Empir Econ (2019) 57:971–990
https://doi.org/10.1007/s00181-018-1464-1

On measuring economic growth from outer space: a single country approach

Víctor M. Guerrero[1] · **Juan A. Mendoza**[2]

Received: 14 June 2016 / Accepted: 26 April 2018 / Published online: 1 June 2018

Abstract This article proposes a simple statistical approach to combine nighttime light data with official national income growth figures. The suggested procedure arises from a signal-plus-noise model for official growth along with a constant elasticity relation between observed night lights and income. The methodology implemented in this paper differs from the approach based on panel data for several countries at once that uses World Bank ratings of income data quality for the countries under study to produce an estimate of true economic growth. The new approach: (a) leads to a relatively simple and robust statistical method based only on time series data pertaining to the country under study and (b) does not require the use of quality ratings of official income statistics. For illustrative purposes, some empirical applications are made for Mexico, China and Chile. The results show that during the period of study there was underestimation of economic growth for both Mexico and Chile, while official figures of China over-estimated true economic growth.

The authors thank two anonymous referees and the editor of this journal for their valuable comments and suggestions. Víctor M. Guerrero acknowledges financial support from Asociación Mexicana de Cultura, A. C. to carry out this work.

✉ Juan A. Mendoza
juanmendozap@uchicago.edu

Víctor M. Guerrero
guerrero@itam.mx

[1] Department of Statistics, Instituto Tecnológico Autónomo de México (ITAM), Río Hondo No. 1, Col. Progreso Tizapán, Mexico City 01080, Mexico

[2] Financial Mathematics, University of Chicago, 5727 South University Avenue, Chicago, IL 60637, USA

Keywords Digital number · Generalized least squares · Median · Night lights · Tchebysheff's theorem

JEL Classification C2 · C4 · O4

1 Introduction

Accurate measurement of gross domestic product (GDP) is one of the biggest challenges faced by National Statistical Agencies (NSA) around the world, albeit this task is even more important in developing countries where unobserved economic activity can be substantial and a reliable calculation could be thwarted. In particular, investment decisions, either domestic or external, could be hampered by the lack of credibility in government's information. Hence, NSAs must follow the guidelines put forward by the United Nations Statistical Commission, where the European Commission, the Organization for Economic Co-Operation and Development, the International Monetary Fund and the World Bank, have played important roles to develop the statistical framework for a System of National Accounts (SNA), see United Nations, European Commission, International Monetary Fund, OCDE and WB (2009), and World Bank (2002).

Despite the updating versions of the SNA which emphasize the new features of the economies, such as informality, financial sector improvements and issues related to globalization, uncertainty related to the official GDP growth figures reported by several countries prevails. Nevertheless, innovative approaches for alternative GDP growth calculations have been proposed, such as Aruoba et al. (2013), who focused on a forecast combination methodology using expenditure and income measures of GDP. Similarly, Henderson et al. (2012) (HSW hereafter) suggested an alternative and innovative approach to measure economic growth, using night lights time series data provided by satellite images. In fact, those authors developed a procedure to combine that kind of data with official growth figures.

It is important to remark that HSW focused on the analysis of data from middle income countries, in contrast with those of wealthier countries where official data have full credibility. Moreover, the contribution by Chen and Nordhaus (2011) indicates that night lights satellite data provide useful information for low income countries, which usually have poor statistical systems. Later on, Nordhaus and Chen (2014) found that nighttime lights data are mainly useful for developing countries and carried out a thorough statistical study regarding the contribution of lights data to improve on official output figures.

Some other uses of nightlight imagery appear in studies at regional or subnational level, many of them are about conflicts in Africa, e.g., Harari and La Ferrara (2013) used night lights data to analyze civil conflicts. (They used weather observations in cells of 1° of latitude by 1° of longitude to relate crops and conflicts.) Similarly, Hodler and Raschky (2014a) employed nighttime light intensity as a measure of economic activity at the subnational level to study the effect of economic shocks on the probability of civil conflicts. Also, Michalopoulos and Papaioannou (2014) employed satellite images of light density at night to quantify the impact of national institutions on regional development.

Another use of nighttime lights appears in Hodler and Raschky (2014b) who found that regions around the world have more intense luminosity when being the birth region of the current political leader (thus providing evidence for widespread regional favoritism). Bertinelli and Strobl (2013) made use of nightlight satellite imagery as a measure of local economic activity to assess the impact of hurricane strikes on local economic growth in the Caribbean region. In the same fashion, Elliot et al. (2015) used satellite imagery data and combined it with historical and simulated typhoon tracks to estimate the impact of typhoons on local economic activity along the coast of China. Some other uses of nighttime image data have appeared in the geoscience field (e.g., Doll et al. 2000; Seo 2011).

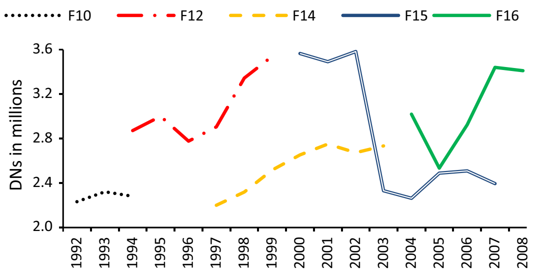
We base our work upon HSW's article, which used panel data for 188 countries and relied on World Bank (WB) ratings of income data quality for the countries under study. Our proposal differs from HSW's in that we make use only of time series and satellite data for each country individually and propose a statistically robust method for estimating economic growth. Thus, our proposal is in line with the official calculation of GDP that relies only on domestic information, not on information from other countries. As a referee of this work pointed out, this is contrary to the usual HSW's approach, since most analysts have used nighttime lights data to get a proxy for GDP at subnational scale, where official estimates are often unavailable. Moreover, the same imagery data can also be used for other economic purposes, e.g., to get a proxy for electrification rates.

Since the 1970s, the United States Air Force Defense Meteorological Satellite Program (DMSP)—an official organization whose main purpose is the design, construction and monitoring of oceanographic and meteorological environments—developed an Operational Linescan System that uses observable and infrared sensors in the satellites to calculate nightlights luminescence. At first, implementation of that system relied on monitoring the earth's cloud distribution, but after several studies scientists at DMSP discovered that luminescence activities such as human lights settlements, northern and southern lights, natural fires and gas flaring could be recorded by using this technique.

Initially, people working for the DMSP faced a conundrum when trying to record only city lights, because this kind of luminescence was hampered by other types of light, such as fires, gas flaring, auroral activity and the effects of summer months in human activities. The National Oceanic and Atmospheric Administration (NOAA) along with the National Geophysical Data Center (NGDC) solved this problem by isolating natural light activity, thus providing stable lights datasets to the public. The publicly available dataset consists of a yearly dataset-grid in which every grid reports the light intensity as a six bit digital number (DN), for every 30 arc second output pixel between 65° south and 75° north latitude. The DN is an integer number in the interval [0, 63], where 0 denotes no light activity and 63 is the strongest recorded intensity "generally in rich and dense areas." These issues, among others, are discussed by HSW.

The analysis carried out in the present study makes use of HSW's database (see HSW for a detailed explanation of the database). It is important to remark that satellites in orbit have a lifetime between 10 and 15 years approximately, due to fuel availability that allows them to stay fix in a specific low orbit in order to pre-

Fig. 1 Satellite data used for Mexico's DNs calculations. *Source*: Author's elaboration with data from NOAA



vent movements produced by solar dust, space debris and solar winds, among other factors. We found in HSW's database that information on DNs was excluded for several countries in year 2003 because satellite F15 (where F stands for flight number) had a shortfall in the measurement of night light activity, as can be appreciated for Mexico in Fig. 1. The measurements of satellites introduced in different years, F10, F12, F14 and F16 that appear also in that figure show alternative measurements whose average leads to the DN data employed by HSW. The level of light activity is clearly different for each satellite, but it is growth what really matters. For detailed information on how the satellite data were produced we recommend the reader to see Elvidge et al. (2009). The shortfall of satellite F15 is a finding that led us to look for a robust statistical procedure based on medians that could mitigate the effect of outliers on growth estimation, instead of excluding information, as HSW did.

Previous work by Chen and Nordhaus (2011) and HSW found similar results for countries with high-quality data (generally high-income countries) where nighttime lights add little value, if any, to official growth measurement figures. In fact, Chen and Nordhaus (2011) indicate that lights add low value to official growth figures in countries with the best statistical systems because, as they established: "the lights data have high measurement errors while the estimated measurement errors in the standard economic data are relatively small." On the other hand, lights data are more useful for improving growth measurement in countries with the poorest statistical systems. Underreporting in most cases (see, for instance, Ghosh et al. 2009) and over-estimation to hit official GDP targets (as some people think it happens in China, e.g., Rawski 2001) in some others are likely causes for such a behavior in developing countries, but finding an explanation for this statistical fact is beyond the scope of the present work.

This paper is organized as follows. Section 2 briefly presents the original HSW model specification and its statistical estimation procedure. Section 3 contains our proposal for model estimation and its statistical justification. The basic idea is to use time series data for a single country at a time. For illustrative purposes, in Sect. 4 we present the results of applying the suggested methodology to data from Mexico, China and Chile. Some findings are as follows: the official figures for Mexico and Chile tend to underestimate true GDP growth, whereas those of China over-estimate it. Section 5 concludes with some final remarks.

2 HSW model specification

Here, we specify HSW's model with a slightly different notation from that used in the original paper. For country j and year t, let $Y_{j,t}$ be the true real GDP, $Z_{j,t}$ be the officially reported real GDP figure and $X_{j,t}$ be the measured lights value. We write in small letters the same variables expressed in natural logs, that is, $y_{j,t} = \log(Y_{j,t})$, and similarly for $z_{j,t}$ and $x_{j,t}$. The corresponding growths are expressed as log-differences, that is, $Dy_{j,t} = y_{j,t} - y_{j,t-1}$ for all time periods under study, $t = 2,\ldots, N$ since $y_{j,0}$ is unavailable. Besides, we omit the time index when we deem it unnecessary for the exposition. Now, according to HSW, the growths are linked by means of

$$Dz_{j,t} = Dy_{j,t} + \varepsilon_{Dz,j,t}, \tag{1}$$

where $\varepsilon_{Dz,j,t}$ is a random measurement error with variance σ^2_{Dz}, and by

$$Dx_{j,t} = \beta Dy_{j,t} + \varepsilon_{Dx,j,t}, \tag{2}$$

with $\varepsilon_{Dx,j,t}$ another random error with variance σ^2_{Dx} and uncorrelated with $\varepsilon_{Dz,j,t}$.

To predict income growth from light growth HSW employed the following equation,

$$\widehat{Dz}_{j,t} = \widehat{\Psi} Dx_{j,t}, \tag{3}$$

where $\widehat{\Psi} = \widehat{\mathrm{Cov}}\,(Dx, Dz)\,/\hat{\sigma}^2_{Dx}$ is the ordinary least squares (OLS) estimator of the inverse elasticity of lights with respect to income. This estimator was shown by HSW to be biased and inconsistent for $1/\beta$ since

$$p\lim\left(\widehat{\Psi}\right) = \frac{1}{\beta}\left(\frac{\beta^2\sigma^2_{Dy}}{\beta^2\sigma^2_{Dy} + \sigma^2_{Dx}}\right). \tag{4}$$

Nevertheless, to take advantage of the lights for predicting income growth, HSW considered the growth values produced by (3) as *proxies* and used them to improve on the growth obtained from official figures by means of a linear combination of $Dz_{j,t}$ and $\widehat{Dz}_{j,t}$, without further justification. That is, they suggested using the weighted average

$$\widehat{Dy}_{j,t} = \lambda Dz_{j,t} + (1-\lambda)\,\widehat{Dz}_{j,t} \quad \text{with} \quad \lambda \in (0, 1)\,. \tag{5}$$

Of course, many other types of averages could be used, e.g., a weighted harmonic average $\widetilde{Dy}_{j,t} = 1/[(\lambda/Dz_{j,t} + (1-\lambda)\,/\widehat{Dz}_{j,t})]$ with $\lambda \in (0, 1)$ or even a simple arithmetic average $\widetilde{Dy}_{j,t}\left(Dz_{j,t} + \widehat{Dz}_{j,t}\right)/2$.

A nice feature of using (5) is that an optimal λ value can be chosen as the minimizer of the variance of the prediction error, given by

$$\mathrm{Var}\left(\widehat{Dy}_{j,t} - Dy_{j,t}\right) = \lambda^2\sigma^2_{Dz} + (1-\lambda)^2\,\frac{\sigma^2_{Dy}\sigma^2_{Dx}}{\beta^2\sigma^2_{Dy} + \sigma^2_{Dx}}, \tag{6}$$

so that the optimal value becomes

$$\widetilde{\lambda} = \frac{\sigma_{Dy}^2 \sigma_{Dx}^2}{\sigma_{Dz}^2 \left(\beta^2 \sigma_{Dy}^2 + \sigma_{Dx}^2\right) + \sigma_{Dy}^2 \sigma_{Dx}^2}. \tag{7}$$

To use this expression, HSW defined the "signal-to-total variance" ratio

$$\phi = \frac{\sigma_{Dy}^2}{\sigma_{Dy}^2 + \sigma_{Dz}^2} \tag{8}$$

in such a way that the value of $\widetilde{\lambda}$ can get fixed by choosing the value of ϕ. In order to assign values to ϕ for each country, they made use of the relative quality ratings of national income data provided by the IMF and the WB (where countries are classified as good data countries or bad data countries). Once the value of ϕ was chosen, the linear combination was applied to estimate GDP growth. For details on the derivation of the expressions shown in this section, the reader should refer to the original HSW paper.

We stress the fact that the estimation procedure employed by HSW relies on the use of relative quality ratings of national income data to weight the evidence provided by both official figures. We consider that fact unnecessary, since the data for each individual country have enough information to assign an appropriate weight to the available statistical evidence, as shown in the next section. Besides, the very idea of using data from other countries to improve on the estimation of a given country's GDP is contrary to the spirit underlying National Accounts calculations, which should be based only on data pertaining to the country under study.

3 Estimating true GDP growth of an individual country

In this section, we derive country-specific optimal weights λ without relying on any auxiliary World Bank or IMF estimate of the country's statistics quality. To that end, we consider the existence of yearly time series data on official GDP and nighttime lights for a gin country and follow the basic idea proposed by HSW to estimate true real GDP growth, but we are more explicit about the statistical procedure involved. In fact, our interest lies in the statistical justification of the linear combination (5) and its empirical application to data from a single country at a time, rather than to a panel of countries.

On the one hand, we assume the following signal-plus-noise model for the official GDP figures,

$$Dz_t = Dy_t + \eta_t \quad \text{for} \quad t = 2, \ldots, N, \tag{9}$$

where η_t is the noise that basically obscures the signal Dy_t at each time t, with $\eta_2, \ldots, \eta_N$ a sequence of random errors such that $\mathrm{Cov}\,(\eta_t, \eta_{t'}) = 0$ if $t \neq t'$, with

$E(\eta_t) = 0$ and $\mathrm{Var}(\eta_t) = \sigma^2$. We also assume that $\mathrm{Cov}(Dy_t, \eta_t) = 0$ and $\mathrm{Var}(Dy_t) = \sigma^2_{Dy}$ for $t = 2, \ldots, N$.

The errors in (9) are implicitly assumed to be stationary and explicitly uncorrelated, implying that the discrepancy between official growth and true (unobserved) growth does not carry over from one period to the next one. We believe that in some countries the discrepancy in the level of official and true real GDP does carry over from one period to the next one, but this is not necessarily reflected as a systematic under- or over-reporting of growth, because taking differences tends to reduce or even cancel autocorrelation as shown by Granger and Newbold (1974). Even so, if error autocorrelation is feared to occur, model (9) could be extended to account for an autoregressive structure in the errors. In practice, a significance test of residual autocorrelation should be applied to justify the inclusion of an autoregressive parameter in the error structure.

On the other hand, as in HSW, we suppose a constant elasticity relation between observed nighttime lights and income of the country under study and assume that $X_t = KY_t^{\beta}$ holds, with K a positive constant and β the elasticity of lights with respect to income. Hence, it follows that

$$Dx_t = \beta Dy_t + \varepsilon_t \;\; \text{for} \;\; t = 2, \ldots, N, \tag{10}$$

where $\varepsilon_2, \ldots, \varepsilon_N$ are uncorrelated random errors, with $E(\varepsilon_t) = 0$, $\mathrm{Var}(\varepsilon_t) = \sigma^2_{\varepsilon}$, $\mathrm{Cov}(Dy_t, \varepsilon_t) = 0$ and $\mathrm{Cov}(\eta_t, \varepsilon_t) = 0$ for $t = 2, \ldots, N$. It may happen that the errors in an equation relating x_t to y_t, as indicated by Nordhaus and Chen (2014), show some autocorrelation structure because satellites change slowly over time, for example in the timing of their orbits. However, we suppose again that the time differences cancel such potential autocorrelation and therefore the errors in (10) are assumed uncorrelated. Now, if we let $D\mathbf{y} = (Dy_2, \ldots, Dy_N)'$, $D\mathbf{z} = (Dz_2, \ldots, Dz_N)'$ and $D\mathbf{x} = (Dx_2, \ldots, Dx_N)'$ be $(N-1)$-dimensional column vectors containing growths of true GDP, official GDP and nighttime lights, respectively, we can write (9) and (10) as a system of linear equations (see "Appendix A"). Then, by assuming that the parameters $\beta, \sigma^2_{\varepsilon}$ and $\alpha = \sigma^2/\sigma^2_{\varepsilon}$ are known, we are led to the use of generalized least squares (GLS) to obtain the best linear unbiased estimator (BLUE) of $D\mathbf{y}$, which is given by

$$\widehat{D\mathbf{y}} = \lambda D\mathbf{z} + (1-\lambda)\,\widetilde{D\mathbf{z}} \tag{11}$$

with $\widetilde{D\mathbf{z}} = \beta^{-1} D\mathbf{x}$ and $\lambda = \frac{\alpha^{-1}}{\alpha^{-1}+\beta^2} \in (0, 1)$. Guerrero (2007) provides a justification for using GLS to estimate a random vector rather than a fixed parameter vector. (His deduction was given in a context different from the one considered here, but it is equally valid for this work.) The derivation of expression (11) appears in "Appendix A," and below we focus on a feasible GLS procedure that provides estimates of the unknown parameters involved. We call $\widehat{D\mathbf{y}}$ estimator, but a more adequate term could be predictor since $D\mathbf{y}$ is a random variable, not a constant parameter. Moreover, the variance–covariance matrix of the prediction error is directly obtained from the GLS application, that is

$$\mathrm{Var}\left(\widehat{D\mathbf{y}} - D\mathbf{y}\right) = \sigma_{\varepsilon}^{2}\left(\alpha^{-1} + \beta^{2}\right)^{-1} I_{N-1}. \tag{12}$$

We should notice that (11) is basically the same result at which HSW arrived, but we followed a different statistical approach that: (a) makes efficient use of both sources of information; (b) justifies the linear combination of official growth and nighttime lights growth; and (c) shows how the weight λ is related to β and α. Besides, the GLS application also leads to the following estimator (see "Appendix B")

$$\hat{\sigma}_{\varepsilon}^{2} = \frac{\hat{\lambda}\hat{\beta}^{2}}{N-3}\left(D\mathbf{z} - \widetilde{D\mathbf{z}}\right)'\left(D\mathbf{z} - \widetilde{D\mathbf{z}}\right). \tag{13}$$

Here, it is interesting to notice that $\hat{\sigma}_{\varepsilon}^{2} \downarrow 0$ as $\hat{\lambda} \downarrow 0$ (or, equivalently, $\hat{\alpha} \uparrow \infty$). This fact will be used later to produce an estimator of α.

Now, in order to estimate β we use Eqs. (9) and (10) to obtain

$$Dz_t = \beta^{-1}\left(Dx_t - \varepsilon_t\right) + \eta_t = \beta_1 Dx_t + \gamma_t, \tag{14}$$

where

$$\beta_1 = \beta^{-1} \text{ and } \gamma_t = \eta_t - \beta^{-1}\varepsilon_t, \tag{15}$$

with $E(\gamma_t) = 0$ and $\mathrm{Var}(\gamma_t) = \sigma_{\varepsilon}^{2}(\alpha + \beta^{-2})$. Next, we realize, as did HSW, that

$$\mathrm{Cov}\left(Dx_t, \gamma_t\right) = \mathrm{Cov}\left(Dx_t, \eta_t - \beta^{-1}\varepsilon_t\right) = -\beta^{-1}\sigma_{\varepsilon}^{2} \tag{16}$$

So that the OLS estimator $\hat{\beta}_{1,\mathrm{OLS}} = \widehat{\mathrm{Cov}}(Dx_t, Dz_t)/\widehat{\mathrm{Var}}(Dx_t)$ involves $\widehat{\mathrm{Cov}}(Dx_t, Dz_t) = \beta_1\widehat{\mathrm{Var}}\left(Dx_t\right) - \beta_1\hat{\sigma}_{\varepsilon}^{2}$. Hence, it follows that

$$E\left(\hat{\beta}_{1,\mathrm{OLS}}\right) = \beta_1 E\left(\frac{\widehat{\mathrm{Var}}\left(Dx_t\right) - \hat{\sigma}_{\varepsilon}^{2}}{\widehat{\mathrm{Var}}\left(Dx_t\right)}\right) \neq \beta_1, \tag{17}$$

so that $\hat{\beta}_{1,\mathrm{OLS}}$ is a biased estimator of β^{-1} unless $\hat{\sigma}_{\varepsilon}^{2} = 0$.

To correct for this situation, we now look for an unbiased estimator of β_1. First, let us recall that our interest lies in estimating the growth of true GDP, so that expression (14) must not to produce spurious results. This requirement amounts to saying that the sequence of errors $\{\gamma_t\}$ is a weakly stationary time series, which should be considered an assumption to be verified with the data at hand. We now assume that the nighttime lights data $D\mathbf{x}$ are fixed and write model (14) conditional on this information, as

$$E\left(Dz_t | D\mathbf{x}\right) = \beta_1 Dx_t \text{ with } E\left(\gamma_t | D\mathbf{x}\right) = 0 \text{ for } t = 2, \ldots, N. \tag{18}$$

Then, by averaging all the observations we get

$$E\left(\sum_{t=2}^{N} Dz_t/\left(N-1\right)|D\mathbf{x}\right) = \beta_1 \sum_{t=2}^{N} Dx_t/\left(N-1\right) \tag{19}$$

so that a very simple estimator of β_1 is obtained as the ratio of growth averages of official GDP and nighttime lights, that is,

$$\tilde{\beta}_1 = \frac{\sum_{t=2}^{N} Dz_t / (N-1)}{\sum_{t=2}^{N} Dx_t/(N-1)} = \frac{z_N - z_1}{x_N - x_1}, \tag{20}$$

which is unbiased as long as $E\left(\gamma_t | D\mathbf{x}\right) = 0$ for $t = 2, \ldots, N$. However, it only makes use of the long-term growth of both variables involved; therefore, it cannot be considered reliable for estimating annual GDP growth.

As an alternative to the use of $\tilde{\beta}_1$, we propose using the median rather than the arithmetic average of nighttime lights growth, that is, for some integer m we define the median of $D\mathbf{x}$ as

$$\text{Med}\left(D\mathbf{x}\right) = \left\{ \begin{array}{ll} Dx_{(m+1)} & \text{if } N-1 = 2m+1 \\ \left(Dx_{(m)} + Dx_{(m+1)}\right)/2 & \text{if } N-1 = 2m \end{array} \right\}. \tag{21}$$

By using the median of nighttime lights growth, we seek for protection against the influence of anomalous satellite measurements that are likely to occur, as mentioned in the introduction. Hence, the proposed estimator becomes

$$\hat{\beta}_1 = \left\{ \begin{array}{ll} \frac{Dz\left\{Dx_{(m+1)}\right\}}{Dx_{(m+1)}} & \text{if } N-1 = 2m+1 \\ \frac{Dz\left\{Dx_{(m)}\right\} + Dz\left\{Dx_{(m+1)}\right\}}{Dx_{(m)} + Dx_{(m+1)}} & \text{if } N-1 = 2m \end{array} \right\} \tag{22}$$

where $Dz\{Dx_{(t)}\}$ denotes the official GDP growth corresponding to the (ordered) night lights growth observed at time t. This estimator is clearly unbiased given $D\mathbf{x}$, because $E\left(\hat{\beta}_1 | D\mathbf{x}\right) = \frac{1}{Dx_{(m+1)}} E\left(Dz\{Dx_{(m+1)}\} | D\mathbf{x}\right) = \beta_1$ if $N-1 = 2m+1$, and similarly when $N-1 = 2m$.

Once the parameter estimate $\hat{\beta} = \hat{\beta}_1^{-1}$ is obtained, we can generate the unbiased *proxy*

$$\widetilde{Dz}_t = \hat{\beta}_1 Dx_t \quad \text{for} \quad t = 2, \ldots, N, \tag{23}$$

which will be used only as a preliminary estimate since it will be combined with the official GDP growth figures by means of expression (11). In order to apply that expression, we have to provide an estimate of the remaining parameter α. Then, it is tempting to use an automatic method to estimate it, such as generalized cross-validation (GCV), see Ruppert et al. (2003, p. 117). Such a method indicates to choose α by minimizing the goodness of fit criterion given by

$$\text{GCV}\left(\alpha\right) = \frac{\text{RSS}\left(\alpha\right)}{\left[1 - \text{tr}\left(H\right)/\left(N-1\right)\right]^2} \tag{24}$$

where RSS(α) is the residual sum of squares that depends on the parameter α. Here, tr(H) denotes the trace of the matrix H known as the "Hat matrix" and it is shown in

"Appendix B" that $\mathrm{tr}(H)=N-1$, so that $\mathrm{GCV}(\alpha)$ gets undetermined in the present situation.

Hence, we propose to estimate α by using the expression for the weight derived from GLS, i.e., $\lambda = \alpha^{-1}/\left(\alpha^{-1}+\beta^2\right)$, so that, once $\hat{\beta}^2$ has been obtained, we can get the estimate

$$\hat{\alpha} = (1-\lambda)/\hat{\beta}^2\lambda \tag{25}$$

by selecting an appropriate value of $\lambda \in (0, 1)$. To do this, we should look at the sensitivity of the results for different choices of λ. Thus, we suggest to take into account the information provided by two standard error intervals for true GDP growth and select λ as the smallest value that makes true the probabilistic assertion provided by Tchebysheff's theorem (see, e.g., Wackerly et al. 2002, Ch. 4). This choice of λ will render the smallest value for the variance estimate $\hat{\sigma}_\varepsilon^2$ as we established after Eq. (13). Thus, for $t=2, \ldots,$ N, it must be true that

$$\Pr\left[\left|\widehat{Dy}_t - Dy_t\right| \geq 2\hat{\sigma}_\varepsilon\left(\hat{\alpha}^{-1}+\hat{\beta}^2\right)^{-1/2}\right] \leq 1/4, \tag{26}$$

in such a way that a time series of official growth figures $\{Dz_t\}$ is close enough to $\{Dy_t\}$ (so as to be considered a good representation of true GDP growth) if at most ¼ of the observations of $\{Dz_t\}$ are outside the intervals given by

$$\widehat{Dy}_t \pm 2\sqrt{\hat{\sigma}_\varepsilon^2\left(\hat{\alpha}^{-1}+\hat{\beta}^2\right)^{-1}} \quad \text{for } t=2,\ldots,N. \tag{27}$$

It should be noticed that (12) and (13) imply that the variance of the prediction error is given by

$$\begin{aligned}\mathrm{Var}\left(\widehat{Dy}_t - Dy_t\right) &= \hat{\sigma}_\varepsilon^2\left(\hat{\alpha}^{-1}+\hat{\beta}^2\right)^{-1} \\ &= \frac{\hat{\lambda}\hat{\beta}^2}{(N-3)\left(\hat{\alpha}^{-1}+\hat{\beta}^2\right)}\left(D\mathbf{z}-\widehat{D\mathbf{z}}\right)'\left(D\mathbf{z}-\widehat{D\mathbf{z}}\right)\end{aligned} \tag{28}$$

then, since $\hat{\alpha}^{-1} = \hat{\lambda}\hat{\beta}^2/\left(1-\hat{\lambda}\right)$, we get

$$\mathrm{Var}\left(\widehat{Dy}_t - Dy_t\right) = \frac{\hat{\lambda}\left(1-\hat{\lambda}\right)}{N-3}\left(D\mathbf{z}-\widetilde{D\mathbf{z}}\right)'\left(D\mathbf{z}-\widetilde{D\mathbf{z}}\right), \tag{29}$$

so that $\mathrm{Var}(\widehat{Dy}_t - Dy_t)$ grows to $\frac{0.25}{N-3}\left(D\mathbf{z}-\widetilde{D\mathbf{z}}\right)'\left(D\mathbf{z}-\widetilde{D\mathbf{z}}\right)$ as $\hat{\lambda} \to 1/2$ and it decreases to 0 as $\hat{\lambda} \downarrow 0$ or $\hat{\lambda} \uparrow 1$. That is, maximum uncertainty occurs when $\hat{\lambda} \to 1/2$, corresponding to the case of equal weighting of both economic growths (measured with satellite data and with official figures), while the uncertainty gets reduced symmetrically when more weight is assigned to either one of the two sources.

It is convenient to state the method for obtaining $\hat{\lambda}$ explicitly.

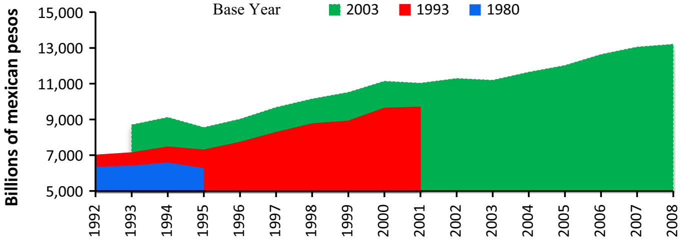


Fig. 2 Annual GDP for Mexico with three different base Years (at 2008 prices). *Source*: Authors' elaboration with data from Banco de Información Económica, INEGI: http://www.inegi.org.mx/sistemas/bie/. Accessed on May 02, 2014

Step 1: Estimate $\hat{\beta} = \hat{\beta}_1^{-1}$ where $\hat{\beta}_1$ is given by (22).
Step 2: Choose a λ value between 0 and 1.
Step 3: Obtain $\hat{\alpha}$ via expression (25) and $\widehat{D\mathbf{y}}$ by means of (11).
Step 4: Calculate 2 standard error intervals for true GDP growth via expression (27) and count how many elements of $\{Dz_t\}$ lie outside the intervals.
Step 5: Repeat steps 2, 3 and 4 for various values of λ. The smallest λ that leaves ¼ of the elements of $\{Dz_t\}$ outside the Tchebysheff bounds is the optimal $\hat{\lambda}$.

Finally, we stress the following statistical facts: first, the estimator required to calculate the *proxy* from light growth is robust against the presence of outliers in the nighttime lights data. Second, since no distributional assumption is made for estimating true GDP growth, we cannot make statistical inferences, such as confidence intervals for parameters or prediction intervals for true annual GDP growth.

4 Empirical applications

4.1 Mexican GDP growth

Mexico presents a situation where the official figures of GDP growth can be appropriately complemented by luminosity data. We think so because revisions of official figures have indicated higher economic growth than originally reported. This can be seen in Fig. 2, where GDP is plotted with three different base years (1980, 1993 and 2008) for the sample period 1992–2008 that covers the data used to illustrate our proposed method). What we want to stress is that the revised GDP figures are higher than the original ones, not how much the revisions have changed the original figures.

Figure 3 shows the light growth in Mexico during years 1993–2008, where the trough in years 1996–1997 is reasonable because an economic crisis took place in 1995. On the contrary, the trough in year 2003 is not associated with an economic crisis but to a failure of the satellite that measured the DNs, as mentioned in the introduction. Therefore, to estimate true GDP growth for Mexico we use the robust estimation procedure that seeks to minimize the influence of such anomalous observation rather than eliminating it. We should recall that the suggested procedure uses the median of nighttime lights growth and the corresponding GDP growth value(s) associated with

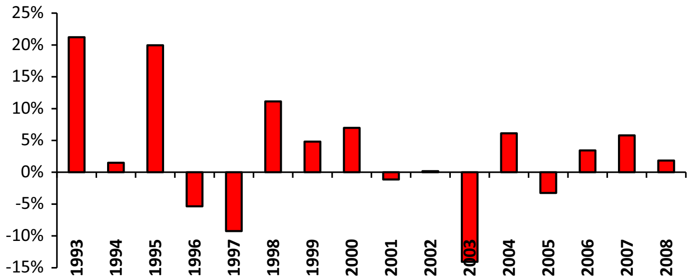


Fig. 3 Annual growth of average DNs for Mexico during 1993–2008

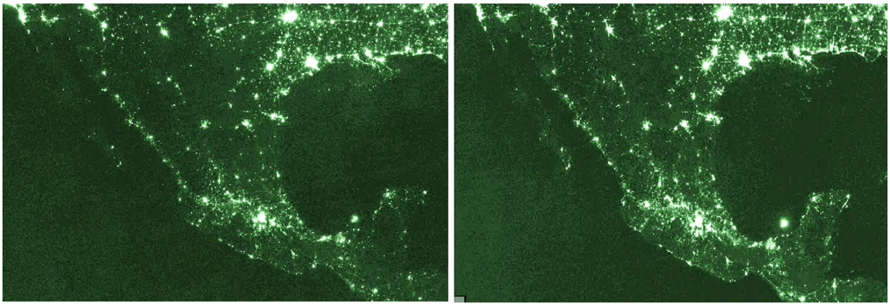


Fig. 4 Nighttime lights in Mexico, 1992 (left) and 2008 (right)

it to estimate the inverse elasticity parameter appearing in Eq. (10). The original data employed in this application appear in "Appendix C".

Figure 4 allows us to get a visual appreciation of the nighttime lights growth between 1992 and 2008 in Mexico. No light intensity is provided, but it should be clear that the whole country got brighter over time and a heterogeneous change can be appreciated at the subnational level (even though the focus of our paper is on national change).

For years 1992–2008, Med$(D\mathbf{x})$ was given by the average of nighttime light growths for 2006 and 2008, so that $\hat{\beta}_1 = 1.2377$; hence, light elasticity with respect to income was estimated as $\hat{\beta} = 0.808$. To verify the assumption of no error serial correlation, we estimated the autoregressive coefficient of order one for the residuals and obtained the value -0.24 with standard error 0.29 and p value 0.42, so that this coefficient cannot be considered significantly different from zero, a fact that lends empirical support to the assumption. Next, in Table 1 we show the results of $\hat{\alpha}$ and the estimated error variance $\hat{\sigma}_\varepsilon^2$ for different choices of the weight λ using the previously obtained elasticity estimate, where we see that $\hat{\sigma}_\varepsilon^2$ gets smaller as λ decreases, as expected.

Figure 5 presents ± 2 standard error intervals obtained with four different choices of λ. For large values of this weight, the estimated true GDP lies very close to the official GDP figure, while the opposite occurs when λ gets smaller. Thus, we can see that with $\lambda = 0.7$ no observation lies outside the band, with $\lambda = 0.5$ only one observation is outside the band, with $\lambda = 0.3$ there are four observations outside the band and with $\lambda = 0.1$ there are five.

Table 1 Estimation results for Mexico with $\hat{\beta} = 0.808$ and different λ values

λ	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
$\hat{\alpha}$	0.170	0.383	0.656	1.021	1.532	2.298	3.574	6.127	13.786
$\hat{\sigma}_{\varepsilon}^{2}$	0.010	0.009	0.008	0.007	0.006	0.004	0.003	0.002	0.001

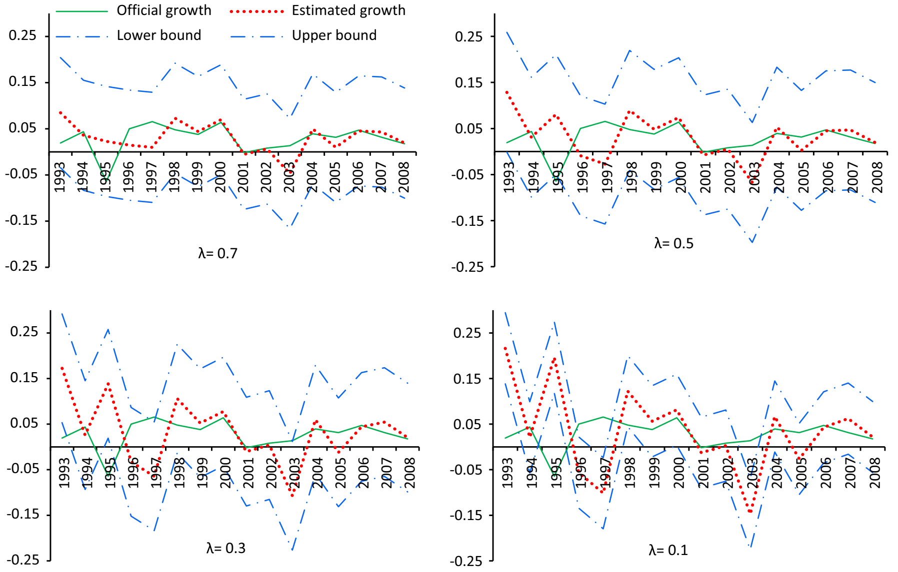


Fig. 5 ±2 Standard error intervals for λ = 0.7, 0.5, 0.3 and 0.1

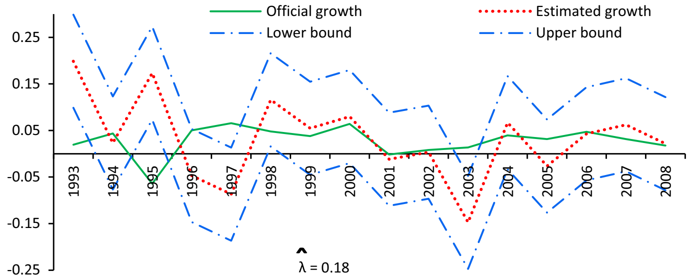


Fig. 6 ±2 Standard error intervals for Mexico's true GDP with $\hat{\lambda} = 0.18$

In Fig. 6 we can see four observations outside the band, which is the number we are looking for, an integer value no larger than $(N-1)/4 = 4$. In this case, $\hat{\lambda} = 0.18$ is the smallest value that produces such a result—with $\hat{\lambda} = 0.17$ there are five values outside the intervals; thus, we are led to conclude that this is an appropriate estimate.

The estimated true real GDP plotted in Fig. 6 allows us to calculate the average estimated true GDP growth for years 1992–2008 as 3.27%, whereas the corresponding official GDP average growth is 2.82%. Hence, Mexico's economic growth has been

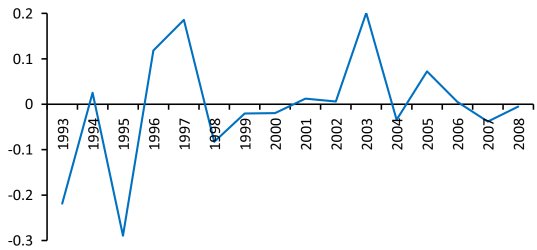


Fig. 7 Residual series $\{Dz_t - \widetilde{Dz_t}\}$ for Mexico's true GDP

underestimated in average 0.45% per year during the period under study. For comparative purposes, we should recall that HSW obtained an underestimation of 0.43% per year, so that their results and ours are very close to each other.

Now, to validate the previous application of the statistical procedure with respect to the stationarity assumption, we checked for the presence of unit roots in the residual series $\{Dz_t - \widetilde{Dz_t}\}$ shown in Fig. 7. We did this by applying the nonparametric test of Phillips and Perron (1988) and obtained the following calculated statistics: with 0 lags, -5.31, with 1 lag, -5.34, and with 2 lags, -5.67, all of which are significant and led us to reject the null hypothesis of unit root at the 1% significance level.

Moreover, it was also assumed that the error term in expression (14) has mean zero and this assumption was verified by means of a Wilcoxon signed-rank two-sided test with paired data $(Dz_t, \widetilde{Dz_t})$, see Wackerly et al. (2002, pp. 716–719). The result of such a test is that the sum of positive ranks equals 68, whereas the critical 10% value is 37 for $n = N - 1 = 16$ pairs of observations. Thus, there is no evidence against the null hypothesis that both series have the same location, leading to the conclusion that the difference is not significantly different from zero, even at the 10% level. Therefore, the assumptions underlying the statistical procedure were validated empirically.

4.2 The cases of China and Chile

Mexico was our leading case study, but the same procedure employed with Mexico can be applied to other countries as well. In recent years, the credibility of China's official figures has been in doubt, e.g., Rawski (2001) carried out an investigation aimed at identifying the most salient features of China's statistics. He found several inconsistencies such as a fall in electricity consumption accompanied by a rise in GDP, while in the 1950s the same country showed positive growth in both GDP and electricity consumption. This could be a good reason to employ China's data to illustrate the usefulness of the proposed procedure.

Another illustrative application is made with data from Chile, a developing Latin-American country with high economic growth. From 1985 to 1997, Chile experienced a "golden age" growth; nevertheless, global institutions like the IMF anticipate that emerging countries that basically export raw materials are prone to have a fall in their GDP growth. The original data employed in these applications appear in "Appendix C".

Table 2 Parameter estimation results for China and Chile

China	Med $(D\mathbf{x})$ is the average of 1993 and 2000 growths, $\hat{\beta}_1 = 1.204$
Chile	Med $(D\mathbf{x})$ is the average of 2001 and 2007 growths, $\hat{\beta}_1 = 1.139$

Table 3 Estimation results for China and Chile, for different λ values

λ	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
China with $\hat{\beta} = 0.831$									
$\hat{\alpha}$	0.161	0.362	0.621	0.966	1.449	2.174	3.381	5.796	13.042
$\hat{\sigma}_\varepsilon^2$	0.006	0.006	0.005	0.004	0.003	0.003	0.002	0.002	0.001
Chile with $\hat{\beta} = 0.878$									
$\hat{\alpha}$	0.144	0.324	0.556	0.865	1.298	1.947	3.028	5.191	11.680
$\hat{\sigma}_\varepsilon^2$	0.012	0.010	0.009	0.008	0.007	0.005	0.004	0.003	0.001

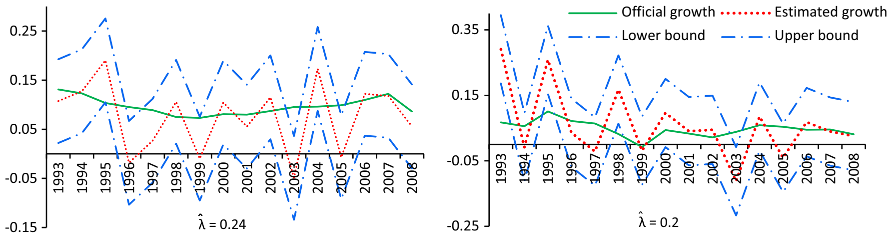


Fig. 8 ±2 Standard error intervals for true GDP of China with $\hat{\lambda} = 0.24$ and Chile with $\hat{\lambda} = 0.2$

In Table 2, we present the parameter estimates of Eq. (21) from which the light elasticity is estimated as $\hat{\beta} = 0.831$ for China and $\hat{\beta} = 0.878$ for Chile. The assumption of no serial correlation was also verified by calculating the autoregressive coefficient of order one for the residuals. We obtained -0.48 and -0.54 with standard errors 0.23 and 0.29, and p values 0.06 and 0.08 for China and Chile, respectively, so that none of those coefficients is significantly different from zero, a fact that again lends empirical support to the assumption.

Next, in Table 3 we show the estimation results of the parameters α and σ_ε^2, for different choices of the weight λ. It is worth noticing that the estimate $\hat{\sigma}_\varepsilon^2$ reduces its value around 85–90% as the weight diminishes from 0.9 to 0.1.

To decide an appropriate weight for the combination of nighttime lights growth with official GDP growth, we searched for the weight that leaves 4 data points of estimated log(GDP) outside the ± 2 standard error intervals. That happened with $\hat{\lambda} = 0.24$ for China and with $\hat{\lambda} = 0.2$ for Chile. The corresponding plots appear in Fig. 8.

The underlying statistical assumptions for these two applications were validated as with the Mexican case by means of a unit-root Phillips–Perron test and a Wilcoxon signed-rank two-sided test with paired data. The results of these tests are summarized

Table 4 Validating the underlying assumptions for the models of China and Chile

Phillips–Perron				Wilcoxon	
Lags	0	1	2	Sum of negative ranks	Critical 10% value
China	−8.01	−8.52	−9.53	55	37
Chile	−7.16	−7.01	−6.91	63	37

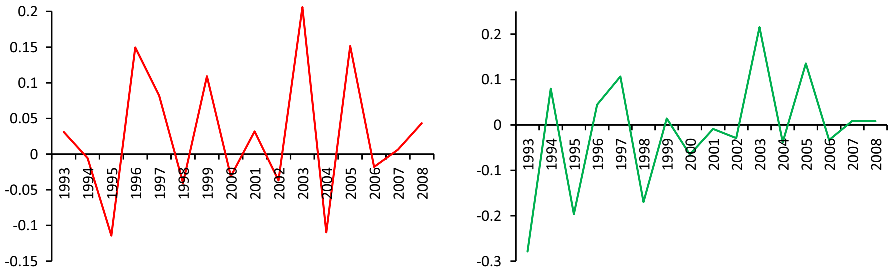


Fig. 9 Residual series $\{Dz_t - \widetilde{Dz_t}\}$ for China (left) and Chile (right)

in Table 4, where we see that the calculated statistics are highly significant (at the 1% level the critical value is -3.92) for the null hypothesis of unit root, so that such a hypothesis is clearly rejected. The corresponding plots of the residual series appear in Fig. 9, where we can see a reasonable stationary behavior for the two series. On the other hand, Table 4 also allows us to see that the assumption of zero mean for the residual series cannot be rejected even at the 10% significance level.

The estimated true GDP for China with the optimal parameter values previously decided produced an estimated true GDP growth of 7.61% which, when compared with the official GDP growth of 9.66%, yields an average over-estimation of 2.06% per year. This result reinforces the lack of credibility on the economic growth of this Asian country.

On the contrary, for the same period, the Chilean figures are: official average GDP growth of 4.72% and estimated true GDP growth of 5.94%, yielding a yearly sub-estimation of 1.22%. These results lead us to say that the method is able to detect both underestimation and over-estimation of official GDP figures and helps to correct both types of bias. As a final comment, we mention that the standard errors of the estimated true GDP for Mexico, China and Chile are, respectively, 0.050, 0.043 and 0.052 which are very close to each other, indicating basically the same uncertainty in the estimation results for these three countries.

5 Final remarks

Measuring GDP growth is a laborious task that gets even more complicated in countries where the informal sector represents a considerable proportion of economic activity, as seen in developing countries. HSW's model provides a way to improve on official GDP growth figures, but as we have previously stated we propose to follow an approach that employs the same variable (light intensity) for all countries in an individual manner. Instead of requiring quality data ratings or the same weights for nighttime lights growth for all countries, our method allows for estimation of parameters and weights for each country individually and is robust enough to make comparisons among economies. By using this method, we argue that more credibility is placed on the use of satellite measurements for estimating true GDP growth.

Economic indicators provided by some countries lack credibility because macroeconomic analysts tend to believe that lower income countries with high corruption levels have incentives to misrepresent their true economic conditions. This problem could be reduced by using nighttime lights data combined with official figures for every country as a complementary measure of growth. An assessment of our proposal was obtained with data from three different economies, and our results showed some underestimation of GDP growth for two of them (Mexico and Chile) and over-estimation for the other one (China).

The basis of our combined GDP growth estimation method is HSW's model, but it improves the statistical estimation by providing robust estimates of the parameters and the *proxy* that is combined with official GDP figures. Moreover, data requirements are also more concise using the method we propose since only data for the country under study are employed and no official data quality ratings or data from other countries are needed.

Finally, it should be clear that nighttime lights data may also be useful to measure economic growth in sub- or supranational regions. In fact, we support HSW's idea that "empirical growth analyses need no longer be tied exclusively to the availability of national income data" because of new developments of spatial analytical tools.

Appendix A: Obtaining the BLUE of $D\mathbf{y}$

Let I be the $(N-1)$-dimensional identity matrix, $\boldsymbol{\eta} = (\eta_2, \ldots, \eta_N)'$ and $\boldsymbol{\varepsilon} = (\varepsilon_2, \ldots, \varepsilon_N)'$, so that models (9) and (10) can be written as the system of linear equations

$$\begin{pmatrix} D\mathbf{z} \\ D\mathbf{x} \end{pmatrix} = U D\mathbf{y} + \begin{pmatrix} \boldsymbol{\eta} \\ \boldsymbol{\varepsilon} \end{pmatrix}$$

with $U = \begin{pmatrix} I \\ \beta I \end{pmatrix}$, $E\begin{pmatrix} \boldsymbol{\eta} \\ \boldsymbol{\varepsilon} \end{pmatrix} = \mathbf{0}_{2(N-1)}$ the zero vector of size $2(N-1)$ and

$$\mathrm{Var}\begin{pmatrix} \boldsymbol{\eta} \\ \boldsymbol{\varepsilon} \end{pmatrix} = \begin{pmatrix} \sigma^2 I & 0_{N-1} \\ 0_{N-1} & \sigma_\varepsilon^2 I \end{pmatrix} = \sigma_\varepsilon^2 \begin{pmatrix} \alpha I & 0_{N-1} \\ 0_{N-1} & I \end{pmatrix} = \sigma_\varepsilon^2 \Omega,$$

where 0_{N-1} is the $(N-1)$-dimensional zero matrix and $\alpha = \sigma^2/\sigma_\varepsilon^2$.

Then, by assuming that the parameters β, σ_ε^2 and α are known, we apply GLS to obtain the best linear unbiased estimator (BLUE) of $D\mathbf{y}$, which is given by

$$\widehat{D\mathbf{y}} = \left(\mathrm{U}'\Omega^{-1}U\right)^{-1} U'\Omega^{-1} \begin{pmatrix} D\mathbf{z} \\ D\mathbf{x} \end{pmatrix} = \left(\alpha^{-1} + \beta^2\right)^{-1} \left(\alpha^{-1} D\mathbf{z} + \beta D\mathbf{x}\right),$$

that leads us to expression (11).

Appendix B: Unbiased estimation of σ_ε^2

Within the context of GLS, we can get an unbiased estimator of σ_ε^2 by defining the residual vectors of models (9) and (10) as

$$\hat{\boldsymbol{\eta}} = D\mathbf{z} - \widehat{D\mathbf{y}} \text{ and } \hat{\boldsymbol{\varepsilon}} = D\mathbf{x} - \beta\widehat{D\mathbf{y}}$$

so that the stacked residual vector becomes

$$\begin{aligned} \begin{pmatrix} \hat{\boldsymbol{\eta}} \\ \hat{\boldsymbol{\varepsilon}} \end{pmatrix} &= \begin{pmatrix} D\mathbf{z} \\ D\mathbf{x} \end{pmatrix} - U\widehat{\mathrm{D}\mathbf{y}} \\ &= \left[\begin{pmatrix} I & 0_{N-1} \\ 0_{N-1} & I \end{pmatrix} - \left(\alpha^{-1} + \beta^2\right)^{-1} U \left(\alpha^{-1} I \beta I\right)\right] \begin{pmatrix} D\mathbf{z} \\ D\mathbf{x} \end{pmatrix} \\ &= \left[\begin{pmatrix} I & 0_{N-1} \\ 0_{N-1} & I \end{pmatrix} - \left(\alpha^{-1} + \beta^2\right)^{-1} \begin{pmatrix} \alpha^{-1} I & \beta I \\ \beta\alpha^{-1} I & \beta^2 I \end{pmatrix}\right] \left[U D\mathbf{y} + \begin{pmatrix} \boldsymbol{\eta} \\ \boldsymbol{\varepsilon} \end{pmatrix}\right] \end{aligned}$$

Now, since

$$\left(\alpha^{-1} + \beta^2\right)^{-1} \begin{pmatrix} \alpha^{-1} I & \beta I \\ \beta\alpha^{-1} I & \beta^2 I \end{pmatrix} U = U$$

we get

$$\begin{pmatrix} \hat{\boldsymbol{\eta}} \\ \hat{\boldsymbol{\varepsilon}} \end{pmatrix} = \left[\begin{pmatrix} \mathrm{I} & 0_{\mathrm{N-1}} \\ 0_{\mathrm{N-1}} & \mathrm{I} \end{pmatrix} - \mathrm{H}\right] \begin{pmatrix} \boldsymbol{\eta} \\ \boldsymbol{\varepsilon} \end{pmatrix},$$

with

$$H = \left(\alpha^{-1} + \beta^2\right)^{-1} \begin{pmatrix} \alpha^{-1} I & \beta I \\ \beta\alpha^{-1} I & \beta^2 I \end{pmatrix}$$

a symmetric and idempotent matrix whose trace, $\mathrm{tr}(H) = N - 1$, is the number of degrees of freedom for the residual sum of squares, given by

$$\begin{pmatrix} \hat{\boldsymbol{\eta}} \\ \hat{\boldsymbol{\varepsilon}} \end{pmatrix}' \begin{pmatrix} \alpha^{-1} I & 0_{N-1} \\ 0_{N-1} & I \end{pmatrix}^{-1} \begin{pmatrix} \hat{\boldsymbol{\eta}} \\ \hat{\boldsymbol{\varepsilon}} \end{pmatrix} = \alpha^{-1} \hat{\boldsymbol{\eta}}'\hat{\boldsymbol{\eta}} + \hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}.$$

The matrix H is called the "hat matrix" because it transforms the original data $D\mathbf{z}$ and $D\mathbf{x}$ into the estimated data $\widehat{D\mathbf{z}}$ and $\widehat{D\mathbf{x}}$ by means of the following relationship

$$\begin{pmatrix} \widehat{D\mathbf{z}} \\ \widehat{D\mathbf{x}} \end{pmatrix} = U\widehat{D\mathbf{y}} = H \begin{pmatrix} D\mathbf{z} \\ D\mathbf{x} \end{pmatrix}.$$

Thus, if α and β were known, we obtain the following unbiased estimator

$$\hat{\sigma}_{\varepsilon}^{2} = \frac{1}{N-1}\left(\alpha^{-1}\hat{\boldsymbol{\eta}}'\hat{\boldsymbol{\eta}} + \hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}\right).$$

In practice, we have to use the estimators $\hat{\alpha}^{-1}$ and $\hat{\beta}^2$, in place of the true parameters α^{-1} and β^2, so that two degrees of freedom are lost and the variance estimator becomes expression (13), where we should notice that $\hat{\boldsymbol{\varepsilon}} = \hat{\beta}(\widehat{D\mathbf{z}} - \widehat{D\mathbf{y}})$.

Appendix C: Original data employed in the empirical applications

Year	Mexico		China		Chile	
	Ln(DN)	Ln(GDP)	Ln(DN)	Ln(GDP)	Ln(DN)	Ln(GDP)
1992	0.356540	29.3690	−0.349342	29.1549	−1.201073	31.0451
1993	0.548981	29.3883	−0.266380	29.2859	−0.896996	31.1126
1994	0.563912	29.4319	−0.159449	29.4090	−0.918416	31.1681
1995	0.745930	29.3677	0.021317	29.5125	−0.657251	31.2691
1996	0.690892	29.4179	−0.023675	29.6078	−0.633634	31.3406
1997	0.593879	29.4834	−0.017993	29.6967	−0.671262	31.4046
1998	0.699476	29.5313	0.078083	29.7718	−0.494210	31.4364
1999	0.746560	29.5693	0.048182	29.8451	−0.513416	31.4288
2000	0.813918	29.6333	0.141074	29.9257	−0.417595	31.4727
2001	0.802614	29.6317	0.180832	30.0055	−0.380645	31.5059
2002	0.804416	29.6399	0.283795	30.0926	−0.336200	31.5275
2003	0.652805	29.6533	0.191841	30.1879	−0.491460	31.5659
2004	0.712251	29.6927	0.362953	30.2841	−0.406507	31.6246
2005	0.679279	29.7242	0.319271	30.3830	−0.477887	31.6787
2006	0.713111	29.7712	0.425282	30.4928	−0.409212	31.7236
2007	0.769451	29.8027	0.521681	30.6150	−0.376859	31.7693
2008	0.787806	29.8203	0.557354	30.7012	−0.356730	31.8004

References

Aruoba BS, Diebold FX, Nalewaik J, Shorfheide F, Song D (2013) Improving GDP measurements: a measurement-error perspective. National Bureau of Economic Research, Working paper 18954:1–34

Bertinelli L, Strobl E (2013) Quantifying the local economic growth impact of hurricane strikes: an analysis from outer space for the Caribbean. J Appl Meteorol Climatol 52:1688–1697

Chen X, Nordhaus W (2011) Using luminosity data as a proxy for economic statistics. Proc Natl Acad Sci 108(21):8589–8594

Doll C, Muller JP, Elvidge CD (2000) Nightime imagery as a tool for global mapping of socio-economic parameters and greenhouse gas emissions. Ambio 29:157–162

Elliot RJR, Strobl E, Sun P (2015) The local impact of typhoons on economic activity in China: a view from outer space. J Urban Econ 88:50–66

Elvidge CD, Ziskin D, Baugh KE, Tuttle BT, Ghosh T, Pack DW, Erwin EH, Zhizhin M (2009) A fifteen year record of global natural gas flaring derived from satellite data. Energies 2(3):595–622

Ghosh T, Sutton P, Powell R, Anderson S, Elvidge ChD (2009) Estimation of Mexico's informal economy and remittances using nighttime imagery. Remote Sens 1(3):418–444

Granger CWJ, Newbold P (1974) Spurious regressions in econometrics. J Econom 2:111–120

Guerrero VM (2007) Time series smoothing by penalized least squares. Stat Probab Lett 77(12):1225–1234

Harari M, La Ferrara E (2013) Conflict, climate and cells: a disaggregated analysis. Center for Economic Policy Research (CEPR) Discussion paper no. DP9277. SSRN: https://ssrn.com/abstract=2210247. Accessed 21 Oct 2015

Henderson JV, Storeygard A, Weil DN (2012) Measuring economic growth from outer space. Am Econ Rev 102(2):994–1028

Hodler R, Raschky PA (2014a) Economic shocks and civil conflict at the regional level. Econ Lett 124:530–533

Hodler R, Raschky PA (2014b) Regional favoritism. Q J Econ 129:995–1033

Michalopoulos S, Papaioannou E (2014) National institutions and subnational development in Africa. Q J Econ 129(1):151–213

Nordhaus W, Chen X (2014) A sharper image? Estimates of the precision of nighttime lights as a proxy for economic statistics. J Econ Geogr 15:217–246

Phillips PCB, Perron P (1988) Testing for a unit root in time series regression. Biometrika 75:335–346

Rawski TG (2001) What's happening to China's GDP statistics. China Econ Rev 12(4):12–14

Ruppert D, Wand MP, Carroll RJ (2003) Semiparametric regression. Cambridge University Press, Cambridge

Seo N (2011) The impacts of climate change on Australia and New Zealand: a gross cell product analysis by land cover. Aust J Agric Resour 55:220–238

United Nations, European Commission, International Monetary Fund, OCDE and WB (2009) System of national accounts 2008. European Communities, International Monetary Fund, Organisation for Economic Co-Operation and Development, United Nations and World Bank, New York

Wackerly D, Mendenhall W III, Scheaffer RL (2002) Mathematical statistics with applications, 6th edn. Thomson/Brooks-Cole, Grove

World Bank (2002) Building statistical capacity to monitor development progress. World Bank, Washington