May 2014	WDI Nov 2007	PWT 6.1
Mrkt Exchange Rates Data Source	WDI May 2014	WDI Nov 2007	WDI Apr 2005
No Fixed Effects			
Log GDP at PPP	.95***	.01	.25
	(.30,1.7)	(-.43,.40)	(-.28,.63)
Log GDP at Market Exchange Rates	.04	.98***	.74***
	(-.77,.69)	(.59,1.4)	(.36,1.2)
Number of Obs.	1422	1120	589
Number of Clusters	75	75	75
Year Fixed Effects			
Log GDP at PPP	.93***	.01	.24
	(.29,1.7)	(-.42,.39)	(-.28,.60)
Log GDP at Market Exchange Rates	.06	.98***	.75***
	(-.73,.70)	(.60,1.4)	(.39,1.2)
Number of Obs.	1422	1120	589
Number of Clusters	75	75	75
Year Fixed Effects + Controls			
Log GDP at PPP	1.23***	.00	.19
	(.64,1.9)	(-.69,.42)	(-.36,.66)
Log GDP at Market Exchange Rates	-.23	.99***	.80***
	(-.99,.35)	(.57,1.6)	(.33,1.3)
Number of Obs.	1422	1120	589
Number of Clusters	75	75	75

Table IIB presents estimates of the optimal weights from the regressions of log nighttime lights on log GDP at PPP and log GDP at market exchange rates as described in Section 4. Country block-bootstrapped 95% confidence intervals for the weights are in parentheses. Data on nighttime lights from the NOAA, data on measures of log GDP at PPP and at market exchange rates are from the World Development Indicators. The covariates are described in Section 2.4. The sample comprises of all countries and years with full information on covariates, and with the country's PWT 9.0 GDP per capita at 2011 PPP below the median of its distribution.

Table III (III)

Comparing PPP and Market Exchange Rates: Countries in 1996 ICP Only Lights Regression-Based Optimal Weights *Dependent Variable is Log Light Intensity*			
	(1)	(2)	(3)
PPP Vintage	2011 PPP	1996 PPP	1996 PPP
PPP Data Source	WDI May 2014	WDI Nov 2007	PWT 6.1
Mrkt Exchange Rates Data Source	WDI May 2014	WDI Nov 2007	WDI Apr 2005
No Fixed Effects			
Log GDP at PPP	1.03***	.90***	.81***
	(.80,1.2)	(.55,1.2)	(.53,1.04)
Log GDP at Market Exchange Rates	-.03	.09	.18
	(-.25,.19)	(-.21,.44)	(-.04,.46)
Number of Obs.	1782	1439	801
Number of Clusters	94	97	97
Year Fixed Effects			
Log GDP at PPP	1.03***	.90***	.80***
	(.80,1.2)	(.55,1.2)	(.52,1.03)
Log GDP at Market Exchange Rates	-.03	.09	.19
	(-.26,.19)	(-.21,.44)	(-.03,.47)
Number of Obs.	1782	1439	801
Number of Clusters	94	97	97
Year Fixed Effects + Controls			
Log GDP at PPP	1.41***	.61	.84**
	(1.1,1.7)	(-.18,1.2)	(.38,1.2)
Log GDP at Market Exchange Rates	-.41***	.38	.15
	(-.70,-.17)	(-.25,1.1)	(-.22,.61)
Number of Obs.	1782	1439	801
Number of Clusters	94	97	97

Table III presents estimates of the optimal weights from the regressions of log nighttime lights on log GDP at PPP and log GDP at market exchange rates as described in Section 4. Country block-bootstrapped 95% confidence intervals for the weights are in parentheses. Data on nighttime lights from the NOAA, data on measures of log GDP at PPP and at market exchange rates are from the World Development Indicators. The covariates are described in Section 2.4. The sample comprises of all countries and years with full information on covariates, and with direct (not model-based) price data from the 1996 ICP.

Table IV (IV)

Comparing Successive PPPs: Lights Regression-Based Optimal Weights						
Dependent Variable is Log Light Intensity						
	(1)	(2)	(3)	(4)	(5)	(6)
New PPP	2005 PPP		2011 PPP		2011 PPP	
Old PPP	1996 PPP		1996 PPP		2005 PPP	
New PPP Data	WDI Apr 2008		WDI May 2014		WDI May 2014	
Old PPP Data	WDI Nov 2007		WDI Nov 2007		WDI Apr 2014	
Specification	Baseline	LCU Control	Baseline	LCU Control	Baseline	LCU Control
No Fixed Effects						
New PPP	1.43***	1.45***	1.19***	1.11***	.98***	.88***
	(1.1,1.7)	(1.1,1.7)	(.94,1.4)	(.93,1.3)	(.35,1.4)	(.29,1.2)
Old PPP	-.43***	-.45***	-.19*	-.11	.01	.11
	(-.79,-.12)	(-.76,-.14)	(-.48,.05)	(-.35,.06)	(-.46,.64)	(-.23,.70)
Number of Obs.	2043	2043	2026	2026	2615	2615
Number of Clusters	137	137	137	137	138	138
Year Fixed Effects						
New PPP	1.42***	1.44***	1.19***	1.10***	.97***	.88***
	(1.1,1.7)	(1.1,1.7)	(.93,1.4)	(.92,1.3)	(.34,1.4)	(.29,1.2)
Old PPP	-.42***	-.44**	-.19*	-.10	.02	.11
	(-.78,-.11)	(-.75,-.13)	(-.48,.06)	(-.35,.07)	(-.46,.65)	(-.22,.70)
Number of Obs.	2043	2043	2026	2026	2615	2615
Number of Clusters	137	137	137	137	138	138
Year Fixed Effects + Controls						
New PPP	1.18***	1.27***	1.10***	1.14***	1.33***	1.11***
	(.81,1.5)	(.92,1.7)	(.83,1.5)	(.89,1.5)	(.72,1.8)	(.58,1.6)
Old PPP	-.18	-.27	-.10	-.14	-.33	-.11
	(-.57,.18)	(-.71,.07)	(-.51,.16)	(-.58,.10)	(-.88,.27)	(-.67,.41)
Number of Obs.	2043	2043	2026	2026	2615	2615
Number of Clusters	137	137	137	137	138	138

Table IV presents estimates of the optimal weights from the regressions of log nighttime lights on log GDP at PPP from different vintages of the WDI as described in Section 5. Columns 2, 4, and 6 include log GDP at local currency from the different vintages of the WDI as controls in all specifications. Country block-bootstrapped 95% confidence intervals for the weights are in parentheses. Data on nighttime lights from the NOAA, data on measures of log GDP at PPP and at local currency are from the World Development Indicators. The covariates are described in Section 2.4. The sample comprises of all countries and years with full information on covariates.

Table V (V)

Comparing Successive PPPs: Countries in 1996 ICP Only Lights Regression-Based Optimal Weights *Dependent Variable is Log Light Intensity*			
	(1)	(2)	(3)
New PPP	2005 PPP	2011 PPP	2011 PPP
Old PPP	1996 PPP	1996 PPP	2005 PPP
New PPP Data	WDI Apr 2008	WDI May 2014	WDI May 2014
Old PPP Data	WDI Nov 2007	WDI Nov 2007	WDI Apr 2014
No Fixed Effects			
New PPP	1.52***	1.18***	.84**
	(.98,2.0)	(.83,1.5)	(.20,1.4)
Old PPP	-.52*	-.18	.15
	(-1.08,.01)	(-.58,.16)	(-.47,.79)
Number of Obs.	1403	1401	1763
Number of Clusters	94	95	93
Year Fixed Effects			
New PPP	1.52***	1.18***	.84**
	(.99,2.0)	(.83,1.5)	(.20,1.4)
Old PPP	-.52*	-.18	.15
	(-1.08,.00)	(-.57,.16)	(-.47,.79)
Number of Obs.	1403	1401	1763
Number of Clusters	94	95	93
Year Fixed Effects + Controls			
New PPP	1.42***	1.13***	.87**
	(.92,2.0)	(.64,1.6)	(.08,1.5)
Old PPP	-.42*	-.13	.12
	(-1.08,.07)	(-.62,.35)	(-.56,.91)
Number of Obs.	1403	1401	1763
Number of Clusters	94	95	93

Table IV presents estimates of the optimal weights from the regressions of log nighttime lights on log GDP at PPP from different vintages of the WDI as described in Section 5. Columns 2, 4, and 6 include log GDP at local currency from the different vintages of the WDI as controls in all specifications. Country block-bootstrapped 95% confidence intervals for the weights are in parentheses. Data on nighttime lights from the NOAA, data on measures of log GDP at PPP and at local currency are from the World Development Indicators. The covariates are described in Section 2.4. The sample comprises of all countries and years with full information on covariates and with available price data in the 1996 ICP.

Table VI (VI)

Comparing Underlying National Accounts in WDI Lights Regression-Based Optimal Weights *Dependent Variable is Log Light Intensity*			
	(1)	(2)	(3)
PPP	1996 PPP	2005 PPP	2011 PPP
New NA	WDI Nov 2007	WDI Apr 2014	WDI May 2017
Old NA	WDI Apr 2005	WDI Apr 2008	WDI May 2014
No Fixed Effects			
New NA	.17	-.34	-1.00
	(-1.6,1.4)	(-1.2,.54)	(-2.1,.28)
Old NA	.82	1.34***	2.00***
	(-.48,2.6)	(.45,2.2)	(.71,3.1)
Number of Obs.	1631	2076	2597
Number of Clusters	138	140	138
Year Fixed Effects			
New NA	.18	-.32	-.99
	(-1.6,1.4)	(-1.1,.54)	(-2.0,.29)
Old NA	.81	1.32***	1.99***
	(-.48,2.6)	(.45,2.1)	(.70,3.0)
Number of Obs.	1631	2076	2597
Number of Clusters	138	140	138
Year Fixed Effects + Controls			
New NA	-1.22	.03	-.53
	(-4.0,.55)	(-.90,.83)	(-1.8,.99)
Old NA	2.22***	.96**	1.53**
	(.44,5.0)	(.16,1.9)	(.00,2.8)
Number of Obs.	1631	2076	2597
Number of Clusters	138	140	138

Table VI presents estimates of the optimal weights from the regressions of log nighttime lights on log GDP at PPP from different vintages of the WDI as described in Section 5. Country block-bootstrapped 95% confidence intervals for the weights are in parentheses. Data on nighttime lights from the NOAA, data on measures of log GDP at PPP are from the World Development Indicators. The covariates are described in Section 2.4. The sample comprises of all countries and years with full information on covariates.

Table VII (VII)

Multiple Benchmark vs. Most Recent PPP Series in PWT Lights Regression-Based Optimal Weights *Dependent Variable is Log Light Intensity*						
	(1)	(2)	(3)	(4)	(5)	(6)
PWT Vintage	PWT 9.0			PWT 8.0		
Specification	Baseline	Inter Polated	Median Reg	Baseline	Inter Polated	Median Reg
No Fixed Effects						
Multiple Benchmark GDP	.00	.08	-.08	.23	-.11	-.02
	(-.27,.34)	(-.28,.46)	(-.37,.23)	(-.32,.84)	(-.58,.39)	(-.58,.59)
GDP at Most Recent PPP	.99***	.91***	1.08***	.76**	1.11***	1.02**
	(.65,1.2)	(.53,1.2)	(.76,1.3)	(.15,1.3)	(.60,1.5)	(.40,1.5)
Number of Obs.	2717	1943	2717	2660	1940	2660
Number of Clusters	143	143	143	140	140	140
Year Fixed Effects						
Multiple Benchmark GDP	.03	.13	-.05	.23	-.13	-.01
	(-.23,.37)	(-.22,.49)	(-.33,.24)	(-.32,.83)	(-.56,.31)	(-.57,.55)
GDP at Most Recent PPP	.96***	.86***	1.05***	.76**	1.13***	1.01**
	(.62,1.2)	(.50,1.2)	(.75,1.3)	(.16,1.3)	(.68,1.5)	(.44,1.5)
Number of Obs.	2717	1943	2717	2660	1940	2660
Number of Clusters	143	143	143	140	140	140
Year Fixed Effects + Controls						
Multiple Benchmark GDP	-.06	.03	-.15	-.03	-.20	-.20
	(-.39,.22)	(-.35,.37)	(-.43,.20)	(-.42,.70)	(-.51,.22)	(-.55,.49)
GDP at Most Recent PPP	1.06***	.96***	1.15***	1.03***	1.20***	1.20**
	(.77,1.3)	(.62,1.3)	(.79,1.4)	(.29,1.4)	(.77,1.5)	(.50,1.5)
Number of Obs.	2717	1943	2717	2660	1940	2660
Number of Clusters	143	143	143	140	140	140

Table VII presents estimates of the optimal weights from the regressions of log nighttime lights on log GDP at latest PPP and log GDP at the multiple benchmark PPP as described in Section 6. Country block-bootstrapped 95% confidence intervals for the weights are in parentheses. Data on nighttime lights from the NOAA, data on measures of log GDP are from the Penn World Tables. The covariates are described in Section 2.4. The sample comprises of all countries and years with full information on covariates. The interpolated specifications (columns 2 and 5) include only years that lie between two price surveys for a country. The median regression specifications (columns 3 and 6) use median regression in lieu of OLS to reduce the influence of outliers.

Table VIII (VIII)

Multiple Benchmark vs. Most Recent PPP Series in PWT Countries in 1996 ICP Only Lights Regression-Based Optimal Weights *Dependent Variable is Log Light Intensity*		
	(1)	(2)
PWT Vintage	PWT 9.0	PWT 8.0
No Fixed Effects		
Multiple Benchmark GDP	.06	-.34
	(-.30,.46)	(-.69,.07)
GDP at Most Recent PPP	.93***	1.34***
	(.53,1.3)	(.92,1.6)
Number of Obs.	1843	1843
Number of Clusters	97	97
Year Fixed Effects		
Multiple Benchmark GDP	.08	-.36*
	(-.28,.45)	(-.73,.06)
GDP at Most Recent PPP	.91***	1.36***
	(.54,1.2)	(.93,1.7)
Number of Obs.	1843	1843
Number of Clusters	97	97
Year Fixed Effects + Controls		
Multiple Benchmark GDP	.04	-.22
	(-.26,.44)	(-.61,.24)
GDP at Most Recent PPP	.95***	1.22***
	(.55,1.2)	(.75,1.6)
Number of Obs.	1843	1843
Number of Clusters	97	97

Table VII presents estimates of the optimal weights from the regressions of log nighttime lights on log GDP at latest PPP and log GDP at the multiple benchmark PPP as described in Section 6. Country block-bootstrapped 95% confidence intervals for the weights are in parentheses. Data on nighttime lights from the NOAA, data on measures of log GDP are from the Penn World Tables. The covariates are described in Section 2.4. The sample comprises of all countries and years with full information on covariates and with available price data in 1996 ICP.

Figures

Figure 1 (1)

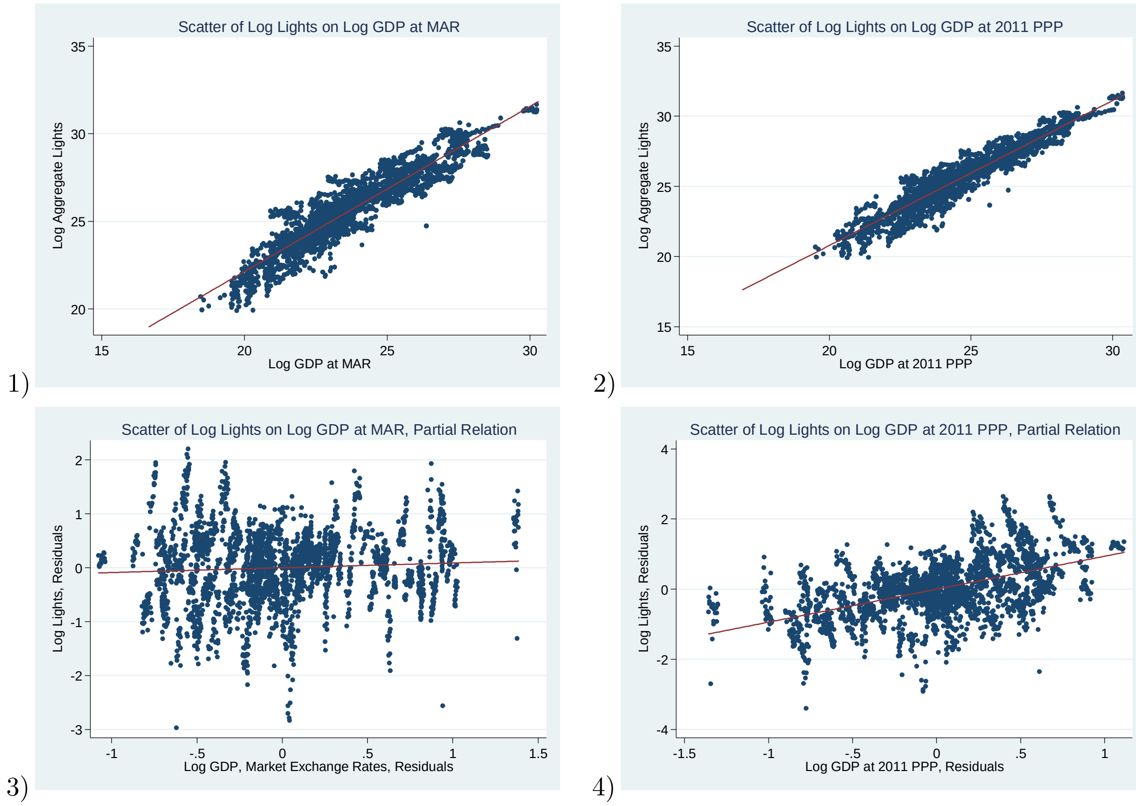


Figure 2 (2)

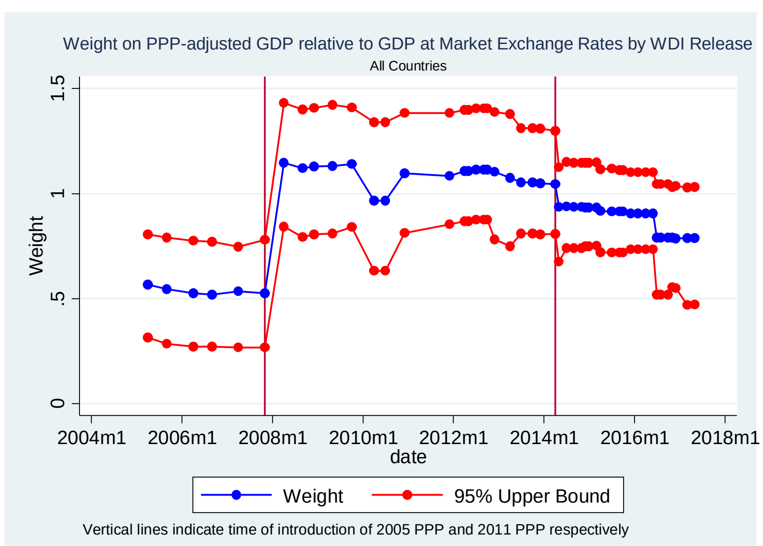
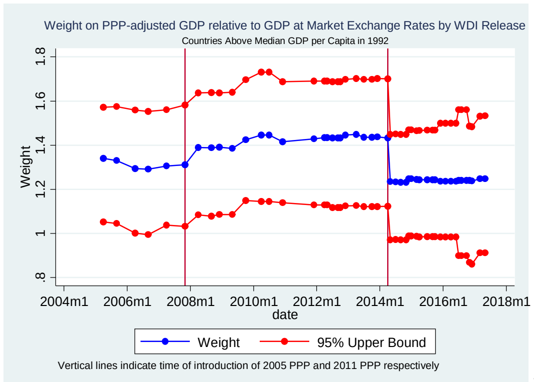
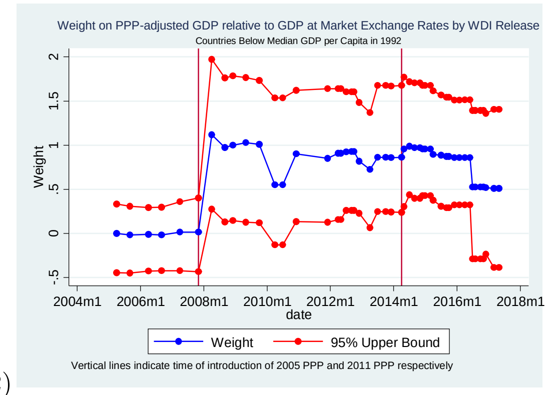


Figure 3 (3)

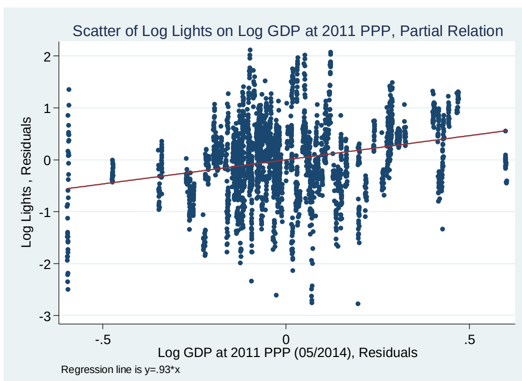


1)

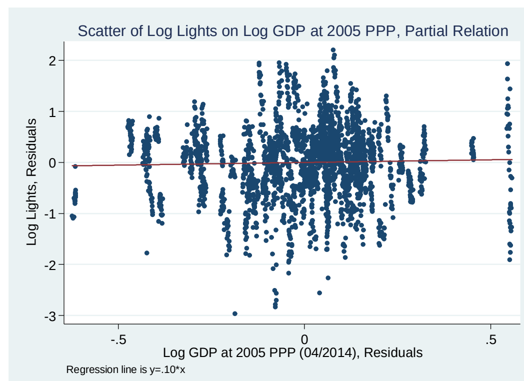


2)

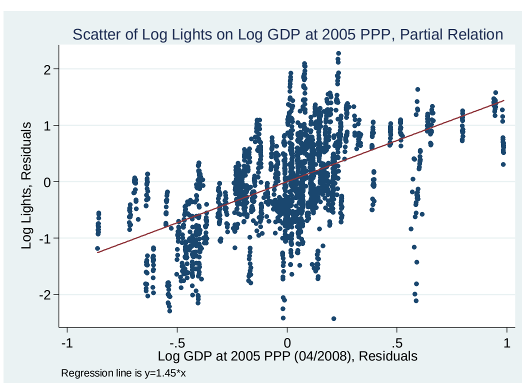
Figure 4 (4)



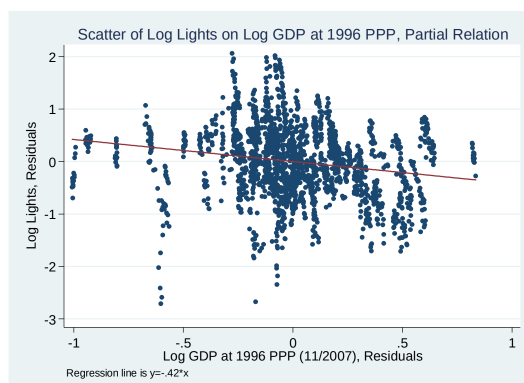
1)



2)



3)



4)

Figure 5 (5)

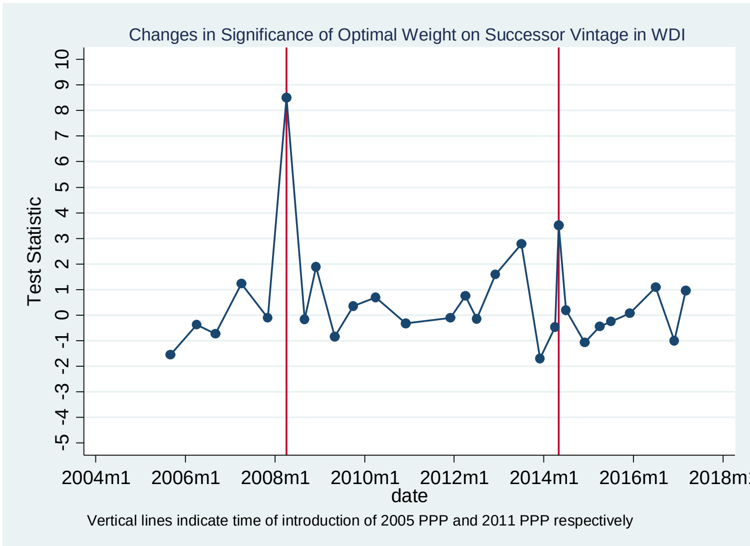
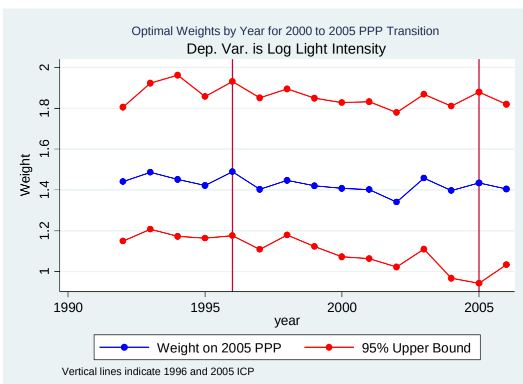


Figure 6 (6)

1) 

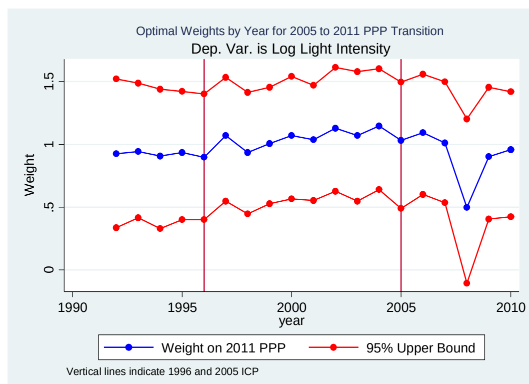
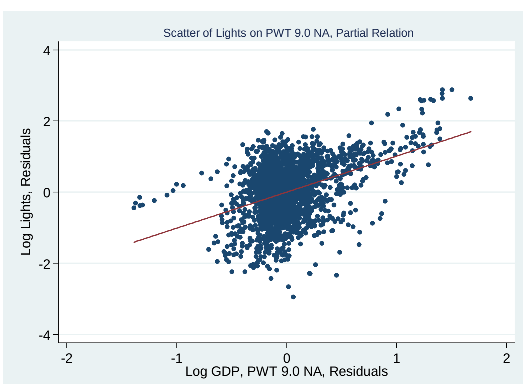
2) 

Figure 7 (7)



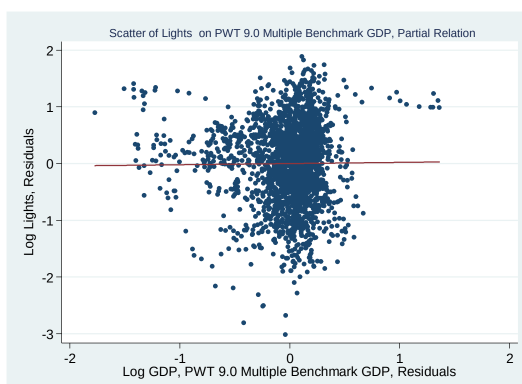
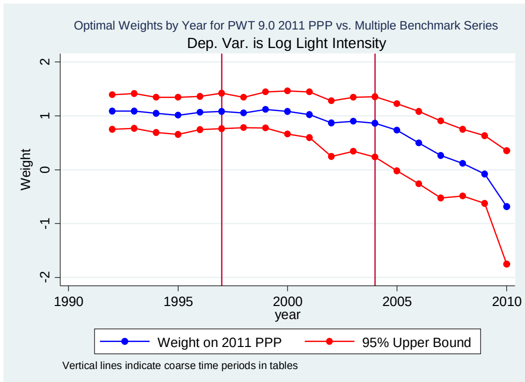
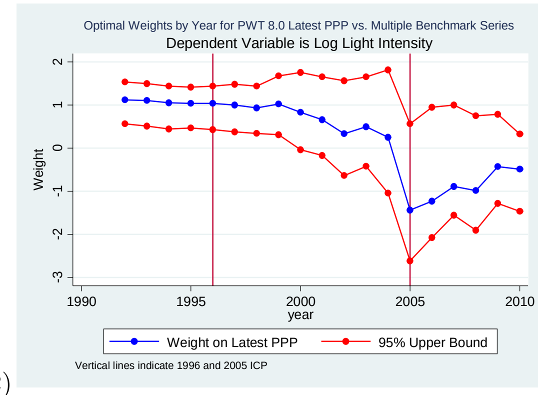


Figure 8 (8)

1) 
Optimal Weights by Year for PWT 9.0 2011 PPP vs. Multiple Benchmark Series
Dep. Var. is Log Light Intensity
Weight
year
Weight on 2011 PPP
95% Upper Bound
Vertical lines indicate coarse time periods in tables

2) 
Optimal Weights by Year for PWT 8.0 Latest PPP vs. Multiple Benchmark Series
Dependent Variable is Log Light Intensity
Weight
year
Weight on Latest PPP
95% Upper Bound
Vertical lines indicate 1996 and 2005 ICP