

**WP/19/77**

# IMF Working Paper

---

## Illuminating Economic Growth

by Yingyao Hu and Jiaxiong Yao

***IMF Working Papers* describe research in progress by the author(s) and are published to elicit comments and to encourage debate.** The views expressed in IMF Working Papers are those of the author(s) and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.

I N T E R N A T I O N A L M O N E T A R Y F U N D

**IMF Working Paper**

African Department

**Illuminating Economic Growth**

Prepared by Yingyao Hu and Jiaxiong Yao\*

Authorized for distribution by Celine Allard

April 2019

**IMF Working Papers describe research in progress by the author(s) and are published to elicit comments and to encourage debate.** The views expressed in IMF Working Papers are those of the author(s) and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.

**Abstract**

This paper seeks to illuminate the uncertainty in official GDP per capita measures using auxiliary data. Using satellite-recorded nighttime lights as an additional measurement of true GDP per capita, we provide a statistical framework, in which the error in official GDP per capita may depend on the country's statistical capacity and the relationship between nighttime lights and true GDP per capita can be nonlinear and vary with geographic location. This paper uses recently developed results for measurement error models to identify and estimate the nonlinear relationship between nighttime lights and true GDP per capita and the nonparametric distribution of errors in official GDP per capita data. We then construct more precise and robust measures of GDP per capita using nighttime lights, official national accounts data, statistical capacity, and geographic locations. We find that GDP per capita measures are less precise for middle and low income countries and nighttime lights can play a bigger role in improving such measures.

JEL Classification Numbers: E01, E23, C18

Keywords: Nighttime lights; measurement error; GDP per capita.

Author's E-Mail Address: [yhu@jhu.edu](mailto:yhu@jhu.edu); [JYao@imf.org](mailto:JYao@imf.org)

---

\*The authors are grateful to Robert Barbera, Olivier Blanchard, Xiaodong Zhu, and seminar participants at International Monetary Fund, Michigan State University, University of Colorado Boulder, Tsinghua University, Beijing University of Aeronautics and Astronautics, and SEA 2018 annual meeting. The views expressed in this paper are those of the authors and do not necessarily reflect those of the International Monetary Fund, its Board of Directors or the countries they represent. All remaining errors are ours.

<b>Contents</b>		<b>Page</b>
I.	Introduction . . . . .	4
II.	Related Literature . . . . .	8
III.	Statistical Framework . . . . .	10
	A. Baseline Setup . . . . .	10
	B. Sieve Maximum Likelihood Estimation . . . . .	11
	C. Constructing Better Measures of GDP . . . . .	13
	1. Optimal Linear Measure . . . . .	13
	2. Semiparametric Conditional Mean . . . . .	15
IV.	Data . . . . .	16
	A. Nighttime Lights . . . . .	16
	B. Administrative Maps and Country Location . . . . .	17
	C. GDP, Population, and Statistical Capacity . . . . .	18
V.	Results . . . . .	20
	A. Nighttime Lights' Elasticity of Real GDP per capita . . . . .	20
	B. Uncertainty in Real GDP per capita . . . . .	21
	C. New Measures of Real GDP per capita . . . . .	22
	D. Official vs. New Measures . . . . .	25
VI.	Conclusion . . . . .	28
	References . . . . .	29
 Appendices		
A.	Mathematical Proofs . . . . .	31
	A.1. Assumptions . . . . .	31
	A.2. Nonparametric Identification . . . . .	31
	A.3. Proofs . . . . .	33
	A.4. Identification of the Simple Specification . . . . .	35
	A.5. Sieve Maximum Likelihood Estimation . . . . .	36
	A.6. Consistency . . . . .	38
	A.7. Convergence Rates and Asymptotic Normality . . . . .	38
	A.7.1. Convergence Rates of Nonparametric Part . . . . .	39
	A.7.2. Asymptotic Normality of Parametric Part . . . . .	40
B.	Robustness Checks: Simulations, Specifications, and Estimates . . . . .	43
	B.1. Simulations . . . . .	43
	B.2. Alternative Specifications . . . . .	45
	B.3. Gas Flaring in the Data . . . . .	48
	B.4. Optimal Linear Measure for More Countries . . . . .	49
	B.5. Semiparametric Conditional Mean for Robustness Check . . . . .	50

C.	More Descriptive Details on Data . . . . .	50
	C.1. Nighttime Lights vs. Real GDP per capita . . . . .	50
	C.2. Statistical Capacity and Latitude . . . . .	51

#### List of Tables

1.	Summary Statistics . . . . .	19
2.	Summary Statistics . . . . .	19
3.	Estimated Light Production Function . . . . .	20
4.	Estimated Light Production Function . . . . .	20
5.	Official Real GDP per capita and Nighttime Lights 1992-2013 . . . . .	23
6.	Optimal Weights and Mean Squared Error by Income Status . . . . .	25
7.	Parameterization in Simulations . . . . .	43
8.	Simulation Results . . . . .	44
9.	Estimated Light Production Function: Parsimonious Error Structure . . . . .	47
10.	Full Parametric Specification . . . . .	48
11.	Fraction of Nighttime Lights in Gas Flare Shapefiles . . . . .	50

#### List of Figures

1.	Examples of Satellite Images of Nighttime Lights . . . . .	6
2.	Nighttime Lights vs. GDP (1992-2017) . . . . .	11
3.	Map of Nighttime Lights in 2010 . . . . .	17
4.	Estimation Results with Quadratic Light Production Function . . . . .	21
5.	Distribution of Measurement Errors of Real GDP per capita . . . . .	22
6.	Optimal Weights and Real GDP per capita: DMSP/OLS 1992-2013 . . . . .	24
7.	Real GDP per capita: Economic Disruption and Restoration . . . . .	26
8.	New Measures and the Informal Economy: 1992-2013 . . . . .	27
9.	Example of Density Function Approximation . . . . .	44
10.	Semiparametric Conditional Mean and Optimal Linear Measure Performance in Simulations . . . . .	46
11.	Linear Specification . . . . .	47
12.	Quadratic Specification with More Parsimonious Error Structure . . . . .	47
13.	Gas Flares in Nigeria . . . . .	49
14.	New Measures for Low Income Countries . . . . .	51
15.	Semiparametric Conditional Mean and Optimal Linear Measure . . . . .	52
16.	Nighttime Lights vs. Real GDP per capita (1992-2013) . . . . .	53
17.	Real GDP per capita vs. Statistical Capacity by Location (1992-2013) . . . . .	54
18.	Centroids of Selected European Countries . . . . .	55
19.	Real GDP per capita vs. Latitude by Statistical Capacity (1992-2013) . . . . .	56

## I. INTRODUCTION

Real Gross Domestic Product (GDP) is at the heart of macroeconomic analysis and policy-making. It is the basis for measuring national economic development and comparing living standards across countries, and it often serves as a reference point for other economic variables.

Measures of real GDP, however, can be quite uncertain. Lack of statistical capacity, mismeasurement of the economy, and the existence of informal economy, among others, can all subject real GDP measures to substantial revision. This problem becomes more acute for low and middle income countries where the data collection and compilation process is less sophisticated. Understanding the uncertainty of these measures and constructing more accurate measures are therefore of great importance to assess economic performance, facilitate cross-country comparisons, and inform policy decisions.

This paper attempts to use satellite-recorded nighttime lights to illuminate the uncertainty of official measures of real GDP. Mostly generated by human activity, nighttime lights are visible from outer space and recorded by satellites. They have been shown to be correlated with economic activity.<sup>1</sup> Their global coverage and exogenous nature<sup>2</sup> make them attractive as a supplementary measure of real GDP and they are increasingly used in the economics literature. However, despite their economic relevance, nighttime lights may not have a straightforward relationship with real GDP. Meanwhile, like official measures of real GDP, they are subject to measurement errors as well, which further complicates the estimation of the functional relationship.

To illustrate such issues, Figure 1 compares satellite images of nighttime lights for mainland China, the lower 48 states of the United States, and Africa between 1992 and 2013. While all of them became brighter at night in 2013, China's transformation was most visible. Variation in nighttime lights may thus contain useful information on China's real economic growth. In contrast, the United States was already bright enough in 1992. The small change in the intensity of lights over this period may not correspond well to economic growth, most of which likely happened on the scientific and technological frontier rather than on infrastructure development. While the latter was captured by satellite, the former was certainly not. Most countries in Africa, despite their fast growth, started from low levels of income and inade-

---

<sup>1</sup>See, for example, [Elvidge and others \(1997\)](#), [Ghosh and others \(2010\)](#), [Henderson, Storeygard, and Weil \(2012\)](#), [Pinkovskiy and Sala-i Martin \(2016\)](#), among others.

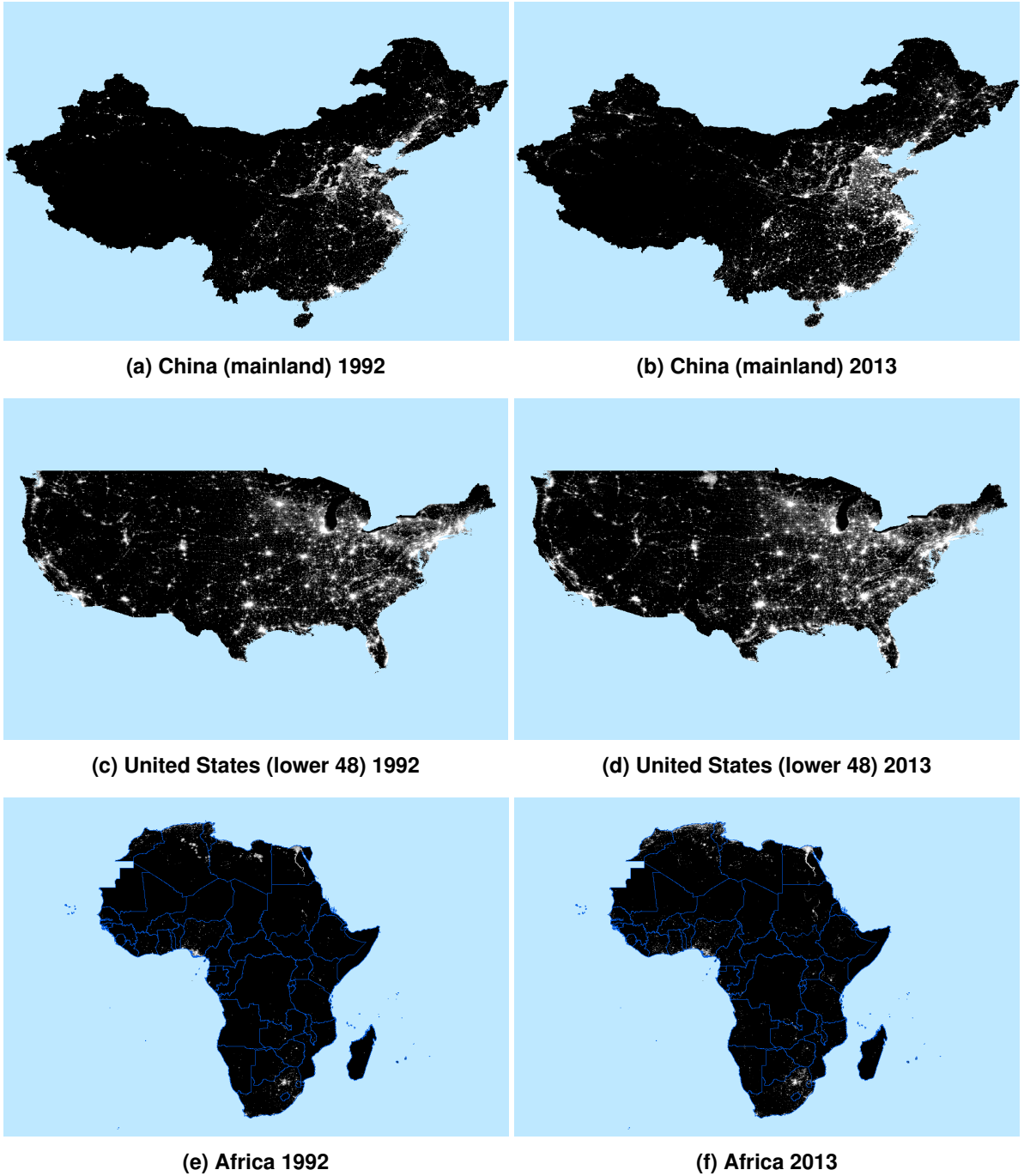
<sup>2</sup>The measurement errors of nighttime lights are independent of those of economic variables.

quate access to electricity. As a result, they were still mostly dark in 2013 and the information contained in nighttime lights may be insufficient for accurately assessing economic growth. Figure 1 highlights that the relationship between nighttime lights and real GDP may be nonlinear, that the relative accuracy of nighttime lights to real GDP may change, and that the extent to which nighttime lights are useful as proxy for real economic activity may differ over time and across countries.

In this paper, we address the aforementioned issues by uncovering the distribution of measurement errors in both official GDP measures and nighttime lights as well as their functional relationship simultaneously. We provide a statistical framework based on nonclassical and nonlinear measurement error models, in which the error in official GDP per capita may depend on the country's statistical capacity and the relationship between nighttime lights and true GDP per capita can be nonlinear and vary with geographic location. Using variation across geographic locations and different levels of statistical capacity, we establish identification of the distribution function of nighttime lights conditional on real GDP under fairly weak and reasonable statistical assumptions based on recently developed results for measurement error models.

Given our nonparametric identification results, estimates of nighttime lights' elasticity with respect to real GDP at different levels of income are naturally obtained. With the estimated distributions of measurement errors, we assess the relative uncertainty of nighttime lights and real GDP and the extent to which nighttime lights can be useful to improve real GDP measures. We then construct new real GDP measures by optimally combining official measures, information in nighttime lights, statistical capacity and geographic location. We focus on two new measures. One is an optimal linear combination of official data and prediction by nighttime lights, for which we provide an estimate of the optimal weight on nighttime lights for each observation of real GDP figure. We show that this optimal linear measure performs very well across countries. The other measure is the semiparametric conditional mean that is based on the full conditional distribution of true GDP given all the observables.

To our best knowledge, this is the first paper to estimate the distribution of measurement errors in official GDP and nighttime lights directly from data. The error distributions are crucial for both understanding the uncertainty in official GDP measures and constructing more precise measures. Intuitively, we can use nighttime lights to infer the accuracy of the official GDP. This is because nighttime lights reflect real economic activities, and therefore, are correlated with the true GDP. In the meantime, nighttime lights are independent of the measure-

**Figure 1. Examples of Satellite Images of Nighttime Lights**

ment error in the official GDP. After we identify how much nighttime lights a given amount of GDP may produce, i.e., the production function of nighttime lights, we may infer the distribution of the true GDP from the observed joint distribution of nighttime lights and the official GDP. Comparing the distribution of the true GDP with the observed distribution of the

official GDP, we may pin down the signal-to-noise ratio of the official GDP. Therefore, we can provide better measures of the true GDP using additional information, such as nighttime lights.

We conduct our analysis using two nighttime lights datasets. The first one spans from 1992 to 2013 and the second from 2013 to 2017. Recorded by different satellite systems, the two datasets are not only useful in extending the analysis of nighttime lights and real GDP to more recent date, but they also serve a purpose of validating the robustness of our identification and estimation strategy.

There are four main findings in this paper. First, the relationship between nighttime lights and real GDP is nonlinear. In our baseline specification, we estimate that the elasticity of nighttime lights with respect to real GDP per capita is around 2.5 for low income countries and close to 0 for high income countries. The elasticity steadily decreases as real GDP per capita increases, reflecting different developing mode at different stages of economic development. Intuitively, countries at early stages of development tend to build more infrastructure that generates lights at night, such as buildings and roads; countries at more advanced stages tend to focus more on technological innovation, which is less associated with lights at night.

Second, we find, perhaps not surprisingly, that measurement errors in real GDP per capita (and hence real GDP) are bigger for countries whose income is lower. There is a sharp distinction between high income countries and the rest of the countries. While high income countries have measurement errors of real GDP concentrated at zero, low and middle income countries have fat tails in the distribution of measurement errors. In other words, there is greater uncertainty in the latter countries' GDP measures and the measurement error can be substantial at times. The distinction between those with high and low statistical capacity among low and middle income countries, however, is blurred.

Third, nighttime lights are most useful for assessing and augmenting measures of real GDP in low and middle income countries. We find that the optimal weight of our new measure of real GDP on light-predicted GDP reaches 70% for middle income countries, but it declines for countries at either end of the income spectrum. For countries with extremely low levels of real GDP per capita, it is rather dark at night and as such the uncertainty in light-predicted GDP can be quite high. In contrast, for countries with high levels of real GDP per capita, nighttime lights are bright enough to reach the saturation level of satellite sensors and hence may not adequately reflect variations in economic activities. More fundamentally, limited access to electricity for low income countries and post-industrialization of high income countries are likely to disassociate their economic development from nighttime lights.



Finally, comparing our new measures with official measures of real GDP, we find that countries disrupted by conflicts and political instability often underestimated the deterioration of the economy during downturns and its recovery afterwards. It is likely that periods of economic disruption made it difficult to track the economy accurately and the emergence of informal economy in subsequent restoration did not enter national accounts. For example, conflicts might simply disrupt tax registration of firms that would otherwise have been recorded.

It should be emphasized that our new measures complement official GDP measures in gauging an economy, rather than substitute them. The construction of our new measures depends on the very existence of official GDP measures. Nevertheless, incorporating information from nighttime lights may provide useful insights into the size of an economy where official measures are uncertain.

The rest of the paper is organized as follows. Section II briefly reviews related literature. Section III and Section IV describe our statistical framework and data, respectively. We present our results in Section V. Section VI concludes. Robustness checks, including simulations and different estimates, are discussed in Appendix B.

## II. RELATED LITERATURE

This paper is closely related to several strands of literature.

First, we contribute to the growing literature on understanding economic growth through the lens of satellite-recorded nighttime lights. Since the seminal work of [Henderson, Storeygard, and Weil \(2012\)](#), nighttime lights have been increasingly used as a proxy for economic activity. For instance, [Pinkovskiy and Sala-i Martin \(2016\)](#) assess the relative quality of GDP per capita and survey means by comparing them to nighttime lights; [Storeygard \(2016\)](#) investigates the role of transport costs on the economic activity of cities proxied by nighttime lights; [Alesina, Michalopoulos, and Papaioannou \(2016\)](#) use nighttime lights to study ethnic inequality; [Henderson and others \(2018\)](#) studies the spatial distribution of economic activity proxied by nighttime lights. While most of the literature use nighttime lights directly as an alternative measure of real economic activity, we show that the relationship between nighttime lights and real GDP may differ for countries at different stages of development. Nighttime lights have also been used for forecast of GDP levels or growth rates. One of the approaches in such forecast is to use the elasticity between nighttime lights and real GDP ([WORLD BANK \(2017\)](#)). Our estimates of the elasticity, which varies across countries and over time, could be useful

for forecasting economic growth of low and middle income countries given the timeliness of recent nighttime lights data.

From a statistical perspective, [Henderson, Storeygard, and Weil \(2012\)](#) and [Pinkovskiy and Sala-i Martin \(2016\)](#) construct new measures of real GDP growth and real GDP per capita, respectively, through the combination of nighttime lights and official or survey-based GDP measures. While they obtain constant weights on nighttime lights, we take a step forward and show that the information content on real GDP from nighttime lights differ for each observation. For each country at each point in time, our optimal linear measure uses a different optimal weight on light-predicted GDP, and this can only be achieved when we uncover the entire distribution of measurement errors in both nighttime lights and official measures of real GDP.

Second, this paper is related to the measurement error literature on identification and estimation of measurement error models. Our statistical framework is based on recently developed results for nonclassical measurement error models. Since [Hu and Schennach \(2008\)](#), we can generally identify and estimate nonlinear models with nonclassical measurement errors in a continuous variable. When there are only two continuous measurements for a continuous latent variable as in the current paper, nonparametric identification requires additional data information or extra restrictions. [Carroll, Chen, and Hu \(2010\)](#) use a secondary survey sample to achieve nonparametric identification, which can be interpreted as identification with two continuous measurements and two discrete instruments. [Schennach and Hu \(2013\)](#) impose additivity and independence to show identification is feasible with two continuous measurements only.<sup>3</sup> Our method relies on the latter two papers with official GDP and nighttime lights as two continuous measurements and statistical capacity and geographic location being two discrete instruments. For high income countries, our identification of error distributions relies on additivity and independence assumptions.

Third, we contribute to the literature on improving the measurement of the economy from a measurement-error perspective. [Aruoba and others \(2016\)](#) improves historical United States' GDP growth at relatively high frequency and find the persistence of aggregate output dynamics to be stronger than previously thought. [Feng and Hu \(2013\)](#) show that the official US unemployment rate substantially underestimates the true level of unemployment. This paper aims to improve annual real GDP estimates in a measurement error model setting for low and middle income countries.

---

<sup>3</sup>See [Hu \(2017\)](#) for a short survey of the recent developments in this literature.

Finally, we make a contribution to the burgeoning literature on bringing satellite data to economic analysis. [Donaldson and Storeygard \(2016\)](#) provide a comprehensive review of applications of satellite data in economics. While many applications focus on converting satellite images to physical quantities relevant for economics, such as nighttime lights, greenness, or temperature, we focus on examining the relationship between such quantities and economic variables of interest from an econometric perspective. Our method can be applied broadly to a wide range of remote sensing data, as they inevitably contain measurement errors and their relationship with economic variables of interest may not be simple and linear. One of the benefits of our method is that we uncover the relationship with little information, which is often unavailable – the lack of alternative data sources is the very reason that satellite data are being used. Moreover, our method allows for flexible relationship between the latent true value and the measurement error.

### III. STATISTICAL FRAMEWORK

In this section, we present a statistical framework to analyze the nighttime lights data in relation to GDP. Under fairly weak assumptions, the nonlinear relationship between nighttime lights and GDP and the distribution of measurement errors in each of them are identified. Elasticities of nighttime lights with respect to GDP per capita are naturally obtained. We then construct a more accurate and robust measure of GDP per capita by optimally combining national accounts data and prediction by nighttime lights.

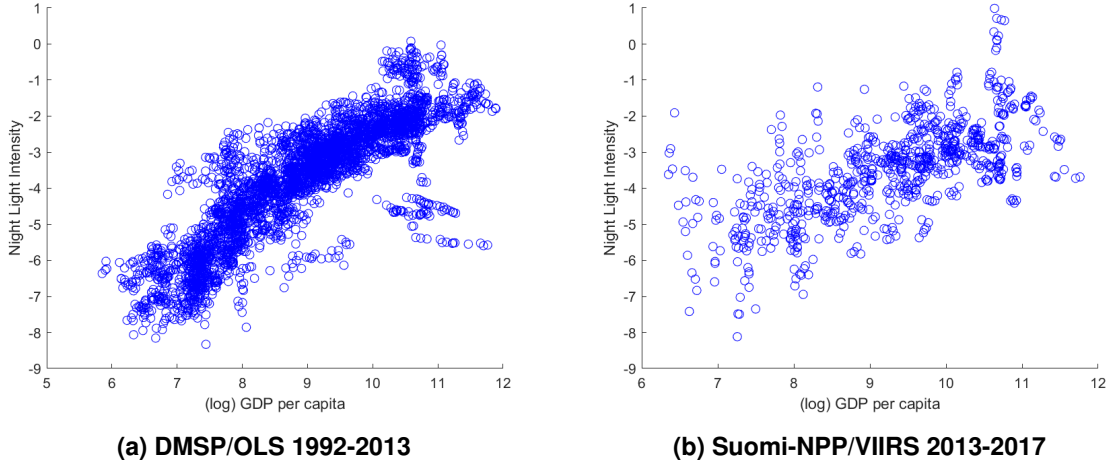
#### A. Baseline Setup

Let  $y_{i,t}^*$  denote the true real GDP per capita in logarithm for country  $i$  in year  $t$ . It is measured as  $y_{i,t}$  with error. Let  $s_{i,t}$  stand for the statistical capacity of country  $i$  at time  $t$ . Let  $z_{i,t}$  denote nighttime lights per capita in logarithm. It is related to the true real GDP per capita but also contains measurement error. Let  $l_i$  stand for the latitude of the country.<sup>4</sup>

In Appendices [A.1](#) and [A.2](#), we provide a set of sufficient conditions under which all the distributions containing the latent true GDP can be nonparametrically identified by the observed joint distribution of GDP and nighttime lights from countries with different statistical capacity and at different locations. For tractability, here we provide a simple specification that

---

<sup>4</sup>The satellites that recorded nighttime lights are polar orbiting. As such we choose the latitude of a country's centroid as the variable for its geographical location.

**Figure 2. Nighttime Lights vs. GDP (1992-2017)**

separates the functional relationship between nighttime lights and GDP per capita and the distribution of measurement errors. We assume that the reported GDP contains an additive measurement error, whose distribution may vary with different statistical capacity as follows:

$$y_{i,t} = y_{i,t}^* + \varepsilon_{i,t}^y(s_{i,t}), \quad (1)$$

In the meanwhile, the nighttime lights is related with the true latent GDP through an unknown production function  $m(\cdot)$  and an additive error term:

$$z_{i,t} = m(y_{i,t}^*) + \varepsilon_{i,t}^z(l_i) \quad (2)$$

The distribution of this error term also varies with the geographic locations.

The specification of the production function  $m(\cdot)$  is informed by the data. As suggested by Figure 2, a quadratic function form is sufficient. The distributions of the error terms  $\varepsilon_{i,t}^y$  and  $\varepsilon_{i,t}^z$  are allowed to be nonparametric. All the assumptions required for identification in Appendices A.1 and A.2 are satisfied under this simple specification, as shown in Appendix A.4.

## B. Sieve Maximum Likelihood Estimation

Given the general nonparametric identification, we provide a seminonparametric estimator as suggested in Carroll, Chen, and Hu (2010). We develop our estimator based on an i.i.d

sample, which can be extended to for time series data. We assume that there is a random sample  $\{z_i, y_i, s_i, l_i\}_{i=1}^n$ .

The specification of the production function  $m(\cdot)$  is informed by the data (Figure 2) and assumed to be a quadratic function. The error terms  $\varepsilon_{i,t}^y$  and  $\varepsilon_{i,t}^z$  are allowed to have a general density function. Therefore, in this empirical study, we adopt a parametric specification of function  $m(\cdot; \theta)$  and leave other elements nonparametrically specified in the simple specification in equations (1) and (2). Let the true value of the unknowns be  $\alpha_0 \equiv (\theta_0^T, f_{y^*|s,l}, f_{\varepsilon^y|s}, f_{\varepsilon^z|l})^T$ , where  $f_{A|B}$  denotes the distribution of  $A$  conditional on  $B$ . We then introduce a sieve MLE estimator  $\hat{\alpha}$  for  $\alpha_0$ , and establish the asymptotic normality of  $\hat{\theta}$ . These results can also be extended to the case where the function  $m$  is misspecified.

In the sieve MLE estimator, we use finite dimensional parametric representations to approximate the nonparametric densities in  $\alpha_0$ , where the dimension may increase with the sample size. Let  $\mathcal{A}$  be the parameter space. The log-joint likelihood for  $\alpha \equiv (\theta^T, f_1, f_2, f_3)^T \in \mathcal{A}$  is given by:

$$\sum_{i=1}^n \log f(z_i, y_i, s_i, l_i) = \sum_{i=1}^n \ell(D_i; \alpha),$$

in which  $D_i = (z_i, y_i, s_i, l_i)$  and

$$\begin{aligned} \ell(D_i; \alpha) &\equiv \ell(z_i, y_i, s_i, l_i; \theta, f_1, f_2, f_3) \\ &= \log \left\{ \int f_1(y^* | s_i, l_i) f_2(y_i - y^* | s_i) f_3(z_i - m(y^*; \theta) | l_i) dy^* \right\} + \log f(s_i, l_i). \end{aligned}$$

Let  $E[\cdot]$  denote the expectation with respect to the underlying true data generating process for  $D_i$ . Then

$$\alpha_0 = \arg \sup_{\alpha \in \mathcal{A}} E[\ell(D_i; \alpha)].$$

We then use a sequence of finite-dimensional sieve spaces  $\mathcal{A}_n$  to approximate the functional space  $\mathcal{A}$ . The seminonparametric sieve MLE  $\hat{\alpha}_n \in \mathcal{A}_n$  is defined as:

$$\hat{\alpha}_n = \arg \max_{\alpha \in \mathcal{A}_n} \sum_{i=1}^n \ell(D_i; \alpha).$$

Under assumptions presented in Appendix A.6, we show the consistency of estimator  $\hat{\alpha}_n$  for  $\alpha_0$  and the convergence rate of the nonparametric components. Furthermore, we show in Appendix A.7 that the sieve MLE  $\hat{\theta}_n$  is asymptotically normally distributed around the true value  $\theta_0$ . When the parametric model  $E[z|y^*, l] = m(y^*; \theta)$  is misspecified, the estimator  $\hat{\theta}_n$  is still

asymptotically normally distributed, but around a pseudo true value. In fact, the estimator  $\hat{\theta}_n$  is semiparametrically efficient for  $\theta_0$ .

### C. Constructing Better Measures of GDP

With the conditional distribution of nighttime lights and official real GDP per capita at hand, we construct new and more accurate measures of real GDP per capita that optimally combine the information in these two measures.

We first follow the practice in the literature and estimate the optimal linear combination of official data and prediction by nighttime lights. With the full conditional distributions, we are able to provide not only the optimal linear measure, but also a measure based on the semi-parametric conditional mean. The former is a convex combination of the official GDP and the GDP predicted by nighttime lights, and therefore, performs very robust across countries. The latter fully makes use of the observed information, but requires more data information so that it may not be robust in the area where observations are sparse.

#### 1. Optimal Linear Measure

Besides the official measure of the GDP, we may also use nighttime lights to predict the GDP, which provides a secondary measure of the GDP. Here we provide an optimal linear combination of the two measures, which is similar in spirit to [Henderson, Storeygard, and Weil \(2012\)](#) and [Pinkovskiy and Sala-i Martin \(2016\)](#). However, there are two important differences. First, the optimal linear combination relies on our estimates of the measurement error distributions in both national accounts GDP data and nighttime lights. This is in contrast to the literature where the optimal weight in the linear combination is obtained either by assumptions about the quality of official data or through the use of auxiliary data beyond nighttime lights. Second, our optimal weights on nighttime light-predicted GDP differ for each individual country in each year. For each observation of the nighttime light and GDP pair, the distributions of the measurement errors in official real GDP per capita and nighttime lights allow us to assess the relative uncertainty of nighttime light-predicted GDP to official GDP, which then permits the construction of an optimal weight that minimizes the conditional mean squared errors of the linear combination.

We proceed in two steps. In the first step, we construct a measure of GDP per capita based on nighttime lights,  $\hat{y}_{i,t}$ . Next, we construct the new measure of GDP per capita  $\hat{y}_{i,t}^*$  by estimat-

ing the optimal weight  $\lambda_{i,t}$  between nighttime lights prediction  $\hat{y}_{i,t}$  and national accounts data  $y_{i,t}$ .

*Real GDP per capita predicted by nighttime lights*

We regress official real GDP per capita on nighttime lights with country and year dummies,

$$y_{i,t} = \beta z_{i,t} + \delta_i D_i^c + \gamma D_t^y + \eta_{i,t}, \quad (3)$$

where  $D^c$  and  $D^y$  are country and year dummies, respectively. The nighttime light-predicted real GDP per capita is defined as

$$\hat{y}_{i,t} = \beta z_{i,t} + \delta_i D_i^c + \gamma D_t^y. \quad (4)$$

The predictive model in equation (4) assumes that controlling for country and year fixed effects, the level of real GDP per capita and nighttime lights have a stable and linear relationship.

Countries may differ in their habits of using lights at night and satellites' sensitivity to lights may change each year. However, after taking these idiosyncrasies into account, we assume that international experience can be relied on to predict real GDP per capita using nighttime lights. It is possible to include higher order terms of nighttime lights in the regression equation (3) and to include interaction terms between country dummies and nighttime lights. However, we prefer equation (3) for its simplicity and as shown later it already explains more than 98% of variation in official real GDP per capita. Including more terms leads to overfitting, making the nighttime light-predicted real GDP per capita  $\hat{y}_{i,t}$  very close to official measures  $y_{i,t}$ , in which case we likely lose the useful information in nighttime lights.

*Optimal linear combination*

We construct a new measure  $\hat{y}^*_{i,t}$  of GDP per capita based on a linear combination of nighttime light-predicted and official GDP per capita, i.e.,

$$\hat{y}^*_{i,t} = \lambda_{i,t} \hat{y}_{i,t} + (1 - \lambda_{i,t}) y_{i,t}, \quad (5)$$

where  $\hat{y}_{i,t}$  is nighttime light-predicted GDP per capita,  $y_{i,t}$  is official GDP per capita, and  $\lambda_{i,t}$  is the weight.

To obtain the optimal linear combination, we minimize the conditional mean squared error of our new measure, i.e.,

$$\lambda_{i,t} = \arg \min_{\lambda} E \left[ (\hat{y}_{i,t}^* - y_{i,t}^*)^2 | z_{i,t}, s_{i,t}, l_i \right], \quad (6)$$

where  $y_{i,t}^*$  is the true GDP. By equations (1) and (2), this conditional mean squared error can be decomposed into two parts:

$$\begin{aligned} & E \left[ (\hat{y}_{i,t}^* - y_{i,t}^*)^2 | z_{i,t}, s_{i,t}, l_i \right] \\ &= \lambda^2 E \left[ (\hat{y}_{i,t} - y_{i,t}^*)^2 | z_{i,t}, s_{i,t}, l_i \right] + (1 - \lambda)^2 E \left[ (\varepsilon^y)^2 | s_{i,t} \right]. \end{aligned} \quad (7)$$

The first term in equation (7) captures the uncertainty in nighttime light-predicted GDP per capita,<sup>5</sup> whereas the second term captures the uncertainty in official GDP per capita.

The optimal weight then depends on the relative uncertainty in these two measures of GDP per capita and equals,

$$\lambda_{i,t} = \frac{E \left[ (\varepsilon^y)^2 | s_{i,t} \right]}{E \left[ (\hat{y}_{i,t} - y_{i,t}^*)^2 | z_{i,t}, s_{i,t}, l_i \right] + E \left[ (\varepsilon^y)^2 | s_{i,t} \right]}. \quad (8)$$

Nighttime lights help provide the optimal weight  $\lambda_{i,t}$  because they reflect real economic activities, and therefore, are correlated with the true GDP. In the meantime, nighttime lights are independent of the measurement error in the official GDP. After we identify the distribution of nighttime lights conditional on the true GDP and the distribution of the true GDP, we may infer the uncertainty in nighttime light-predicted GDP and the uncertainty in the official GDP. We may then pin down the optimal weight in the linear combination of the nighttime light-predicted GDP and the official GDP. Notice that  $\lambda_{i,t}$  is always in  $[0, 1]$ , which makes the optimal linear measure very robust across countries and years.

## 2. Semiparametric Conditional Mean

Given that we have identified the distribution of the true GDP joint with all the observables in the semiparametric specification, a natural measure of the true GDP is the conditional mean

<sup>5</sup>Note that

$$E \left[ (\hat{y}_{i,t} - y_{i,t}^*)^2 | z_{i,t}, s_{i,t}, l_i \right] = \frac{\int (\hat{y}(z_{i,t}) - y^*)^2 f_{\varepsilon^z}(z_{i,t} - m(y^*)) f(y^* | s_{i,t}, l_i) dy^*}{\int f_{\varepsilon^z}(z_{i,t} - m(y^*)) f(y^* | s_{i,t}, l_i) dy^*}$$

is a function of  $z_{i,t}$ . And we treat the country and year dummies as exogenous variables given  $(z_{i,t}, s_{i,t}, l_i)$ .



of the distribution, i.e.,  $E[y_{i,t}^* | y_{i,t}, z_{i,t}, s_{i,t}, l_i]$ . Notice that the conditional mean may also be considered as a minimizer of the mean squared error conditional on all the observables, particularly including  $y_{i,t}$ , which makes it different from the previous optimal linear measure. Under our nonparametric identification and semiparametric specification, it follows that

$$E[y_{i,t}^* | y_{i,t}, z_{i,t}, s_{i,t}, l_i] = \frac{\int y^* f(y^* | s_{i,t}, l_i) f(y_{i,t} - y^* | s_{i,t}) f(z_{i,t} - m(y^*; \theta) | l_i) dy^*}{\int f(y^* | s_{i,t}, l_i) f(y_{i,t} - y^* | s_{i,t}) f(z_{i,t} - m(y^*; \theta) | l_i) dy^*}.$$

Compared to the optimal linear measure, while the optimal measure  $E[y_{i,t}^* | y_{i,t}, z_{i,t}, s_{i,t}, l_i]$  has the clear advantage of making full use of the information in the conditional distributions. However, its nonparametric feature makes it less robust than the previous optimal linear measure because the semiparametric measure requires a large sample size to perform well. Especially in the sparse area of the empirical distribution of  $(y_{i,t}, z_{i,t}, s_{i,t}, l_i)$ , the semiparametric conditional mean can be volatile, while the optimal linear measure remains robust. In the area where the density  $f(y_{i,t}^* | y_{i,t}, z_{i,t}, s_{i,t}, l_i)$  takes a relatively larger value, the conditional mean is actually stable and also close to the optimal linear measure. For this reason, we make the optimal linear measure our choice of new measure for the true GDP.

## IV. DATA

### A. Nighttime Lights

The U.S. Air Force Defense Meteorological Satellite Program (DMSP) Operational Linescan System (OLS) has been collecting global low light imaging data since the 1970s. The National Oceanic and Atmospheric Administration (NOAA) processes the data and hosts a digital archive from 1992 to 2013.<sup>6</sup> DMSP satellites overpass at local time in the 7pm to 9pm range,<sup>7</sup> and nighttime lights are a class of derived products of the low light imaging data in spectral bands where electric lights emissions are observed. NOAA provides cloud-free composites of nighttime lights based on a set of quality criteria that remove observations affected by sunlight, moonlight, glare, aurora, and the edges of the DMSP/OLS swaths.<sup>8</sup> For some years, there were two satellites collecting data and two composites were produced. In those cases, we use the average of the two composites. Figure 3 is an example of the DMSP/OLS

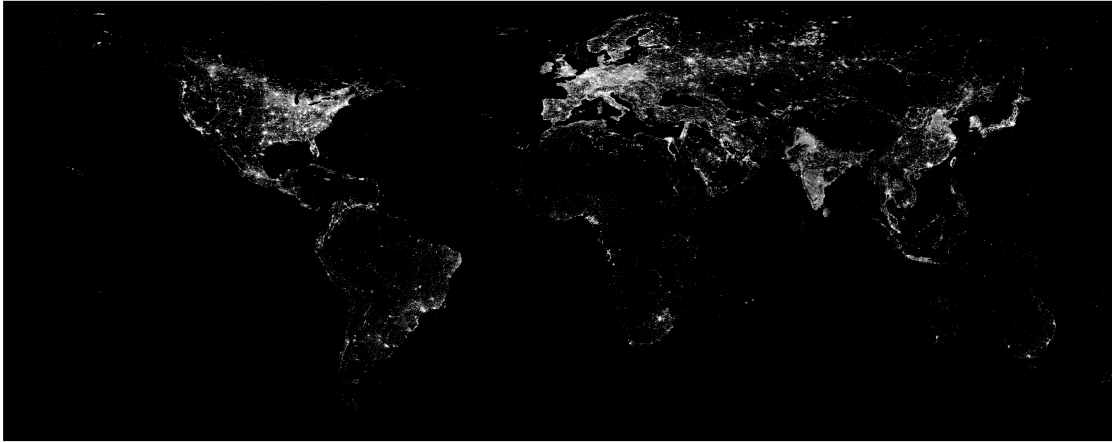
<sup>6</sup> The lights data can be downloaded here: <http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html>.

<sup>7</sup> See Elvidge and others (2009) for an overview.

<sup>8</sup> A detailed description of the selection criteria can be found here: [https://www.ngdc.noaa.gov/eog/gcv4\\_readme.txt](https://www.ngdc.noaa.gov/eog/gcv4_readme.txt).

nighttime lights image in 2010. Each pixel of the DMSP/OLS nighttime lights images is a 30 arc-second grid (a bit less than 1 square kilometer). It is associated with a numerical value of radiance from 0 to 63 that is increasing with brightness.

**Figure 3. Map of Nighttime Lights in 2010**



Since 2013, nighttime lights data are produced monthly by the Visible Infrared Imaging Radiometer Suite (VIIRS) onboard a different satellite (Suomi-NPP).<sup>9</sup> VIIRS nighttime lights have a number of advantages compared to DMSP/OLS, including greater radiometric accuracy, finer geographical resolution, etc., though the satellite's overpass time is after midnight, when outdoor lights are arguably less related to residential economic activity than DMSP/OLS nighttime lights. NOAA have yet to process annual nighttime lights data. We therefore take the 12-month average of monthly data as annual data and use them for analysis.

### **B. Administrative Maps and Country Location**

The Database of Global Administrative Areas (known as GADM) provides administrative shape files for all countries in the world. For each country in each year, we clip the nighttime

---

<sup>9</sup>VIIRS nighttime lights data can be downloaded here: [https://www.ngdc.noaa.gov/eog/viirs/download\\_dnb\\_composites\\_iframe.html](https://www.ngdc.noaa.gov/eog/viirs/download_dnb_composites_iframe.html).

lights map down to their country borders and sum up the numerical value of radiances of all pixels within their borders. As such we obtain a country-year nighttime lights panel.

To obtain a country's latitude, we calculate the centroid of its largest contiguous block. For instance, the United States has a few separate bodies of landmass, such as Alaska, Hawaii, and the lower 48 states. We use the centroid of the lower 48 states as the location of the United States. Appendix C.2 provides an illustration.

For estimation purpose, we discretize country locations into binary values based on whether their centroid's latitude is between the Northern and Southern Tropic or outside. Countries between the tropics generally receive more sunshine than those outside, which might affect the background light in the nighttime lights maps.<sup>10</sup> Though NOAA has gone to great lengths to remove background noises, there might still remain measurement errors related to country locations. Our classification also takes into account the number of countries in each geographical subarea. For example, an alternative classification is to group countries by Northern and Southern Hemisphere, but the Southern Hemisphere contains much less countries.

### C. GDP, Population, and Statistical Capacity

We obtain GDP per capita (PPP, constant 2011 international dollars) and population data from the World Bank. To calculate night light intensity, which we define as (log) total nighttime lights per capita, we divide the radiance sum of lights by population within each country in any given year and use the logarithms. The World Bank also provides statistical capacity ratings for low and middle income countries. Though the ratings change over time, the change is small for most countries. For this reason we group countries into three categories: high income countries, those below the median of statistical capacity rating among the low and middle income countries and those above.

In total we have an unbalanced panel of 182 countries and 3870 observations based on DMSP/OLS nighttime lights data. The majority of countries have data spanning from 1992 to 2013. A similar panel is constructed for Suomi/NPP VIIRS nighttime lights data, where there are 184 countries and 920 observations. The discrepancy between the number of countries in the two datasets arises from the availability of statistical capacity ratings.<sup>11</sup> Table 1 and 2 presents the

<sup>10</sup>In addition, the sun sets later in the summer for countries closer to the arctic circle. For the VIIRS data set, the nighttime light readings are zero for Nordic countries during the summer months. While we use the annual average where such issues are less concerning, differentiating measurement errors for countries at different locations can nevertheless be helpful.

<sup>11</sup>The two countries are Barbados and the Czech Republic.

summary statistics based on nighttime lights data from DMSP/OLS and VIIRS, respectively. As can be seen from the table, richer countries tend to be brighter at night.

**Table 1. Summary Statistics (DMSP/OLS)**

Location	Statistical Capacity	Night lights per 1000 people	real GDP per capita	# of countries	# of obs
Between Tropics	Low	66	9490	12	228
Between Tropics	High	58	9680	36	786
Between Tropics	(High income)	186	40214	34	729
Outside Tropics	Low	17	5311	57	1192
Outside Tropics	High	21	6614	37	812
Outside Tropics	(High income)	78	62074	6	123
Total	-	63	15097	182	3870

**Table 2. Summary Statistics (VIIRS)**

Location	Statistical Capacity	Night lights per 1000 people	real GDP per capita	# of countries	# of obs
Between Tropics	Low	101	9250	10	50
Between Tropics	High	63	12074	30	150
Between Tropics	(High income)	172	40720	45	225
Outside Tropics	Low	28	5169	51	255
Outside Tropics	High	24	8881	35	175
Outside Tropics	(High income)	81	46064	13	65
Total	-	76	18807	184	920

As mentioned in Section III.B, the nonparametric densities in the sieve MLE estimator are approximated by finite dimensional parametric representations, where the dimension depends on the sample size. We find that Hermite orthogonal polynomials work well as basis functions with just a few sieve terms.<sup>12</sup> Given the sample size of our data sets, we conduct simulation studies in Appendix B to choose the smoothing parameters in our sieve MLE estimator. With a sample size similar to the DMSP/OLS sample, our simulation studies show that the estimates are stable with the number of sieve terms used for each density function being around 6. As such we choose 6 for the DMSP/OLS sample. The VIIRS sample has much fewer observations and we reduce the number of sieve terms to 4.

<sup>12</sup>Compared to Hermite polynomials, the drawback for using Legendre polynomials is that they have bounded support and Fourier series require many more sieve terms to approximate density functions well.

## V. RESULTS

### A. Nighttime Lights' Elasticity of Real GDP per capita

Table 3 displays the estimated quadratic nighttime light production function for DMSP/OLS data. The estimates are quite precise with each parameter statistically significant at the 0.05 level. Notably, the parameter estimate on the quadratic term is non-zero, indicating that nighttime lights' elasticity of real GDP per capita is not constant. Table 4 presents the results for VIIRS data. The estimates are much less precise due to the limited number of observations and a linear production function seems adequate. The two production functions are expected to be different because the overpass time of the two satellite systems differs.

**Table 3. Estimated Light Production Function (DMSP/OLS)**

$$m(y^*) = \theta_0 + \theta_1 y^* + \theta_2 (y^*)^2$$

Parameter	$\theta_0$	$\theta_1$	$\theta_2$
Point Estimate	0.398	1.234	-0.244
Standard Error	(0.176)	(0.139)	(0.049)

Data are re-centered at zero. Standard errors are based on 400 sample bootstraps.

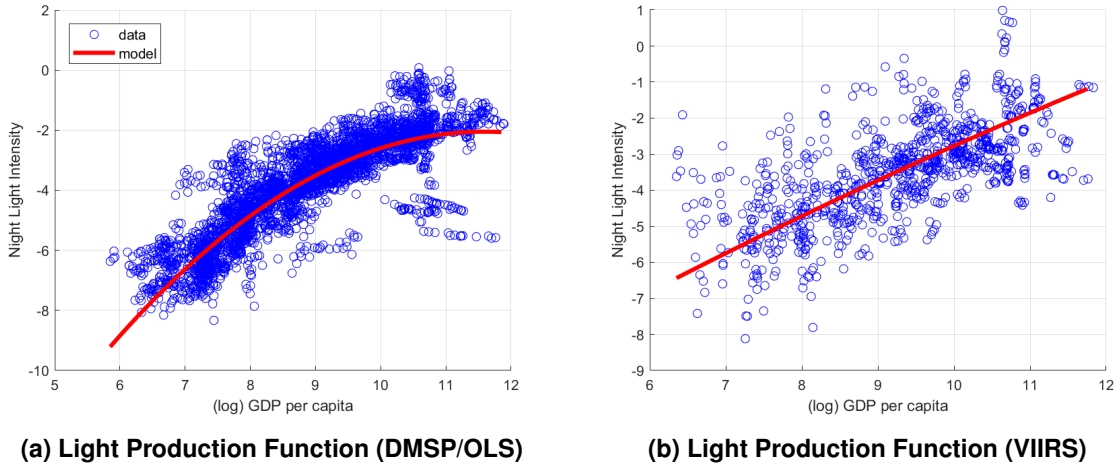
**Table 4. Estimated Light Production Function (VIIRS)**

$$m(y^*) = \theta_0 + \theta_1 y^* + \theta_2 (y^*)^2$$

Parameter	$\theta_0$	$\theta_1$	$\theta_2$
Point Estimate	0.225	0.964	-0.019
Standard Error	(0.170)	(0.248)	(0.157)

Data are re-centered at zero. Standard errors are based on 400 sample bootstraps.

Figure 4 presents the estimation results graphically. Graph (a) superimposes the estimated light production function on the raw DMSP/OLS data. It can be seen that the quadratic func-

**Figure 4. Estimation Results with Quadratic Light Production Function**

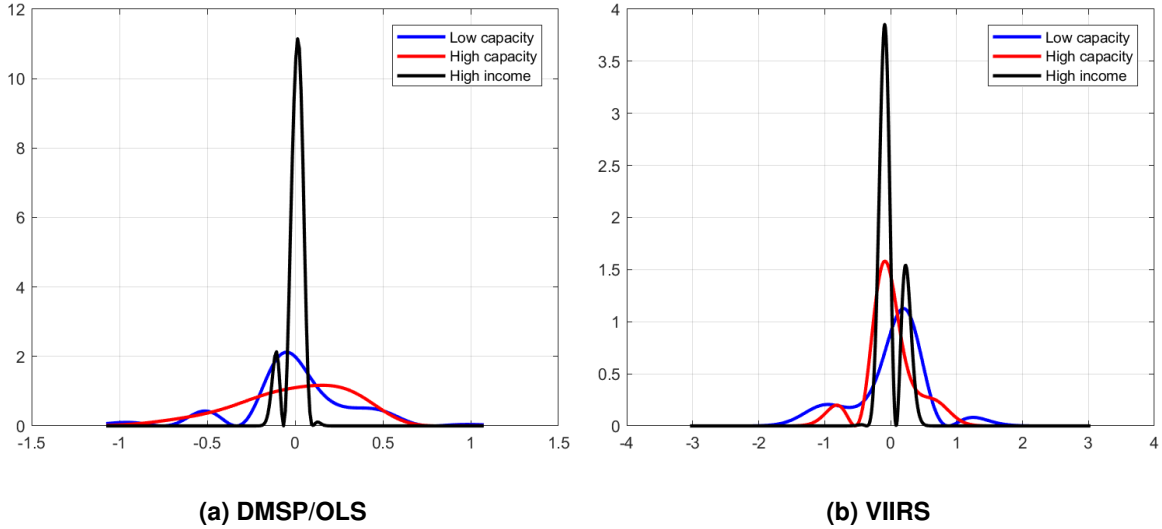
tion form broadly captures the relationship in the data.<sup>13</sup> Nighttime lights' elasticity with respect to real GDP per capita, which is the slope of the production function, steadily decreases as real GDP per capita increases, reflecting that a country moves away from infrastructure development toward technical progress as it develops. Low income countries have an estimated elasticity about of 2.5 whereas high income countries have an elasticity close to 0.

Graph (b) display the results based on VIIRS data. The estimated elasticity is close to 1 for all countries. Interestingly, the elasticities based on VIIRS data are on average smaller than those based on DMSP/OLS data, which is consistent with the overpass time of the two satellite systems. DMSP/OLS satellites' local overpass time was between 7pm and 9pm, whereas that of VIIRS is after midnight. Intuitively, nighttime lights between 7-9 pm have more to do with real economic activities than those after midnight.

## B. Uncertainty in Real GDP per capita

Figure 5 compares the probability density function of measurement errors of real GDP per capita for high income countries and the rest of the countries with relatively high and low statistical capacity. There is a sharp distinction between high income countries and the rest,

<sup>13</sup>One caveat here is that when plotting the raw data, we use official GDP per capita, which contain measurement error, whereas when plotting the light production function, the horizontal grids should be interpreted as true GDP per capita.

**Figure 5. Distribution of Measurement Errors of Real GDP per capita**

whereas the distinction between those with high and low statistical capacity among low and middle income countries is not clear-cut.

The measurement errors of high income countries' real GDP per capita are concentrated at zero, indicating relatively high precision in official figures. In contrast, low and middle income countries' distribution of measurement errors have fat tails, which suggests that measurement errors are generally of bigger size. Among low and middle income countries, both data sets suggest that those with high statistical capacity tend to have similar measurement errors to those with low statistical capacity.

### C. New Measures of Real GDP per capita

We focus on the optimal linear measure of real GDP per capita based on equation (5), where we need nighttime light-predicted real GDP per capita as well as the optimal weights to combine official figures and prediction by nighttime lights.

Table 5 presents results from regressions of official real GDP per capita on nighttime lights and a set of dummies based on DMSP/OLS nighttime lights data. Column (1) is our baseline specification in equation (3). With R square above 0.98, nighttime lights are highly correlated with real economic activity. For the purpose of comparison, we include a quadratic term in column (2). Though the coefficient on the quadratic term is statistically significant, there is

little improvement in R square. In column (3), we further include interaction terms of nighttime lights and countries dummies.

We use prediction by nighttime lights based on column (1). As mentioned before, we want to retain information in nighttime lights and predict real GDP per capita based on cross-country relationship between nighttime lights and real economic activity. While column (2) adds little predictive power by including higher order terms of nighttime lights, column (3) uses within-country variation to predict real GDP per capita, which partly defeats the purpose of relying on cross-country comparisons to estimate real economic activity. In cases where official real GDP per capita is systematically mis-measured, predictions by column (3) will be biased accordingly.

**Table 5. Official Real GDP per capita and Nighttime Lights 1992-2013**

Dependent Variable:	(log) Real GDP per capita		
	(1)	(2)	(3)
(log) Nighttime Lights per capita	0.218*** (0.013)	0.265*** (0.031)	0.373*** (0.110)
(log) Nighttime Lights per capita squared		0.005* (0.003)	
country fixed effect	X	X	X
year fixed effect	X	X	X
(log) Nighttime Lights per capita × country dummies			X
adj. $R^2$	0.981	0.981	0.986
N	3,870	3,870	3,870

Standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

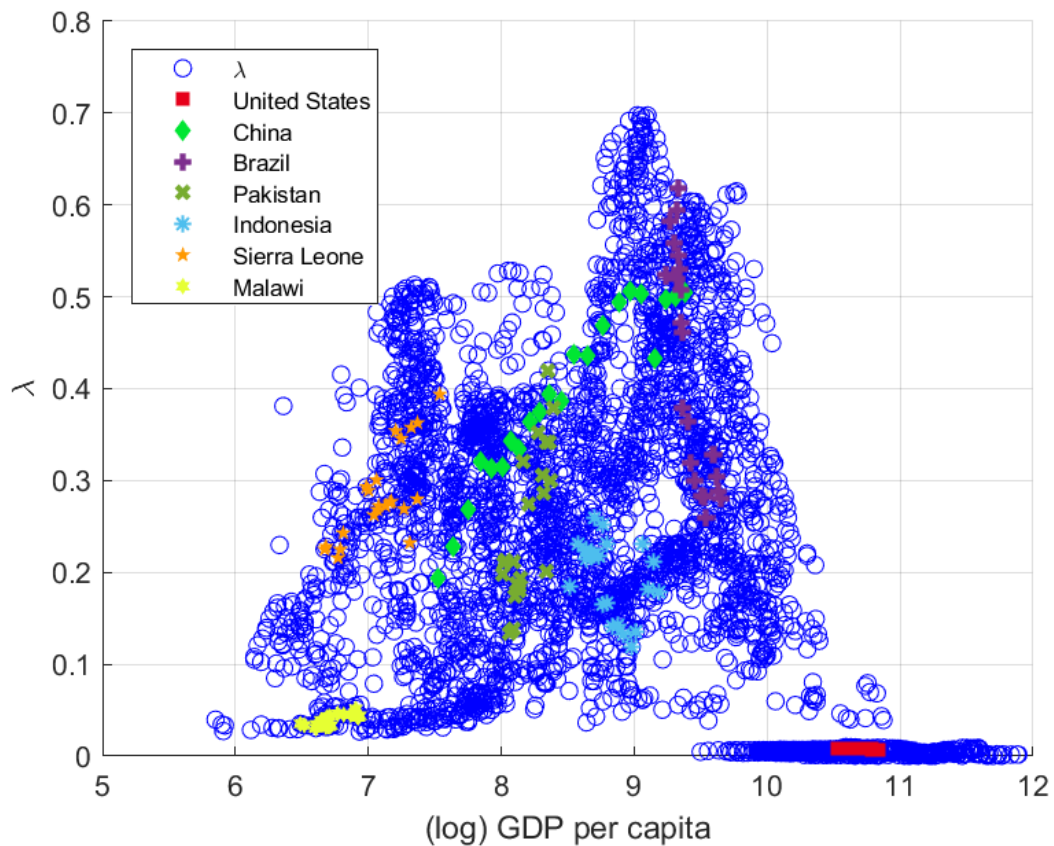
The optimal weights are calculated according to equation (8). Figure 6 contrast the optimal weights against real GDP per capita based on DMSP/OLS data.<sup>14</sup> There is broadly a bell-shaped pattern. Countries with very low or high income tend to have small weights on prediction by nighttime lights, whereas countries in the middle of the income spectrum have comparatively high weights.

The optimal weight makes use of the relative accuracy of nighttime light-predicted GDP and official GDP figures. Intuitively, for countries with extremely low levels of real GDP per

<sup>14</sup>Optimal weights based on VIIRS data display a similar pattern.



**Figure 6. Optimal Weights and Real GDP per capita: DMSP/OLS 1992-2013**



capita, it is rather dark at night and as such the uncertainty in nighttime light-predicted GDP can be quite high. For example, night light intensity for Malawi is almost zero as a result of paltry access to electricity in the country. Our optimal weights on nighttime light-predicted GDP are below 0.05 for Malawi. In contrast, for countries with high levels of real GDP per capita, nighttime lights are bright enough to reach the saturation level of satellite sensors and hence may not adequately reflect variations in economic activity. More fundamentally, post-industrialization of high income countries is likely to disassociate their economic development from nighttime lights. Meanwhile, high income countries tend to have the most accurate national accounts data. Thus it is not surprising that we find almost zero weights for high income countries such as the United States.

Nighttime lights can play a big role in improving real GDP per capita measures for the majority of middle and low income countries. Figure 6 shows that for fragile states like Sierra Leone, emerging markets such as China, Brazil, Indonesia, and Pakistan, our optimal weights on nighttime light-predicted GDP range from 0.2 to 0.6.

It is also useful to examine the absolute accuracy of the two measures. To this end, Table 6 compares the mean squared error of regression (3) of countries with different World Bank income status. The mean squared error of light-predicted GDP per capita for low income countries is almost twice as it is for high income countries, and this is not driven by the number of observations. In other words, there is greater discrepancy between light-predicted GDP per capita and official figures for low income countries.

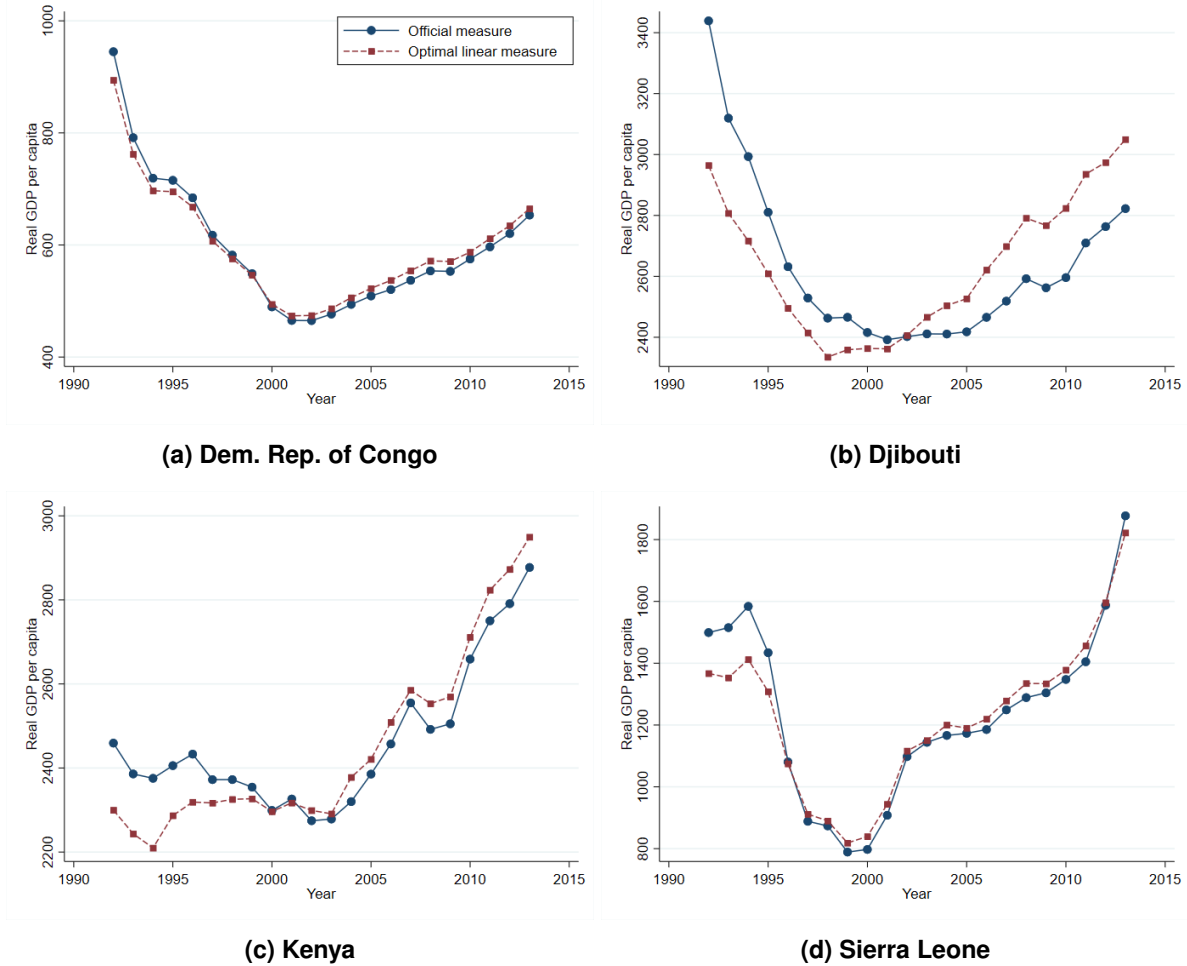
**Table 6. Optimal Weights and Mean Squared Error by Income Status**

World Bank Country Classification	Low Income	Lower Middle Income	Upper Middle Income	High Income
$\lambda$	0.24	0.34	0.28	0.01
Mean squared error	0.02	0.02	0.07	0.01
# observations	1212	1121	673	852

#### D. Official vs. New Measures

Comparing official measures of real GDP per capita to our new measures for every country in our data sets, we find that countries disrupted by conflicts and political instability often underestimated the deterioration of the economy during downturns and its recovery afterwards. For example, Figure 7 contrasts the optimal linear measure against official measure for the Democratic Republic of the Congo, Djibouti, Kenya, and Sierra Leone, respectively. The Democratic Republic of the Congo was in a state of conflict until 2001 when UN peacekeepers arrived. When its economic situation deteriorated between 1992-2001, panel (a) shows that the new measures of GDP per capita were worse than official figures suggested; to the contrary, when the economic situation improved afterwards, the new measures suggested higher living standards. Similarly, panel (b) shows that during the 1990s when armed conflicts routed the economy in Djibouti, the new measure suggested worse situation than official figures; when the economy recovered, the new measures suggested higher GDP per capita. Kenya's economy was afflicted by political instability before 2002 and Sierra Leone by its civil war in the 1990s. Graphs (c) and (d) again display a similar pattern of overestimation of real GDP per capita in the economic downturn and underestimation in the upturn. It is likely that periods of economic disruption made it difficult to track the economy accurately and the emergence of informal economy in subsequent restoration did not enter national accounts.

**Figure 7. Real GDP per capita: Economic Disruption and Restoration**



Appendix B.4 presents the official measures of real GDP per capita and our optimal linear measures for a number of other countries.

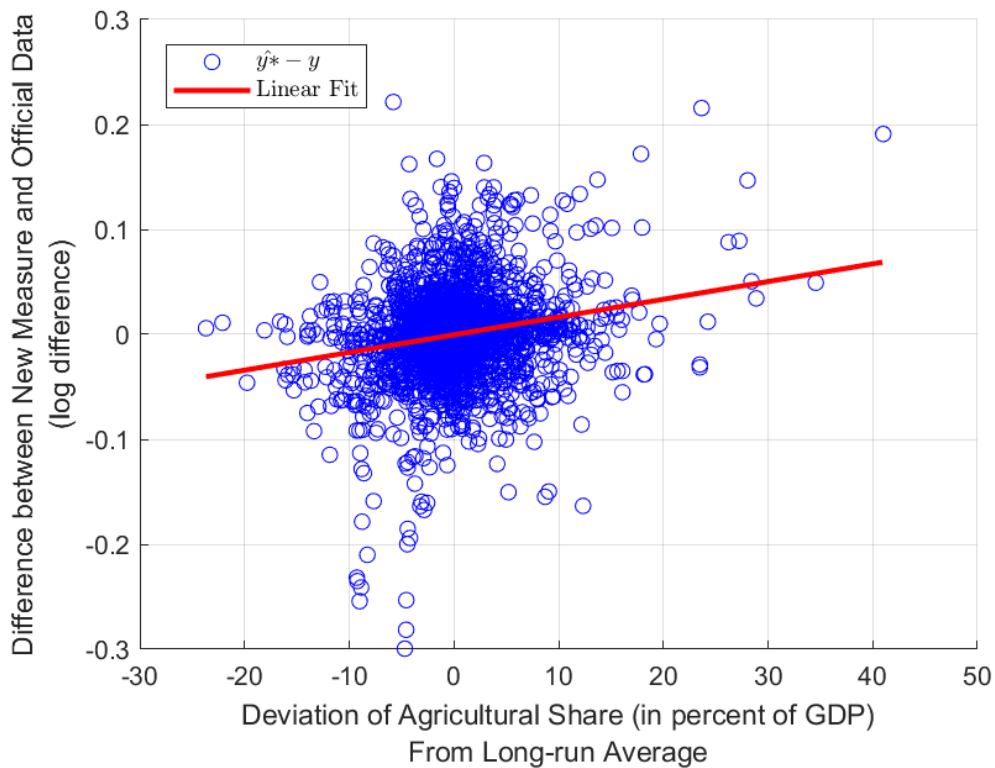
#### *Relation to Informal Economy*

There is a large literature documenting that the size of the informal economy is large and that much of it is related to agriculture (see, for example, [La Porta and Shleifer \(2014\)](#) and [Medina and Schneider \(2018\)](#)). A natural question that arises is the extent to which the discrepancy between the new measures and official data is driven by the existence of the informal economy.

The optimal linear measure, by construction through equations (4) and (5), implies that the discrepancy between the new and official measures should be zero on average over a long period of time. As such the discrepancy does not reflect the absolute size of the informal econ-

omy, but rather could be related to the changing size of the informal economy. To see this, we focus on low and middle income countries and use the agricultural share of the economy<sup>15</sup> as a proxy for the informal economy. Figure 8 contrasts the discrepancy between the new and official measures of real GDP per capita against the deviation of a country's agricultural share from its long-run average. Each blue dot represents a country-year observation. There is clearly a positive relationship. In years where the agricultural share of a country is above its own average, the discrepancy between the two GDP measures tends to increase. However, the correlation is only 0.20, suggesting that there are more dimensions than the informal economy that drive the discrepancy. We leave the precise mechanisms that drive such discrepancy for future research.

**Figure 8. New Measures and the Informal Economy: 1992-2013**



In this paper, we show that nighttime lights can be a useful source of information to improve official real GDP data. To begin with, they can be used to detect the uncertainty in official data and potential mismeasurement of real GDP. Systematic differences between the new and official measures may warrant further investigation as to what contributes to such differences.

<sup>15</sup>Data are from the World Bank.

While nighttime lights can be computationally intensive to process, the method developed in this paper is not limited to nighttime lights. In fact, measures of real GDP that are conditionally independent of official data can be used in a similar fashion.

Several caveats are in order. First, undoubtedly, countries differ in their social norms and habits of using lights at night. To the extent that these country-specific factors are time-invariant, such difference is taken into account in our optimal linear measure through country fixed effects. However, to the extent that countries have undergone structural transformation in a way that is dissimilar to other countries, our new measures may not adequately capture such changes. Second, our model assumes that measurement errors in nighttime lights depend on the latitude of a country's centroid. For countries with a large geographical coverage, this assumption may be strong. As such we view our new measures as more useful and relevant for countries with an average size. Third, while nighttime lights has the benefits of being independent of GDP measurement errors, its relationship with GDP may not be always the same, especially given increasing global awareness of energy conservation and climate change. Our new measures provide new perspectives on understanding economic growth and augmenting official statistics in gauging the economy. Understanding the relationship between nighttime lights and the structure of the economy and utilizing alternative measures to further improve measurements of GDP would be next steps of our future research.

## VI. CONCLUSION

In this paper, we first provide a statistical framework to describe the relationship among nighttime lights, official GDP, and true GDP. We make use of the variation of observed nighttime lights and official GDP across different statistical capacity and geographic location, and provide sufficient conditions under which the joint distribution of observables and the latent true GDP is uniquely determined by the distribution of observables. The paper proposes a semi-nonparametric Sieve MLE and constructs new measures of real GDP per capita and real GDP growth based on estimated distributions. We find that official real GDP per capita measures are less precise for low income countries and nighttime lights can play a bigger role in improving such measures. Comparing our new measures with official measures of real GDP, we find that countries disrupted by conflicts and political instability often underestimated the deterioration of the economy during downturns and its recovery afterwards. We expect our statistical framework and methodology will have a broad impact on measuring GDP using additional information.

## REFERENCES

- Alesina, Alberto, Stelios Michalopoulos, and Elias Papaioannou, 2016, "Ethnic Inequality," *Journal of Political Economy*, Vol. 124, No. 2, pp. 428–488.
- Aruoba, S Borağan, Francis X Diebold, Jeremy Nalewaik, Frank Schorfheide, and Dongho Song, 2016, "Improving GDP Measurement: A Measurement-Error Perspective," *Journal of Econometrics*, Vol. 191, No. 2, pp. 384–397.
- Carroll, Raymond J, Xiaohong Chen, and Yingyao Hu, 2010, "Identification and Estimation of Nonlinear Models Using Two Samples With Nonclassical Measurement Errors," *Journal of Nonparametric Statistics*, Vol. 22, No. 4, pp. 379–399.
- Donaldson, Dave, and Adam Storeygard, 2016, "The View from Above: Applications of Satellite Data in Economics," *Journal of Economic Perspectives*, Vol. 30, No. 4, pp. 171–98.
- Elvidge, Christopher D, Kimberly E Baugh, Eric A Kihn, Herbert W Kroehl, and Ethan R Davis, 1997, "Mapping City Lights With Nighttime Data from the DMSP Operational Linescan System," *Photogrammetric Engineering and Remote Sensing*, Vol. 63, No. 6, pp. 727–734.
- Elvidge, Christopher D, Edward H Erwin, Kimberly E Baugh, Daniel Ziskin, Benjamin T Tuttle, Tilottama Ghosh, and Paul C Sutton, 2009, "Overview of DMSP Nighttime Lights and Future Possibilities," in *Urban remote sensing event, 2009 joint*, pp. 1–5 (IEEE).
- Feng, Shuaizhang, and Yingyao Hu, 2013, "Misclassification Errors and the Underestimation of the US Unemployment Rate," *American Economic Review*, Vol. 103, No. 2, pp. 1054–70.
- Freyberger, Joachim, 2017, "On Completeness and Consistency in Nonparametric Instrumental Variable Models," *Econometrica*, Vol. 85, No. 5, pp. 1629–1644.
- Ghosh, Tilottama, Rebecca L Powell, Christopher D Elvidge, Kimberly E Baugh, Paul C Sutton, and Sharolyn Anderson, 2010, "Shedding Light on the Global Distribution of Economic Activity," *Open Geography Journal*, Vol. 3, pp. 147–160.
- Henderson, J Vernon, Tim Squires, Adam Storeygard, and David Weil, 2018, "The Global Distribution of Economic Activity: Nature, History, and the Role of Trade," *The Quarterly Journal of Economics*, Vol. 133, No. 1, pp. 357–406.
- Henderson, J Vernon, Adam Storeygard, and David N Weil, 2012, "Measuring Economic Growth from Outer Space," *American Economic Review*, Vol. 102, No. 2, pp. 994–1028.
- Hu, Yingyao, 2017, "The Econometrics of Unobservables: Applications of Measurement Error Models in Empirical Industrial Organization and Labor Economics," *Journal of Econometrics*, Vol. 200, No. 2, pp. 154–168.
- Hu, Yingyao, and Susanne M Schennach, 2008, "Instrumental Variable Treatment of Nonclassical Measurement Error Models," *Econometrica*, Vol. 76, No. 1, pp. 195–216.

- La Porta, Rafael, and Andrei Shleifer, 2014, “Informality and development,” *Journal of Economic Perspectives*, Vol. 28, No. 3, pp. 109–26.
- Medina, Leandro, and Friedrich Schneider, 2018, “Shadow Economies Around the World: What Did We Learn Over the Last 20 Years?” *IMF Working Paper 18/17*.
- Pinkovskiy, Maxim, and Xavier Sala-i Martin, 2016, “Lights, Camera . . . Income! Illuminating the National Accounts-Household Surveys Debate,” *The Quarterly Journal of Economics*, Vol. 131, No. 2, pp. 579–631.
- Schennach, Susanne M, and Yingyao Hu, 2013, “Nonparametric Identification and Semiparametric Estimation of Classical Measurement Error Models Without Side Information,” *Journal of the American Statistical Association*, Vol. 108, No. 501, pp. 177–186.
- Shen, Xiaotong, and Wing Hung Wong, 1994, “Convergence Rate of Sieve Estimates,” *The Annals of Statistics*, pp. 580–615.
- Storeygard, Adam, 2016, “Farther on down the Road: Transport Costs, Trade and Urban Growth in Sub-Saharan Africa,” *The Review of Economic Studies*, Vol. 83, No. 3, pp. 1263–1295.
- WORLD BANK, 2017, “Growth Out of the Blue,” *South Asia Economic Focus*.

## APPENDIX A. MATHEMATICAL PROOFS

### A.1. Assumptions

Let  $f(\cdot|\cdot)$  is a generic conditional probability density function. We make the following assumptions.

**Assumption 1.**

$$f(z_{i,t}|y_{i,t}^*, y_{i,t}, s_{i,t}, l_i) = f(z_{i,t}|y_{i,t}^*, l_i). \quad (9)$$

This assumption implies that the nighttime lights recorded by satellite are related to the true real GDP per capita and the geographic location of a country, but have nothing to do with how its GDP is measured or its statistical capacity.

**Assumption 2.**

$$f(y_{i,t}|y_{i,t}^*, s_{i,t}, l_i) = f(y_{i,t}|y_{i,t}^*, s_{i,t}). \quad (10)$$

The statistical capacity captures how accurate GDP is measured regardless of the location of the country.

Under these assumptions above, the joint probability density of nighttime lights per capita, official measures of real GDP per capita, statistical capacity, and location follows:

$$f(z_{i,t}, y_{i,t}, s_{i,t}, l_i) = \int f(z_{i,t}|y_{i,t}^*, l_i) f(y_{i,t}|y_{i,t}^*, s_{i,t}) f(y_{i,t}^*, s_{i,t}, l_i) dy_{i,t}^*$$

The identification results in [Carroll, Chen, and Hu \(2010\)](#) imply that  $f(z_{i,t}|y_{i,t}^*, l_i)$ ,  $f(y_{i,t}|y_{i,t}^*, s_{i,t})$ , and  $f(y_{i,t}^*, s_{i,t}, l_i)$  are uniquely determined by  $f(z_{i,t}, y_{i,t}, s_{i,t}, l_i)$ . In other words, the conditional distributions of nighttime lights and official measures of real GDP per capita are identified. The identification relies on reasonable assumptions, which are discussed below. For middle and low income countries, whose measurement error issue is more significant, our results don't require additivity of measurement errors, nor direct independence between true GDP per capita and its measurement error.

### A.2. Nonparametric Identification

We assume the existence of a random sample from distribution  $f(z_{i,t}, y_{i,t}, s_{i,t}, l_i)$  for country  $i$  in year  $t$ . We provide sufficient conditions for the identification of latent distributions:  $f(z_{i,t}|y_{i,t}^*, l_i)$ ,  $f(y_{i,t}|y_{i,t}^*, s_{i,t})$ , and  $f(y_{i,t}^*, s_{i,t}, l_i)$  from the observed distribution  $f(z_{i,t}, y_{i,t}, s_{i,t}, l_i)$ . In this subsection, we omit the subscript  $(i, t)$  for the simplicity of notations.

Suppose the supports of  $z, y, y^*, s$ , and  $l$  are  $\mathcal{Z} \subseteq \mathbb{R}$ ,  $\mathcal{Y} \subseteq \mathbb{R}$ ,  $\mathcal{Y}^* \subseteq \mathbb{R}$ ,  $\mathcal{S} = \{s_1, s_2, \dots, s_J\}$  with  $J \geq 2$ , and  $\mathcal{L} = \{l_1, l_2, \dots, l_K\}$  with  $K \geq 2$ , respectively. We assume



**Assumption 3.**  $f(z, y, y^*, s, l)$  is positive, bounded on its support  $\mathcal{Z} \times \mathcal{Y} \times \mathcal{Y}^* \times \mathcal{S} \times \mathcal{L}$ , and is continuous in  $(z, y, y^*) \in \mathcal{Z} \times \mathcal{Y} \times \mathcal{Y}^*$ .

**Assumption 4.** (i) for each given  $s$ ,  $\int f(y|y^*, s)h(y^*)dy^* = 0$  for all  $y \in \mathcal{Y}$  for all bounded function  $h$  implies that  $h \equiv 0$  over  $\mathcal{Y}^*$ ; (ii) for each given  $(s, l)$ ,  $\int f(z, y|s, l)h(y)dy = 0$ , i.e.,  $E[h(y)|z, s, l] = 0$ , for all  $z \in \mathcal{Z}$  for all bounded function  $h$  implies that  $h \equiv 0$  over  $\mathcal{Y}$ ;

Assumption 4 (i) is the *bounded completeness* of the conditional density  $f(y^*|y)$ ; see, e.g., Mattner (1993). Assumption 4(ii) is imposed on observables directly and is directly testable under some restrictions (Freyberger (2017)). Nevertheless, comparing with conditions for parametric identification, Assumption 4 is a high level condition mainly because we are achieving a nonparametric identification result.

Define

$$\kappa^{jk}(y^*) \equiv \frac{f(y^*|s_a, l_j) f(y^*|s_b, l_k)}{f(y^*|s_b, l_j) f(y^*|s_a, l_k)} \quad \text{for } y^* \in \mathcal{Y}^*. \quad (11)$$

**Assumption 5.** For any  $y_1^* \neq y_2^*$ , there exist  $j, k \in \{1, 2, \dots, K\}$ , such that  $\kappa^{jk}(y_1^*) \neq \kappa^{jk}(y_2^*)$  and  $\sup_{y^* \in \mathcal{Y}^*} \kappa^{jk}(y^*) < \infty$ .

This assumption requires that the distribution of real GDP varies with countries' statistical capacity and geographic location, which is quite reasonable. In the data, we observe that higher income countries have more effective statistical institutions. In the meantime, it is well known that the GDP is highly correlated with countries' geographic location. Appendix C.2 provides more details on the data that underpin this assumption.

Since  $y^*$  is not observed, we need a normalization assumption as follows:

**Assumption 6.** One of the followings holds for all  $y^* \in \mathcal{Y}^*$ : for some  $s_j$ , (i) (mean)  $E[y|y^*, s_j] = y^*$ ; or (ii) (mode)  $\arg \max_y f(y|y^*, s_j) = y^*$ ; or (iii) (median)  $\inf\{v : \int_{-\infty}^v f(y|y^*) dy \geq 0.5\} = y^*$ .

Assumption 6 says that the reported GDP from some country with statistical capacity  $s_j$  is targeted for the true  $y^*$ . Specifically, either the mean, mode or median of the distribution of  $y$  given  $y^*$  and  $s_j$  is equal to  $y^*$ . This condition is not required for other countries with different statistical capacity.

We summarize the nonparametric identification result as follows:

**Theorem 1.** Suppose Assumptions 1–6 hold. Then, the distribution function  $f(z, y, s, l)$  uniquely determines the joint distribution function  $f(z, y, y^*, s, l)$  satisfying

$$f(z, y, y^*, s, l) = f(z|y^*, l)f(y|y^*, s)f(y^*, s, l). \quad (12)$$

**Proof:** See Appendix A.3.

This theorem presents a set of sufficient conditions under which all the distributions containing the latent true GDP can be uniquely determined by the observed joint distribution of GDP and nighttime lights from countries with different statistical capacity and at different locations. Such a nonparametric identification result implies that consistent estimation is possible for parametric, semiparametric, or nonparametric specifications. In order to focus on the relationship between the nighttime lights and the latent true GDP and also to take the sample size into account, we adopt a simple specification to simplify the measurement error structure.

### A.3. Proofs

We define the integral operator  $L_{y|y^*,s_a} : \mathcal{L}^2(\mathcal{Y}^*) \rightarrow \mathcal{L}^2(\mathcal{Y})$  as:

$$\{L_{y|y^*,s_a} h\}(y) = \int_{\mathcal{Y}^*} f_{y|y^*,s}(y|y^*,s_a) h(y^*) dy^* \quad \text{for any } h \in \mathcal{L}^2(\mathcal{Y}^*), y \in \mathcal{Y}.$$

where  $\mathcal{L}^2(\mathcal{Y})$  denotes the space of functions with  $\int_{\mathcal{Y}} |h(y)|^2 dy < \infty$ . Similarly, we define

$$\begin{aligned} L_{y,z|s_a,l_j} &: \mathcal{L}^2(\mathcal{Z}) \rightarrow \mathcal{L}^2(\mathcal{Y}), & (L_{y,z|s_a,l_j} h)(y) &= \int f_{y,z|s,l}(y,z|s_a,l_j) h(z) dz, \\ L_{z|y^*,l_j} &: \mathcal{L}^2(\mathcal{Z}) \rightarrow \mathcal{L}^2(\mathcal{Y}^*), & (L_{z|y^*,l_j} h)(y^*) &= \int f_{z|y^*,l}(z|y^*,l_j) h(z) dz, \\ D_{y^*|s_a,l_j} &: \mathcal{L}^2(\mathcal{Y}^*) \rightarrow \mathcal{L}^2(\mathcal{Y}^*), & (D_{y^*|s_a,l_j} h)(y^*) &= f_{y^*|s,l}(y^*,s_a,l_j) h(y^*). \end{aligned}$$

Notice that the operator  $D_{y^*|s_a,l_j}$  is diagonal or multiplication operator, and the operator  $L_{y,z|s_a,l_j}$  is observed from the data.

**Proof of Theorem 1:** For each value  $(s,l)$ , assumptions 1 and 2 imply that

$$f_{y,z|s,l}(y,z|s_a,l_j) = \int f_{y|y^*,s}(y|y^*,s_a) f_{z|y^*,l}(z|y^*,l_j) f_{y^*|s,l}(y^*,s_a,l_j) dy^*, \quad (13)$$

$$f_{y,z|s,l}(y,z|s_b,l_j) = \int f_{y|y^*,s}(y|y^*,s_b) f_{z|y^*,l}(z|y^*,l_j) f_{y^*|s,l}(y^*,s_b,l_j) dy^*. \quad (14)$$

By equation (13) and the definition of the operators, we have, for any function  $h \in \mathcal{L}^2(\mathcal{Z})$ ,

$$\begin{aligned}
(L_{y,z|s_a,l_j} h)(y) &= \int f_{y,z|s,l}(y, z|s_a, l_j) h(z) dz \\
&= \int \left( \int f_{y|y^*,s}(y|y^*, s_a) f_{z|y^*,l}(z|y^*, l_j) f_{y^*|s,l}(y^*, s_a, l_j) dy^* \right) h(z) dz \\
&= \int f_{y|y^*,s}(y|y^*, s_a) f_{y^*|s,l}(y^*, s_a, l_j) \left( \int f_{z|y^*,l}(z|y^*, l_j) h(z) dz \right) dy^* \\
&= \int f_{y|y^*,s}(y|y^*, s_a) f_{y^*|s,l}(y^*, s_a, l_j) (L_{z|y^*,l_j} h)(y^*) dy^* \\
&= \int f_{y|y^*,s}(y|y^*, s_a) (D_{y^*|s_a,l_j} L_{z|y^*,l_j} h)(y^*) dy^* \\
&= (L_{y|y^*,s_a} D_{y^*|s_a,l_j} L_{z|y^*,l_j} h)(y).
\end{aligned}$$

This means we have the operator equivalence

$$L_{y,z|s_a,l_j} = L_{y|y^*,s_a} D_{y^*|s_a,l_j} L_{z|y^*,l_j}. \quad (15)$$

Similarly, we have,

$$L_{y,z|s_b,l_j} = L_{y|y^*,s_b} D_{y^*|s_b,l_j} L_{z|y^*,l_j}. \quad (16)$$

Note that the left-hand sides of equations (15) and (16) are observed.

Assumption 4 imply that all the operators involved in equations (15) and (16) are invertible.

Hence

$$L_{y,z|s_a,l_j} L_{y,z|s_b,l_j}^{-1} = L_{y|y^*,s_a} D_{y^*|s_a,l_j} D_{y^*|s_b,l_j}^{-1} L_{y|y^*,s_b}^{-1}. \quad (17)$$

This equation holds for all  $s_j$  and  $s_k$  so that we may then eliminate  $L_{y|y^*,s_b}$  to have

$$L_{y,y}^{jk} \equiv (L_{y,z|s_a,l_j} L_{y,z|s_b,l_j}^{-1}) (L_{y,z|s_a,l_k} L_{y,z|s_b,l_k}^{-1})^{-1} = L_{y|y^*,s_a} D_{y^*}^{jk} L_{y|y^*,s_a}^{-1}, \quad (18)$$

where  $D_{y^*}^{jk} : \mathcal{L}^2(\mathcal{Y}^*) \rightarrow \mathcal{L}^2(\mathcal{Y}^*)$  is still a diagonal operator

$$D_{y^*}^{jk} \equiv D_{y^*|s_a,l_j} D_{y^*|s_b,l_j}^{-1} (D_{y^*|s_a,l_k} D_{y^*|s_b,l_k}^{-1})^{-1}. \quad (19)$$

In fact, this diagonal operator can be defined as  $(D_{y^*}^{jk} h)(y^*) \equiv \kappa^{jk}(y^*) h(y^*)$  with  $\kappa^{jk}$  defined in equation (11). Equation (18) implies a diagonalization of an observed operator  $L_{y,y}^{jk}$ , where an eigenvalue of  $L_{y,y}^{jk}$  equals  $\kappa^{jk}(y^*)$  for a value of  $y^*$  with corresponding eigenfunction  $f_{y|y^*,s}(\cdot|y^*, s_a)$ . Notice that each eigenfunction is a conditional density, and therefore, is automatically normalized.

Equation (18) implies that the operator  $L_{y,y}^{jk}$  has the same spectrum as the diagonal operator  $D_{y^*}^{jk}$ . Since an operator is bounded by the largest element of its spectrum, Assumption 4 guarantees that the operator  $L_{y,y}^{jk}$  is bounded with distinctive eigenvalues. Following theorem

XV.4.3.5 in Dunford and Schwartz (1971), we have that the diagonal decomposition of  $L_{y,y}^{jk}$  is unique up to the index of eigenvalues and eigenfunctions. Notice that Assumption 4 guarantees that, for any two different eigenfunctions  $f_{y|y^*,s}(\cdot|y_1^*, s_a)$  and  $f_{y|y^*,s}(\cdot|y_2^*, s_a)$ , one can always find two subsets with  $l_j$  and  $l_k$  such that the two different eigenfunctions correspond to two different eigenvalues  $\kappa^{jk}(y_1^*)$  and  $\kappa^{jk}(y_2^*)$  and, therefore, are uniquely determined.

In order to fully identify each eigenfunction, we need to identify the exact value of  $y^*$  in each eigenfunction  $f_{y|y^*,s}(\cdot|y^*, s_a)$ . Here we use the ordering assumption in [Hu and Schennach \(2008\)](#), i.e. Assumption 6, to pin down the the exact value of  $y^*$  for each eigenfunction  $f_{y|y^*,s}(\cdot|y^*, s_a)$ . Such an identification procedure can be applied to each subpopulation with a different value of  $s$ . Thus, we have fully identified the conditional density  $f_{y|y^*,s}$ .

Given  $f_{y|y^*,s}$ , other densities containing  $y^*$  can also be identified due to the injectivity of operator  $L_{y|y^*,s}$  as follows:

$$f_{z,y^*,s,l} = L_{y|y^*,s}^{-1} f_{z,y,s,l} \quad (20)$$

In summary, we have shown that the density  $f(z, y, s, l)$  uniquely determines the joint density  $f(z, y, y^*, s, l)$  satisfying  $f(z, y, y^*, s, l) = f(y|y^*, s) f(z, y^*, s, l)$ .

#### A.4. Identification of the Simple Specification

Assumptions 1 and 2 are satisfied when the two error terms  $\varepsilon_{i,t}^y$  and  $\varepsilon_{i,t}^z$  are independent of each other and the latent true GDP. Assumption 3 holds when the distributions of the error terms and the latent true GDP are bounded and continuous and the function  $m$  is bounded and continuous. Assumption 4 is a high-level condition. Part (i) requires that the characteristic function of the error term  $\varepsilon_{i,t}^y$  does not vanish on the real line. Assumption 5 requires that the distribution of latent true GDP varies with statistical capacity and location. Assumption 6 requires that for some category for the statistical capacity the error term  $\varepsilon_{i,t}^y$  has a zero mean, a zero mode, or a zero median.

Ideally, we should observe the statistical capacity  $s_{i,t}$  for all the countries. The world bank, however, only provides this measure for middle and low income countries. Therefore, we assign an additional category for the discretized statistical capacity to represent the high income group. Since this assigned category coincides with the high income group, the support of the true GDP conditional on this category is different from the support of true GDP conditional on different statistical capacity and location in the middle or low income group. That means the assigned category can't serve as  $j$  and  $k$  in Assumption 5.

In addition, identification of the error distribution corresponding to this assigned category has to rely on this simple specification above. We use Theorem 1 in [Schennach and Hu \(2013\)](#) to show that the distributions of  $y^*$ ,  $\varepsilon^y$ , and  $\varepsilon^z$ , and function  $m(\cdot)$  are identified under assumptions as follows: i) the errors  $\varepsilon^y$  and  $\varepsilon^z$  are mutually independent with a zero mean and also jointly independent of  $y^*$ ; ii) the characteristic functions of  $y$  and  $z$  do not vanish on the real

line; and iii) function  $m(\cdot)$  is monotone and continuously differentiable, and does not belong to a particular parametric family, which includes linear functions. Given that these assumptions are relatively mild comparing with those in the existing literature, we adopt this simple specification instead of imputing the statistical capacity for the high income counties.

### A.5. Sieve Maximum Likelihood Estimation

Given the general nonparametric identification, we provide a seminonparametric estimator as suggested in [Carroll, Chen, and Hu \(2010\)](#). We develop our estimator based on an i.i.d sample, which can be extended to for time series data. We assume that there is a random sample  $\{z_i, y_i, s_i, l_i\}_{i=1}^n$ .

We adopt a parametric specification of function  $m(\cdot; \theta)$  and leave other elements nonparametrically specified in the simple specification in equations (1) and (2). Let the true value of the unknowns be  $\alpha_0 \equiv (\theta_0^T, f_{01}, f_{02}, f_{03})^T \equiv (\theta_0^T, f_{y^*|s,l}, f_{\mathcal{E}^y|s}, f_{\mathcal{E}^z|l})^T$ , where  $f_{A|B}$  denotes the distribution of  $A$  conditional on  $B$ . We then introduce a sieve MLE estimator  $\hat{\alpha}$  for  $\alpha_0$ , and establish the asymptotic normality of  $\hat{\theta}$ . These results can also be extended to the case where the function  $m$  is misspecified.

Following [Carroll, Chen, and Hu \(2010\)](#), we consider the widely used Hölder space of functions. Let  $\xi = (\xi_1, \xi_2, \xi_3)^T \in \mathbb{R}^3$ ,  $a = (a_1, a_2, a_3)^T$ , and  $\nabla^a h(\xi) \equiv \frac{\partial^{a_1+a_2+a_3} h(\xi_1, \xi_2, \xi_3)}{\partial \xi_1^{a_1} \partial \xi_2^{a_2} \partial \xi_3^{a_3}}$  denote the  $(a_1 + a_2 + a_3)^{\text{th}}$  derivative. Let  $\|\cdot\|_E$  denote the Euclidean norm. Let  $\mathcal{V} \subseteq \mathbb{R}^3$  and  $\underline{\gamma}$  be the largest integer satisfying  $\gamma > \underline{\gamma}$ . The Hölder space  $\Lambda^\gamma(\mathcal{V})$  of order  $\gamma > 0$  is a space of functions  $h : \mathcal{V} \mapsto \mathbb{R}$ , such that the first  $\underline{\gamma}$  derivatives are continuous and bounded, and the  $\underline{\gamma}^{\text{th}}$  derivative is Hölder continuous with the exponent  $\gamma - \underline{\gamma} \in (0, 1]$ . We define a Hölder ball as  $\Lambda_c^\gamma(\mathcal{V}) \equiv \{h \in \Lambda^\gamma(\mathcal{V}) : \|h\|_{\Lambda^\gamma} \leq c < \infty\}$ , in which

$$\|h\|_{\Lambda^\gamma} \equiv \max_{a_1+a_2+a_3 \leq \underline{\gamma}} \sup_{\xi} |\nabla^a h(\xi)| + \max_{a_1+a_2+a_3 = \underline{\gamma}} \sup_{\xi \neq \xi'} \frac{|\nabla^a h(\xi) - \nabla^a h(\xi')|}{(\|\xi - \xi'\|_E)^{\gamma - \underline{\gamma}}} < \infty.$$

The space containing  $f_{01} = f_{y^*|s,l}$  are assumed to be

$$\mathcal{F}_1 = \left\{ \begin{array}{l} f_1(\cdot|\cdot, \cdot) \in \Lambda_c^{\underline{\gamma}_1}(\mathcal{Y}^* \times \mathcal{S} \times \mathcal{L}) : \text{Assumption 5 holds,} \\ f_3(\cdot|s, l) \text{ is a positive density function for all } s \in \mathcal{S}, l \in \mathcal{L} \end{array} \right\}.$$

Similarly, we assume  $f_{02}$  and  $f_{03}$  are in the following functional spaces

$$\mathcal{F}_2 = \left\{ \begin{array}{l} f_2(\cdot|\cdot) \in \Lambda_c^{\underline{\gamma}_2}(\mathcal{E}^y \times \mathcal{S}) : \text{Assumption 6 holds,} \\ f_2(\cdot|s) \text{ is a positive density function for all } s \in \mathcal{S} \end{array} \right\},$$

and

$$\mathcal{F}_3 = \{f_3(\cdot|\cdot) \in \Lambda_c^{\mathcal{Y}_3}(\mathcal{E}^z \times \mathcal{L}) : f_3(\cdot|l) \text{ is a positive density function for all } l \in \mathcal{L}\},$$

where  $\mathcal{E}^y$  and  $\mathcal{E}^z$  are supports of the error terms in equations (1) and (2), respectively.

Let  $\mathcal{A} = \Theta \times \mathcal{F}_1 \times \mathcal{F}_2 \times \mathcal{F}_3$  as the parameter space. The log-joint likelihood for  $\alpha \equiv (\theta^T, f_1, f_2, f_3)^T \in \mathcal{A}$  is given by:

$$\sum_{i=1}^n \log f(z_i, y_i, s_i, l_i) = \sum_{i=1}^n \ell(D_i; \alpha),$$

in which  $D_i = (z_i, y_i, s_i, l_i)$  and

$$\begin{aligned} \ell(D_i; \alpha) &\equiv \ell(z_i, y_i, s_i, l_i; \theta, f_1, f_2, f_3) \\ &= \log \left\{ \int f_1(y^*|s_i, l_i) f_2(y_i - y^*|s_i) f_3(z_i - m(y^*; \theta)|l_i) dy^* \right\} + \log f(s_i, l_i). \end{aligned}$$

Let  $E[\cdot]$  denote the expectation with respect to the underlying true data generating process for  $D_i$ . Then

$$\alpha_0 = \arg \sup_{\alpha \in \mathcal{A}} E[\ell(D_i; \alpha)].$$

We then use a sequence of finite-dimensional sieve spaces  $\mathcal{A}_n = \Theta \times \mathcal{F}_1^n \times \mathcal{F}_2^n \times \mathcal{F}_3^n$  to approximate the functional space  $\mathcal{A} = \Theta \times \mathcal{F}_1 \times \mathcal{F}_2 \times \mathcal{F}_3$ . The seminonparametric sieve MLE  $\hat{\alpha}_n = (\hat{\theta}^T, \hat{f}_1, \hat{f}_2, \hat{f}_3)^T \in \mathcal{A}_n$  for  $\alpha_0 \in \mathcal{A}$  is defined as:

$$\hat{\alpha}_n = \arg \max_{\alpha \in \mathcal{A}_n} \sum_{i=1}^n \ell(D_i; \alpha).$$

Let  $p^{k_n}(\cdot)$  be a  $k_n \times 1$ -vector of known basis functions, such as power series, splines, Fourier series, Legendre polynomials, Hermite polynomials, etc. We use linear sieves to directly approximate unknown densities:

$$\mathcal{F}_1^n = \left\{ f_1(y^*|s, l) = \left[ \sum_{i=1}^K \sum_{j=1}^J p^{k_{1,n}}(y^*)^T \beta_{1,i,j} I(l = l_i) I(s = s_j) \right]^2 \in \mathcal{F}_1 \right\}$$

$$\mathcal{F}_2^n = \left\{ f_2(e|s) = \left[ \sum_{j=1}^J p^{k_{2,n}}(e)^T \beta_{2,j} I(s = s_j) \right]^2 \in \mathcal{F}_2 \right\}$$

$$\mathcal{F}_3^n = \left\{ f_3(e|l) = \left[ \sum_{i=1}^K p^{k_{3,n}}(e)^T \beta_{3,i} I(l = l_i) \right]^2 \in \mathcal{F}_3 \right\}.$$

Below we present the asymptotic properties of the proposed estimator.

### A.6. Consistency

Here we provide sufficient conditions for the consistency of the sieve estimator  $\widehat{\alpha}_n = \left( \widehat{\theta}^T, \widehat{f}_1, \widehat{f}_2, \widehat{f}_3 \right)^T$ .

**Assumption 7.** (i) All the assumptions in theorem 1 hold; (ii)  $f_{y^*|s,l}(\cdot|s,l) \in \mathcal{F}_1$  with  $\gamma_1 > 1/2$  for all  $s \in \mathcal{S}$  and  $l \in \mathcal{L}$ ; (iii)  $f_{\varepsilon^v|s}(\cdot|\cdot) \in \mathcal{F}_2$  with  $\gamma_2 > 1$ ; (iv)  $f_{\varepsilon^z|l}(\cdot|\cdot) \in \mathcal{F}_3$  with  $\gamma_3 > 1$ .

**Assumption 8.** (i)  $\{z_i, y_i, s_i, l_i\}_{i=1}^n$  is i.i.d.; (ii)  $m(y^*; \theta)$  is continuous in  $\theta \in \Theta$ , and  $\Theta$  is a compact subset of  $\mathbb{R}^{d_\theta}$ ; (iii)  $\theta_0 \in \Theta$  is the unique solution of  $E[z|y^*, l] = m(y^*; \theta)$  over  $\theta \in \Theta$ .

We define a norm on  $\mathcal{A}$  as:  $\|\alpha\|_s = \|\theta\|_E + \|f_1\|_{\infty, \omega_1} + \|f_2\|_{\infty, \omega_2} + \|f_3\|_{\infty, \omega_3}$  in which  $\|h\|_{\infty, \omega_j} \equiv \sup_{\xi} |h(\xi) \omega_j(\xi)|$  with  $\omega_j(\xi) = \left(1 + \|\xi\|_E^2\right)^{-\varsigma_j/2}$ ,  $\varsigma_j > 0$  for  $j = 1, 2, 3$ . We assume

**Assumption 9.** (i)  $-\infty < E[\ell(D_i; \alpha_0)] < \infty$ ,  $E[\ell(D_i; \alpha)]$  is upper semicontinuous on  $\mathcal{A}$  under the metric  $\|\cdot\|_s$ ; (ii) there is a finite  $\tau > 0$  and a random variable  $U(D_i)$  with  $E\{U(D_i)\} < \infty$  such that  $\sup_{\alpha \in \mathcal{A}_n: \|\alpha - \alpha_0\|_s \leq \delta} |\ell(D_i; \alpha) - \ell(D_i; \alpha_0)| \leq \delta^\tau U(D_i)$ .

**Assumption 10.** (i)  $p^{k_{j,n}}(\cdot)$  is a  $k_{j,n} \times 1$ -vector of basis functions on  $\mathbb{R}$  for  $j = 1, 2, 3$ ; (ii)  $\min\{k_{1,n}, k_{2,n}, k_{3,n}\} \rightarrow \infty$  and  $\max\{k_{1,n}, k_{2,n}, k_{3,n}\}/n \rightarrow 0$ .

We then have

**Lemma 1.** Under Assumptions 7–10, we have  $\|\widehat{\alpha}_n - \alpha_0\|_s = o_p(1)$ .

This is a direct extension from [Carroll, Chen, and Hu \(2010\)](#), which uses theorem 3.1 in [Chen \(2007\)](#).

### A.7. Convergence Rates and Asymptotic Normality

The asymptotic properties of our estimator is a direct extension of that in [Carroll, Chen, and Hu \(2010\)](#). We list the conditions below for readers' convenience.

### A.7.1. Convergence Rates of Nonparametric Part

Given the consistency shown in Lemma 1, we focus on a shrinking  $\|\cdot\|_s$ -neighborhood around  $\alpha_0$ . Let  $\mathcal{A}_{0s} \equiv \{\alpha \in \mathcal{A} : \|\alpha - \alpha_0\|_s = o(1), \|\alpha\|_s \leq c_0 < c\}$  and  $\mathcal{A}_{0sn} \equiv \{\alpha \in \mathcal{A}_n : \|\alpha - \alpha_0\|_s = o(1), \|\alpha\|_s \leq c_0 < c\}$ . We assume that both  $\mathcal{A}_{0s}$  and  $\mathcal{A}_{0sn}$  are convex parameter spaces, and that  $\ell(D_i; \alpha + \tau v)$  is twice continuously differentiable at  $\tau = 0$  for almost all  $D_i$  and any direction  $v \in \mathcal{A}_{0s}$ .

Define the pathwise first and second derivatives of the sieve loglikelihood in the direction  $v$  as

$$\frac{d\ell(D_i; \alpha)}{d\alpha} [v] \equiv \left. \frac{d\ell(D_i; \alpha + \tau v)}{d\tau} \right|_{\tau=0}; \quad \frac{d^2\ell(D_i; \alpha)}{d\alpha d\alpha^T} [v, v] \equiv \left. \frac{d^2\ell(D_i; \alpha + \tau v)}{d\tau^2} \right|_{\tau=0}.$$

Mimicing Ai and Chen (2007), for any  $\alpha_1, \alpha_2 \in \mathcal{A}_{0s}$ , we define a pseudo metric  $\|\cdot\|_2$  as

$$\|\alpha_1 - \alpha_2\|_2 \equiv \sqrt{-E \left( \frac{d^2\ell(D_i; \alpha_0)}{d\alpha d\alpha^T} [\alpha_1 - \alpha_2, \alpha_1 - \alpha_2] \right)}.$$

Our goal is to show that  $\widehat{\alpha}_n$  converges to  $\alpha_0$  at a rate faster than  $n^{-1/4}$  under the pseudo metric  $\|\cdot\|_2$ . We make the following assumptions:

**Assumption 11.** (i)  $\varsigma_j > \gamma_j$  for  $j = 1, 2, 3$ ; (ii)  $\max\{k_{1,n}^{-\gamma_1/2}, k_{2,n}^{-\gamma_2/2}, k_{3,n}^{-\gamma_3}\} = o(n^{-1/4})$ .

**Assumption 12.** (i)  $\mathcal{A}_{0s}$  is convex at  $\alpha_0$  and  $\theta_0 \in \text{int}(\Theta)$ ; (ii)  $\ell(D_i; \alpha)$  is twice continuously pathwise differentiable with respect to  $\alpha \in \mathcal{A}_{0s}$ , and  $m(y^*; \theta)$  is twice continuously differentiable at  $\theta_0$ .

**Assumption 13.**  $\sup_{\tilde{\alpha} \in \mathcal{A}_{0s}} \sup_{\alpha \in \mathcal{A}_{0sn}} \left| \frac{d\ell(D_i; \tilde{\alpha})}{d\alpha} \left[ \frac{\alpha - \alpha_0}{\|\alpha - \alpha_0\|_s} \right] \right| \leq U(D_i)$  for a random variable  $U(D_i)$  with  $E\{[U(D_i)]^2\} < \infty$ .

**Assumption 14.** (i)  $\sup_{v \in \mathcal{A}_{0s}: \|v\|_s=1} -E \left( \frac{d^2\ell(D_i; \alpha_0)}{d\alpha d\alpha^T} [v, v] \right) \leq C < \infty$ ; (ii) uniformly over  $\tilde{\alpha} \in \mathcal{A}_{0s}$  and  $\alpha \in \mathcal{A}_{0sn}$ , we have

$$-E \left( \frac{d^2\ell(D_i; \tilde{\alpha})}{d\alpha d\alpha^T} [\alpha - \alpha_0, \alpha - \alpha_0] \right) = \|\alpha - \alpha_0\|_2^2 \times \{1 + o(1)\}.$$

These assumptions are standard in the literature. As a direct application of Theorem 3.2 of Shen and Wong (1994) to the local parameter space  $\mathcal{A}_{0s}$  and the local sieve space  $\mathcal{A}_{0sn}$ , we have

**Theorem 2.** Let  $\gamma \equiv \min\{\gamma_1/2, \gamma_2/2, \gamma_3\} > 1/2$ . Under assumptions 7–14, if  $k_{1,n} = O\left(n^{\frac{1}{\gamma_1+1}}\right)$ ,  $k_{2,n} = O\left(n^{\frac{1}{\gamma_2+1}}\right)$ , and  $k_{3,n} = O\left(n^{\frac{1}{2\gamma_3+1}}\right)$ , then

$$\|\widehat{\alpha}_n - \alpha_0\|_2 = O_P\left(n^{\frac{-\gamma}{2\gamma+1}}\right) = o_P\left(n^{-1/4}\right).$$



### A.7.2. Asymptotic Normality of Parametric Part

This section presents sufficient conditions for the asymptotic normality of the parametric part of the model. Define an inner product corresponding to the pseudo metric  $\|\cdot\|_2$ :

$$\langle v_1, v_2 \rangle_2 \equiv -E \left[ \frac{d^2 \ell(D_i; \alpha_0)}{d\alpha d\alpha^T} [v_1, v_2] \right],$$

where

$$\frac{d^2 \ell(D_i; \alpha_0)}{d\alpha d\alpha^T} [v_1, v_2] \equiv \frac{d^2 \ell(D_i; \alpha_0 + \tau_1 v_1 + \tau_2 v_2)}{d\tau_1 d\tau_2} \Big|_{\tau_1 = \tau_2 = 0}.$$

Let  $\bar{\mathbf{V}}$  denote the closure of the linear span of  $\mathcal{A} - \{\alpha_0\}$  under the metric  $\|\cdot\|_2$ . Then  $(\bar{\mathbf{V}}, \|\cdot\|_2)$  is a Hilbert space. We define  $\bar{\mathbf{V}} = \mathbb{R}^{d_\theta} \times \bar{\mathcal{U}}$  with  $\bar{\mathcal{U}} \equiv \overline{\mathcal{F}_1 \times \mathcal{F}_2 \times \mathcal{F}_3} - \{(f_{01}, f_{02}, f_{03})\}$  and let  $h = (f_1, f_2, f_3)$  denote all the unknown densities. The pathwise first derivative can be written as

$$\begin{aligned} \frac{d\ell(D_i; \alpha_0)}{d\alpha} [\alpha - \alpha_0] &= \frac{d\ell(D_i; \alpha_0)}{d\theta^T} (\theta - \theta_0) + \frac{d\ell(D_i; \alpha_0)}{dh} [h - h_0] \\ &= \left( \frac{d\ell(D_i; \alpha_0)}{d\theta^T} - \frac{d\ell(D_i; \alpha_0)}{dh} [\mu] \right) (\theta - \theta_0), \end{aligned}$$

with  $h - h_0 \equiv -\mu \times (\theta - \theta_0)$ , and in which

$$\begin{aligned} \frac{d\ell(D_i; \alpha_0)}{dh} [h - h_0] &= \frac{d\ell(D_i; \theta_0, h_0(1 - \tau) + \tau h)}{d\tau} \Big|_{\tau=0} \\ &= \frac{d\ell(D_i; \alpha_0)}{df_1} [f_1 - f_{01}] + \frac{d\ell(D_i; \alpha_0)}{df_{1a}} [f_{1a} - f_{01a}] \\ &\quad + \frac{d\ell(D_i; \alpha_0)}{df_2} [f_2 - f_{02}] + \frac{d\ell(D_i; \alpha_0)}{df_{2a}} [f_{2a} - f_{02a}]. \end{aligned}$$

Note that

$$\begin{aligned} &E \left( \frac{d^2 \ell(D_i; \alpha_0)}{d\alpha d\alpha^T} [\alpha - \alpha_0, \alpha - \alpha_0] \right) \\ &= (\theta - \theta_0)^T E \left( \frac{d^2 \ell(D_i; \alpha_0)}{d\theta d\theta^T} - 2 \frac{d^2 \ell(D_i; \alpha_0)}{d\theta dh^T} [\mu] + \frac{d^2 \ell(D_i; \alpha_0)}{dh dh^T} [\mu, \mu] \right) (\theta - \theta_0), \end{aligned}$$

with  $h - h_0 \equiv -\mu \times (\theta - \theta_0)$ , and in which

$$\begin{aligned} \frac{d^2 \ell(D_i; \alpha_0)}{d\theta dh^T} [h - h_0] &= \frac{d(\partial \ell(D_i; \theta_0, h_0(1 - \tau) + \tau h) / \partial \theta)}{d\tau} \Big|_{\tau=0}, \\ \frac{d^2 \ell(D_i; \alpha_0)}{dh dh^T} [h - h_0, h - h_0] &= \frac{d^2 \ell(D_i; \theta_0, h_0(1 - \tau) + \tau h)}{d\tau^2} \Big|_{\tau=0}. \end{aligned}$$

For each component  $\theta^k$  (of  $\theta$ ),  $k = 1, \dots, d_\theta$ , suppose there exists a  $\mu^{*k} \in \overline{\mathcal{U}}$  that solves:

$$\mu^{*k} : \inf_{\mu^k \in \overline{\mathcal{U}}} E \left\{ - \left( \frac{\partial^2 \ell(D_i; \alpha_0)}{\partial \theta^k \partial \theta^k} - 2 \frac{d^2 \ell(D_i; \alpha_0)}{\partial \theta^k d h^T} [\mu^k] + \frac{d^2 \ell(D_i; \alpha_0)}{d h d h^T} [\mu^k, \mu^k] \right) \right\}.$$

Denote  $\mu^* = (\mu^{*1}, \mu^{*2}, \dots, \mu^{*d_\theta})$  with each  $\mu^{*k} \in \overline{\mathcal{U}}$ , and

$$\begin{aligned} \frac{d \ell(D_i; \alpha_0)}{d h} [\mu^*] &= \left( \frac{d \ell(D_i; \alpha_0)}{d h} [\mu^{*1}], \dots, \frac{d \ell(D_i; \alpha_0)}{d h} [\mu^{*d_\theta}] \right), \\ \frac{d^2 \ell(D_i; \alpha_0)}{\partial \theta d h^T} [\mu^*] &= \left( \frac{d^2 \ell(D_i; \alpha_0)}{\partial \theta d h} [\mu^{*1}], \dots, \frac{d^2 \ell(D_i; \alpha_0)}{\partial \theta d h} [\mu^{*d_\theta}] \right), \\ \frac{d^2 \ell(D_i; \alpha_0)}{d h d h^T} [\mu^*, \mu^*] &= \begin{pmatrix} \frac{d^2 \ell(D_i; \alpha_0)}{d h d h^T} [\mu^{*1}, \mu^{*1}] & \dots & \frac{d^2 \ell(D_i; \alpha_0)}{d h d h^T} [\mu^{*1}, \mu^{*d_\theta}] \\ \dots & \dots & \dots \\ \frac{d^2 \ell(D_i; \alpha_0)}{d h d h^T} [\mu^{*d_\theta}, \mu^{*1}] & \dots & \frac{d^2 \ell(D_i; \alpha_0)}{d h d h^T} [\mu^{*d_\theta}, \mu^{*d_\theta}] \end{pmatrix}. \end{aligned}$$

We also define

$$V_* \equiv -E \left( \frac{\partial^2 \ell(D_i; \alpha_0)}{\partial \theta \partial \theta^T} - 2 \frac{d^2 \ell(D_i; \alpha_0)}{\partial \theta d h^T} [\mu^*] + \frac{d^2 \ell(D_i; \alpha_0)}{d h d h^T} [\mu^*, \mu^*] \right). \quad (21)$$

We then consider a linear functional of  $\alpha$ , which is  $\lambda^T \theta$  for any  $\lambda \in \mathbb{R}^{d_\theta}$  with  $\lambda \neq 0$ . Since

$$\begin{aligned} & \sup_{\alpha - \alpha_0 \neq 0} \frac{|\lambda^T (\theta - \theta_0)|^2}{\|\alpha - \alpha_0\|_2^2} \\ &= \sup_{\theta \neq \theta_0, \mu \neq 0} \frac{(\theta - \theta_0)^T \lambda \lambda^T (\theta - \theta_0)}{(\theta - \theta_0)^T E \left\{ - \left( \frac{d^2 \ell(D_i; \alpha_0)}{d \theta d \theta^T} - 2 \frac{d^2 \ell(D_i; \alpha_0)}{d \theta d h^T} [\mu] + \frac{d^2 \ell(D_i; \alpha_0)}{d h d h^T} [\mu, \mu] \right) \right\} (\theta - \theta_0)} \\ &= \lambda^T (V_*)^{-1} \lambda, \end{aligned}$$

the functional  $\lambda^T (\theta - \theta_0)$  is *bounded* if and only if the matrix  $V_*$  is nonsingular.

Suppose that  $V_*$  is nonsingular. For any fixed  $\lambda \neq 0$ , denote  $v^* \equiv (v_\theta^*, v_h^*)$  with  $v_\theta^* \equiv (V_*)^{-1} \lambda$  and  $v_h^* \equiv -\mu^* \times v_\theta^*$ . Then the Riesz representation theorem implies:  $\lambda^T (\theta - \theta_0) = \langle v^*, \alpha - \alpha_0 \rangle_2$  for all  $\alpha \in \mathcal{A}$ . We can show equation (22) as follows:

$$\lambda^T (\hat{\theta}_n - \theta_0) = \langle v^*, \hat{\alpha}_n - \alpha_0 \rangle_2 = \frac{1}{n + n_a} \sum_{i=1}^n \frac{d \ell(D_i; \alpha_0)}{d \alpha} [v^*] + o_p\{n^{-1/2}\}. \quad (22)$$

Denote  $\mathcal{N}_0 = \{\alpha \in \mathcal{A}_{0s} : \|\alpha - \alpha_0\|_2 = o(n^{-1/4})\}$  and  $\mathcal{N}_{0n} = \{\alpha \in \mathcal{A}_{0sn} : \|\alpha - \alpha_0\|_2 = o(n^{-1/4})\}$ . We provide additional sufficient for asymptotic normality of sieve MLE  $\hat{\theta}_n$  as follows:

**Assumption 15.**  $\mu^*$  exists (i.e.,  $\mu^{*k} \in \overline{\mathcal{U}}$  for  $k = 1, \dots, d_\theta$ ), and  $V_*$  is positive-definite.

**Assumption 16.** There is a  $v_n^* \in \mathcal{A}_n - \{\alpha_0\}$ , such that  $\|v_n^* - v^*\|_2 = o(1)$  and  $\|v_n^* - v^*\|_2 \times \|\widehat{\alpha}_n - \alpha_0\|_2 = o_P(\frac{1}{\sqrt{n}})$ .

**Assumption 17.** There is a random variable  $U(D_i)$  with  $E\{[U(D_i)]^2\} < \infty$  and a non-negative measurable function  $\eta$  with  $\lim_{\delta \rightarrow 0} \eta(\delta) = 0$ , such that, for all  $\alpha \in \mathcal{N}_{0n}$ ,

$$\sup_{\overline{\alpha} \in \mathcal{N}_0} \left| \frac{d^2 \ell(D_i; \overline{\alpha})}{d\alpha d\alpha^T} [\alpha - \alpha_0, v_n^*] \right| \leq U(D_i) \times \eta(\|\alpha - \alpha_0\|_s).$$

**Assumption 18.** Uniformly over  $\overline{\alpha} \in \mathcal{N}_0$  and  $\alpha \in \mathcal{N}_{0n}$ ,

$$E \left( \frac{d^2 \ell(D_i; \overline{\alpha})}{d\alpha d\alpha^T} [\alpha - \alpha_0, v_n^*] - \frac{d^2 \ell(D_i; \alpha_0)}{d\alpha d\alpha^T} [\alpha - \alpha_0, v_n^*] \right) = o \left( \frac{1}{\sqrt{n}} \right).$$

**Assumption 19.**  $E \left\{ \left( \frac{d\ell(D_i; \alpha_0)}{d\alpha} [v_n^* - v^*] \right)^2 \right\}$  goes to zero as  $\|v_n^* - v^*\|_2$  goes to zero.

Recall the definitions of Fisher inner product and the Fisher norm:

$$\langle v_1, v_2 \rangle \equiv E \left\{ \left( \frac{d\ell(D_i; \alpha_0)}{d\alpha} [v_1] \right) \left( \frac{d\ell(D_i; \alpha_0)}{d\alpha} [v_2] \right) \right\}, \quad \|v\| \equiv \sqrt{\langle v, v \rangle}.$$

Under correct specification,  $m(y^*; \theta_0) = E(z|y^*, l)$ , it can be shown that  $\|v\| = \|v\|_2$  and  $\langle v_1, v_2 \rangle = \langle v_1, v_2 \rangle_2$ . Thus, the space  $\overline{\mathbf{V}}$  is also the closure of the linear span of  $\mathcal{A} - \{\alpha_0\}$  under the Fisher metric  $\|\cdot\|$ .

Suppose that  $\theta$  has  $d_\theta$  components, and write its  $k^{\text{th}}$  component as  $\theta^k$ . Write  $\mu^* = (\mu^{*1}, \mu^{*2}, \dots, \mu^{*d_\theta})$ , where we compute  $\mu^{*k} \equiv (\mu_1^{*k}, \mu_2^{*k}, \mu_3^{*k})^T \in \overline{\mathcal{U}}$  as the solution to

$$\begin{aligned} & \inf_{\mu^k \in \overline{\mathcal{U}}} E \left\{ \left( \frac{d\ell(D_i; \alpha_0)}{d\theta^k} - \frac{d\ell(D_i; \alpha_0)}{dh} [\mu^k] \right)^2 \right\} \\ &= \inf_{(\mu_1, \mu_2, \mu_3)^T \in \overline{\mathcal{U}}} E \left\{ \left( \begin{array}{c} \frac{d\ell(D_i; \alpha_0)}{d\theta^k} - \frac{d\ell(D_i; \alpha_0)}{df_1} [\mu_1] \\ -\frac{d\ell(D_i; \alpha_0)}{df_2} [\mu_2] - \frac{d\ell(D_i; \alpha_0)}{df_3} [\mu_3] \end{array} \right)^2 \right\}. \end{aligned}$$

This equation also defines  $\frac{d\ell(D_i; \alpha_0)}{dh} [\mu^*]$ . Then  $\mathcal{S}_{\theta_0} \equiv \frac{d\ell(D_i; \alpha_0)}{d\theta^T} - \frac{d\ell(D_i; \alpha_0)}{dh} [\mu^*]$  becomes the semiparametric efficient score for  $\theta_0$ , and

$$I_* \equiv E [\mathcal{S}_{\theta_0}^T \mathcal{S}_{\theta_0}] = V_* \tag{23}$$

becomes the semiparametric information bound for  $\theta_0$ .

Finally, we can show that the sieve MLE  $\widehat{\theta}_n$  is asymptotically normally distributed around  $\theta_0$  as follows:

**Theorem 3.** *Suppose that Assumptions of Lemma 1, and Assumptions 11–19 hold. Then:  $\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} N(0, V_*^{-1} I_* V_*^{-1})$ , with  $V_*$  defined in equation (21) and  $I_*$  given by equation (23).*

## APPENDIX B. ROBUSTNESS CHECKS: SIMULATIONS, SPECIFICATIONS, AND ESTIMATES

In this section, we conduct simulation exercises to confirm our estimation strategy and do a number of robustness checks with respect to data and specification.

### B.1. Simulations

We consider a data generating process similar to equations (1) and (2) with a quadratic nighttime light production function. There are six equal-sized groups of countries based on statistical capacity ( $s_i = 1, 2, 3$ ) and location ( $l_i = 1, 2$ ). Each group's true GDP distribution follows a mixture of two normal distributions. Measurement errors in both GDP per capita and nighttime lights follow normal distributions where the variances  $\sigma(s)$  and  $\sigma(l)$  differ for different groups. The quadratic function  $m(\cdot)$  is assumed to have the same coefficients as point estimates in the DMSP/OLS data. In the simulation, we draw 400 samples with size  $n = 4000$ . Table 7 shows the parameter details in the simulations.

**Table 7. Parameterization in Simulations**

Group	$y^*$	$\sigma(s)$	$\sigma(l)$
1	$0.6N(-0.8, 0.6^2) + 0.4N(0.8, 0.7^2)$	0.4	1
2	$0.7N(-1, 0.9^2) + 0.3N(0.7, 0.8^2)$	0.4	0.8
3	$0.4N(-0.1, 0.6^2) + 0.6N(0.2, 0.8^2)$	0.3	1
4	$0.2N(-0.4, 0.7^2) + 0.8N(0.6, 0.7^2)$	0.3	0.8
5	$0.3N(-0.2, 0.8^2) + 0.7N(1, 0.6^2)$	0.1	1
6	$0.5N(-1, 1) + 0.5N(-0.4, 0.6^2)$	0.1	0.8

Table 8 presents the simulation results with different choices of the number of orthogonal Hermite terms. The coefficients of the quadratic production function are accurately estimated around  $k = 6$ . For this reason, we apply  $k = 6$  to DMSP/OLS data and  $k = 4$  to VIIRS data since the latter has much less observations.

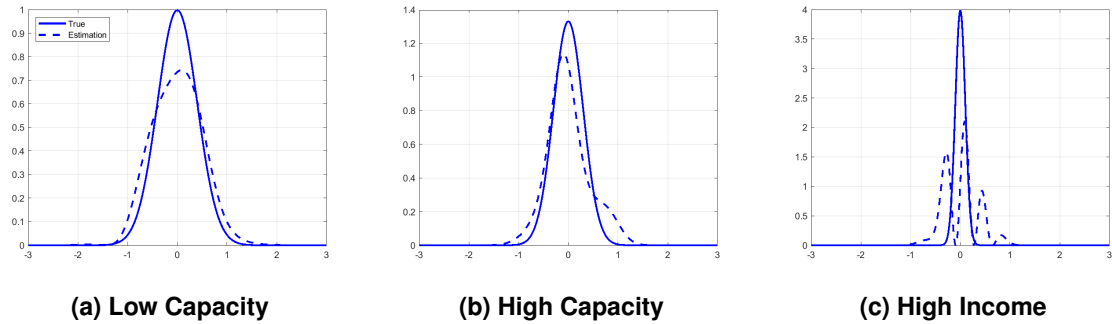
**Table 8. Simulation Results**

$$m(y^*) = \theta_0 + \theta_1 y^* + \theta_2 (y^*)^2$$

Parameter	$\theta_0$	$\theta_1$	$\theta_2$
True values	0.398	1.234	-0.244
$k = 4$	0.364 (0.083)	1.343 (0.075)	-0.229 (0.065)
$k = 5$	0.370 (0.039)	1.334 (0.047)	-0.234 (0.030)
$k = 6$	0.365 (0.041)	1.328 (0.111)	-0.227 (0.030)
$k = 8$	0.357 (0.091)	1.311 (0.140)	-0.221 (0.058)

Standard errors are based on 400 sample bootstraps.

To give an idea of how well orthogonal Hermite series approximate density functions, Figure 9 shows an example of the approximated density functions of measurement errors in GDP per capita in one of the simulations ( $k = 6$ ). The approximation is broadly in line with the true distributions.

**Figure 9. Example of Density Function Approximation**

In Section V.C, we proposed two new measures of real GDP per capita: the optimal linear measure and the semiparametric conditional mean. While the latter makes full use of the information in the conditional distributions, its nonparametric feature nonetheless makes it less robust because it requires a large sample size to perform well. In the sparse area of the empirical distribution of  $(y_{i,t}, z_{i,t}, s_{i,t}, l_i)$ , the semiparametric conditional mean can be volatile, while the optimal linear measure remains robust. In the area where the density  $f(y_{i,t}, z_{i,t}, s_{i,t}, l_i)$

takes a relatively larger value, the conditional mean is actually stable and also close to the optimal linear measure.

Graph (a) of Figure 10 contrasts the difference between the semiparametric conditional mean and the optimal linear measure ( $E[y_{i,t}^* | y_{i,t}, z_{i,t}, s_{i,t}, l_i] - \hat{y}^*$ ) against the empirical kernel density estimates of  $f(y, z, s, l)$  in the simulated data. Notice that the the difference between the semiparametric conditional mean and the optimal linear measure decreases as the empirical density increases. Meanwhile, we actually observe the true value  $y_{i,t}^*$  in the simulation. Graph (b) and (c) show scatter plots of the optimal linear measure and the semiparametric conditional mean against the true values, respectively. Our new measures perform very well in terms of predicting the true value  $y_{i,t}^*$ . The mean squared error for the optimal linear measure is slightly smaller but very close to that of the semiparametric conditional mean in the simulation.

For these reasons, we make the optimal linear measure our choice of the new measure for the true GDP.

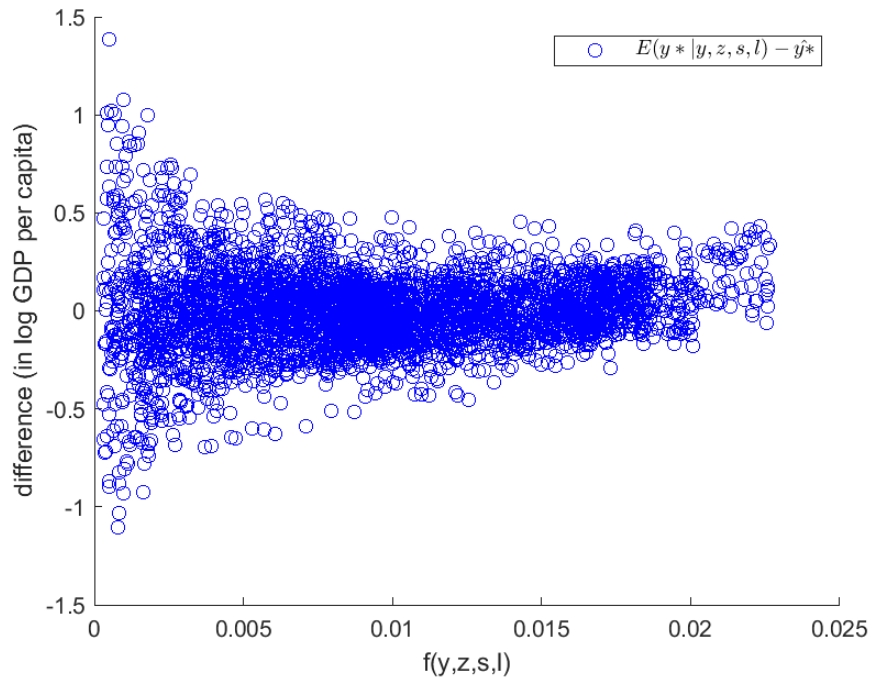
## B.2. Alternative Specifications

We consider a number of alternative specifications of the model.

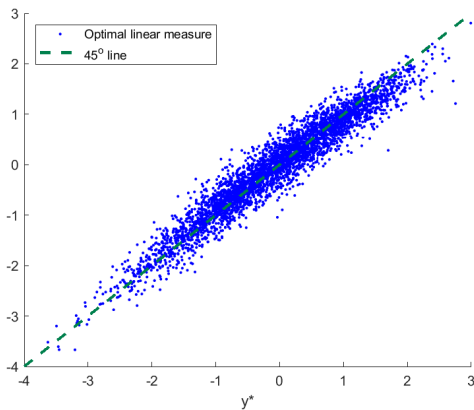
First, we consider cases where the nighttime lights production function  $m(\cdot)$  could be linear or cubic. As suggested by Table 3, the quadratic term is statistically significantly different from zero. However, if we impose that the production function is linear, the estimation results in high income countries having the highest measurement errors. Figure 11 shows that under the linear specification, deviations from the nighttime lights production function are wrongly attributed to measurement errors in official GDP. While the linear specification is mis-specified, the cubic term is not significantly different from zero and it results in unstable performance of the model.

Second, we consider more parametric specification of the error terms in equations (1) and (2). In particular, we assume  $\varepsilon_{i,t}^y(s_i) = \sigma(s)\varepsilon_{i,t}^y$  and  $\varepsilon_{i,t}^z(l_i) = \sigma(l)\varepsilon_{i,t}^z$ . In other words, the distribution of measurement errors is the same up to a change of variance for all three groups of countries with different statistical capacity, and similarly for measurement errors in nighttime lights. This specification has the advantage of having less parameters to estimate. Using the same number of sieve terms  $k = 6$  for the error terms as in our baseline specification (1) and (2), we obtain similar results, as shown in Figure 12 and Table 9. However, we find this specification to be less robust to the choice of the number of sieve terms despite its parsimonious specification. Thus we choose the distribution of measurement errors to be different for countries with different statistical capacity and geographical location groups.

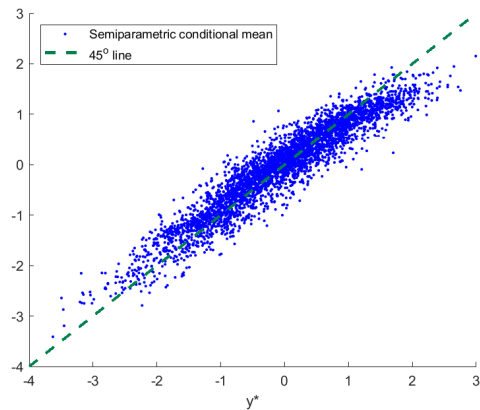
**Figure 10. Semiparametric Conditional Mean and Optimal Linear Measure Performance in Simulations**



**(a) Difference between New Measures**



**(b) Optimal Linear Measure vs. True Values**



**(c) Semiparametric Conditional Mean vs. True Values**

Third, we consider a full parametric specification where the nighttime lights production function is quadratic and the true GDP per capita as well as all measurement errors are normally distributed. Despite being computationally lightweight, the full parametric specification is not very robust to outliers. Using the full DMSP/OLS sample, we obtain the counterintuitive and incorrect results that high income economies have the highest measurement error, as shown in Table 10. However, if we remove Special Administrative Regions such as Hong Kong SAR

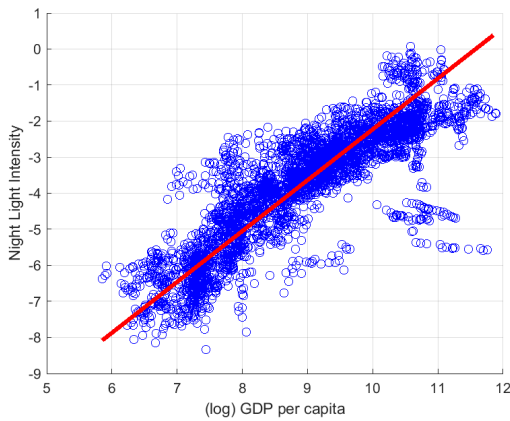
**Table 9. Estimated Light Production Function (DMSP/OLS) with Parsimonious Error Structure**

$$m(y^*) = \theta_0 + \theta_1 y^* + \theta_2 (y^*)^2$$

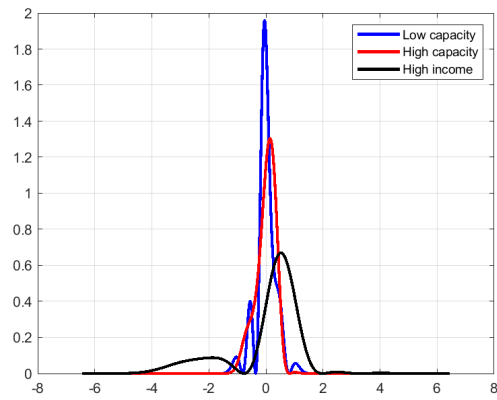
Parameter	$\theta_0$	$\theta_1$	$\theta_2$
Point Estimate	0.326	1.261	-0.212
Standard Error	(0.160)	(0.090)	(0.032)

Data are re-centered at zero. Standard errors are based on 400 sample bootstraps.

**Figure 11. Linear Specification**

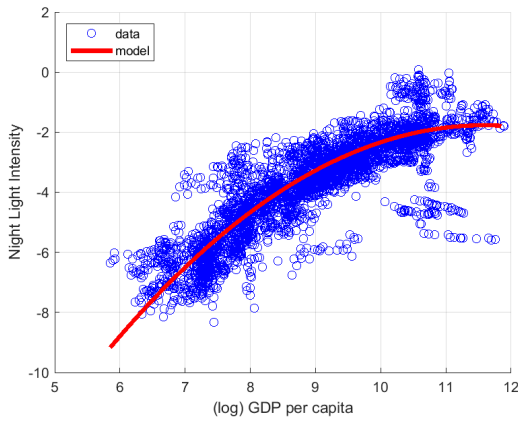


**(a) Model vs. Data**

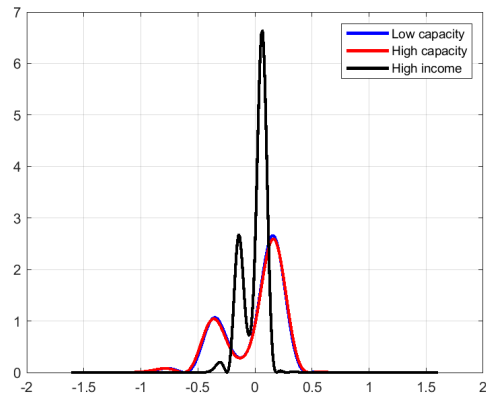


**(b) Distribution of Measurement Error**

**Figure 12. Quadratic Specification with More Parsimonious Error Structure**



**(a) Model vs. Data**



**(b) Distribution of Measurement Error**



and Macau SAR, and city states such as Singapore, where nighttime lights per capita was unusually low as a result of high population density, we obtain the more intuitive results that the variance of measurement errors is much smaller for high income countries. In light of this sensitivity, we prefer the semiparametric specification.

**Table 10. Full Parametric Specification (DMSP/OLS)**

$$m(y^*) = \theta_0 + \theta_1 y^* + \theta_2 (y^*)^2$$

Parameter	$\theta_0$	$\theta_1$	$\theta_2$	$\sigma_1^y$ (low capacity)	$\sigma_2^y$ (high capacity)	$\sigma_3^y$ (high income)
All Data						
Point Estimate	0.336	1.352	-0.256	0.458	0.474	0.791
Standard Error	(0.047)	(0.034)	(0.041)	(0.025)	(0.020)	(0.056)
Excluding Hong Kong SAR, Macau SAR, and Singapore						
Point Estimate	0.316	1.354	-0.268	0.458	0.471	0.050
Standard Error	(0.012)	(0.015)	(0.008)	(0.019)	(0.015)	(0.006)

Data are re-centered at zero. Standard errors are based on 400 sample bootstraps.

Finally, we consider combining the DMSP/OLS and VIIRS datasets by assuming that the measurement errors in official real GDP per capita follow the same distribution in the two datasets. We find that the estimated production functions are similar to our baseline results and there is little change in our new measures of real GDP per capita.

### B.3. Gas Flaring in the Data

While nighttime lights primarily reflect economic activities for a majority of countries, it is recognized in the literature (for example, [Henderson, Storeygard, and Weil \(2012\)](#)) that the flaring of natural gas might make nighttime lights incommensurate with the level of economic development. To examine the extent to which gas flares affect our results, we use gas flare shapefiles (polygons) provided by NOAA<sup>16</sup> and calculate the fraction of nighttime lights within gas flare shapefiles in a country.

As an example, Figure 13 shows a map of Nigeria where the white area is the shapefile that contains gas flares and the green area is the rest of the country. We obtain the fraction of nighttime lights in gas flare shapefiles (denoted by  $\tau$ ) by aggregating the nighttime lights within the white area first and then divide the sum by the total sum of lights in Nigeria.

<sup>16</sup>[https://ngdc.noaa.gov/eog/interest/gas\\_flares\\_countries\\_shapefiles.html](https://ngdc.noaa.gov/eog/interest/gas_flares_countries_shapefiles.html)

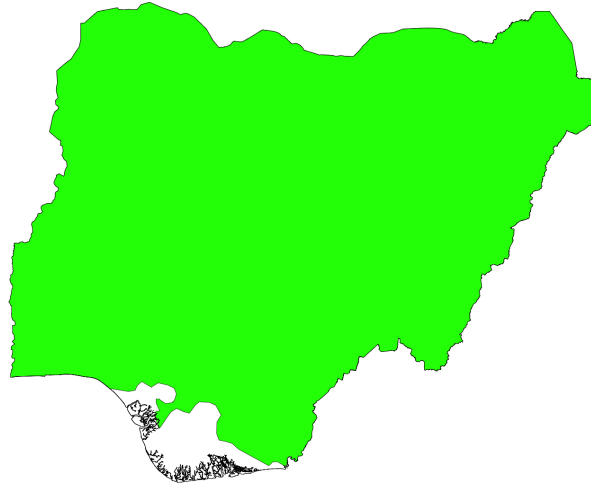
**Figure 13. Gas Flares in Nigeria**

Table 11 ranks countries by the fraction of nighttime lights in gas flare shapefiles in descending order. Among the highest are Equatorial Guinea, Gabon, and Nigeria. Most of the countries with high ranks are African and Middle East oil producers. While China and the United States produce oil, the vast majority of nighttime lights were produced in areas outside of the gas flare shapefiles. In particular, nighttime lights within gas flare shapefiles account for about 1 percent of total nighttime lights in China between 1992-2013 and even less for the United States. When dropping countries with  $\tau \geq 0.02$  (34 countries) and reestimating our model, we find no statistically significant difference in estimated coefficients of the nighttime light production function compared to the full sample.

#### **B.4. Optimal Linear Measure for More Countries**

Given the high uncertainty about official data in low income countries, we present more results comparing the optimal linear measure of real GDP per capita with official data for a number of low income countries in Figure 14. Note that the optimal linear measure being close to official data for some low income countries does not necessarily mean the official data are accurate. More often than not, it means nighttime lights for those countries do not contain much information either as a result of lack of electricity. As such the optimal linear measure puts low weight on nighttime light-predicted GDP. Nevertheless, these estimates show that our proposed optimal linear measure performs very reliably across countries.

**Table 11. Fraction of Nighttime Lights in Gas Flare Shapefiles**

ISO code	$\tau$	ISO code	$\tau$	ISO code	$\tau$	ISO code	$\tau$
GNQ	0.75	RUS	0.17	VEN	0.08	AUS	0.01
GAB	0.69	AGO	0.16	ARE	0.06	CHN	0.01
NGA	0.57	SYR	0.15	SAU	0.05	CMR	0.01
LBY	0.36	ECU	0.15	EGY	0.04	MYS	0.00
IRQ	0.32	TKM	0.15	ARG	0.04	USA	0.00
KWT	0.31	TCD	0.14	SDN	0.03	ZAF	0.00
OMN	0.31	BOL	0.13	COL	0.03	PHL	0.00
DZA	0.30	PNG	0.13	CHL	0.03	BRA	0.00
COG	0.26	COD	0.11	PER	0.02	NOR	0.00
QAT	0.26	IRN	0.10	TUN	0.02	TTO	0.00
KAZ	0.24	UZB	0.09	CAN	0.01	...	...
YEM	0.20	IDN	0.09	AZE	0.01	...	...

Notes.  $\tau$  is the ratio of the sum of nighttime lights in gas flare shapefiles to that in the whole country averaged between 1992-2013.

### B.5. Semiparametric Conditional Mean for Robustness Check

In section III.C.2, we discussed the use of the semiparametric conditional mean as another new measure of real GDP per capita. Despite its clear advantage of making full use of the information in the conditional distributions, its nonparametric feature makes it less robust without a large sample size.

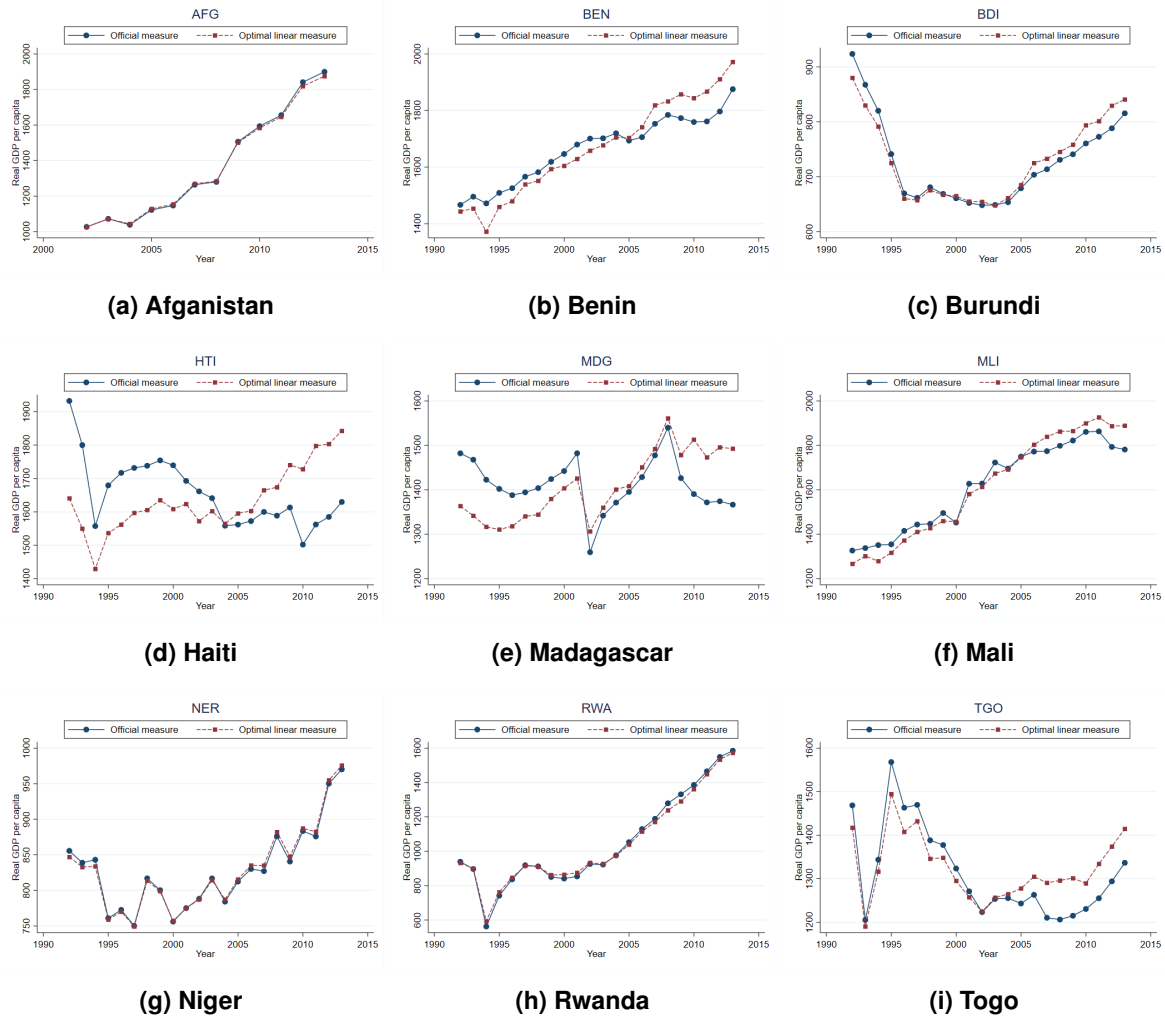
Figure 15 plots the difference between the semiparametric conditional mean and the optimal linear measure ( $E[y_{i,t}^* | y_{i,t}, z_{i,t}, s_{i,t}, l_i] - \hat{y}^*$ ) against the empirical kernel density estimates of  $f(y, z, s, l)$  in DMSP/OLS data. Consistent with the simulation results in Section B.1, as the empirical density increases, the difference between the semiparametric conditional mean and the optimal linear measure decreases. Note that the difference is almost zero for high income countries because their measurement errors that are identifiable by our method are very small.

## APPENDIX C. MORE DESCRIPTIVE DETAILS ON DATA

### C.1. Nighttime Lights vs. Real GDP per capita

Figure 16 presents where countries are located on nighttime lights and real GDP per capita graph. We choose selectively a few countries that together span the real GDP per capita spectrum.

Figure 14. New Measures for Low Income Countries



## C.2. Statistical Capacity and Latitude

We use the World Bank's Statistical Capacity Indicator<sup>17</sup> for developing countries as our variable  $s$  for statistical capacity. The Statistical Capacity Indicator is a composite score assessing the capacity of a country's statistical system, including the following areas: methodology; data sources; and periodicity and timeliness. The scores are based on 25 criteria in these areas and the overall Statistical Capacity score is then being calculated as simple average of all three area scores on a scale of 0-100.

Because the indicator starts in 2004 and the change over time for each country is small, we use the average score during our sample periods for each country. Since high income countries don't have scores, we assign them the highest score and treat them as a separate group.

<sup>17</sup><http://datatopics.worldbank.org/statisticalcapacity/>

**Figure 15. Semiparametric Conditional Mean and Optimal Linear Measure**

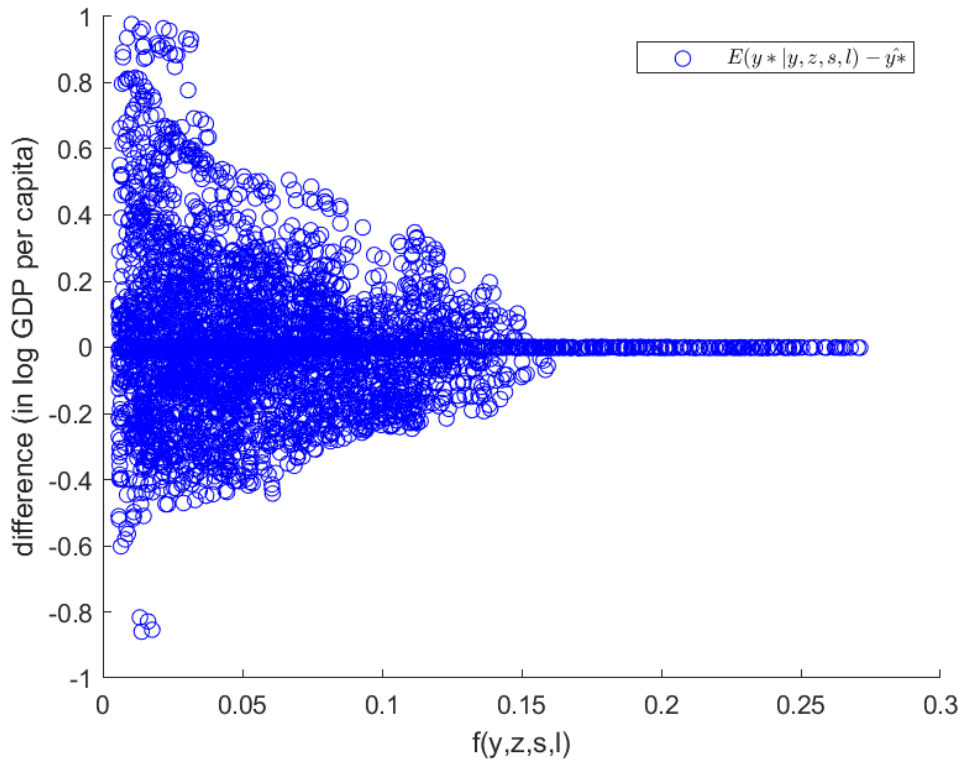
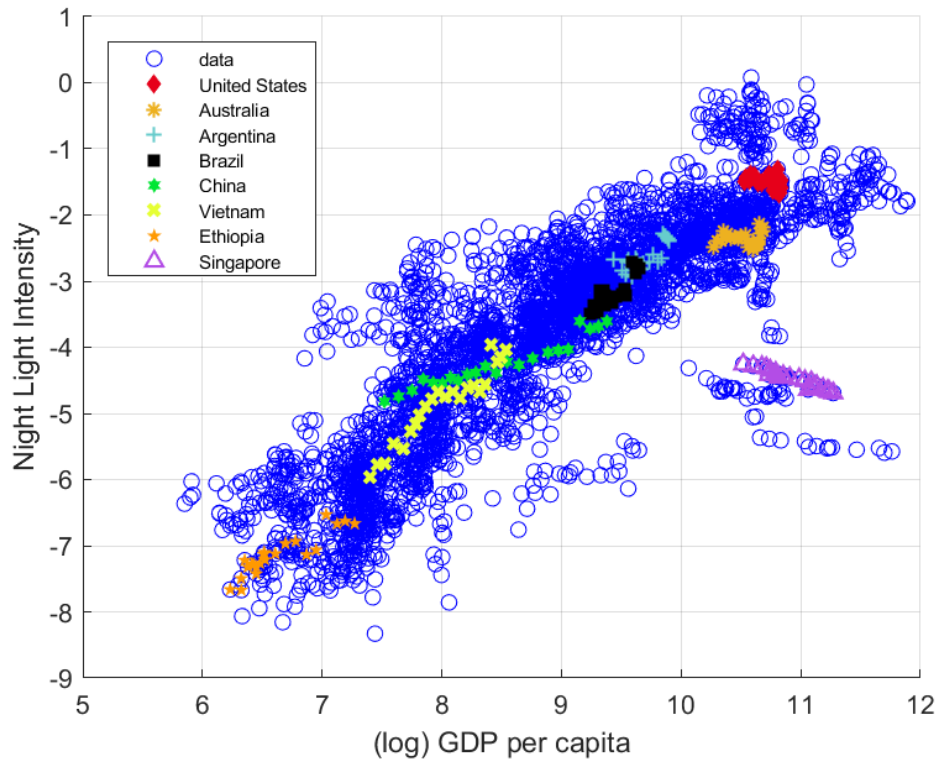


Figure 17 shows that there is a clear positive relationship between real GDP per capita and statistical capacity. Note that high income countries are located on the top right corner where statistical capacity is 100 by construction.

A country's latitude is calculated as the centroid of its largest contiguous block. We focus on contiguity because overseas territories or separate land blocks would complicate the definition of the geographic center of a country. As an example, Figure 18 illustrates the centroids that we use for a few European countries. Figure 19 shows that real GDP per capita varies with latitude. Countries at high latitude in both the Northern and the Southern Hemispheres tend to be rich whereas countries at low latitude tend to be poor. This is the well-known North-South Divide.

Collectively, Figure 17 and 19 indicate that Assumption 5 is fairly reasonable, i.e., the distribution of real GDP varies with countries' statistical capacity and geographic location.

**Figure 16. Nighttime Lights vs. Real GDP per capita (1992-2013)**

**Figure 17. Real GDP per capita vs. Statistical Capacity by Location  
(1992-2013)**

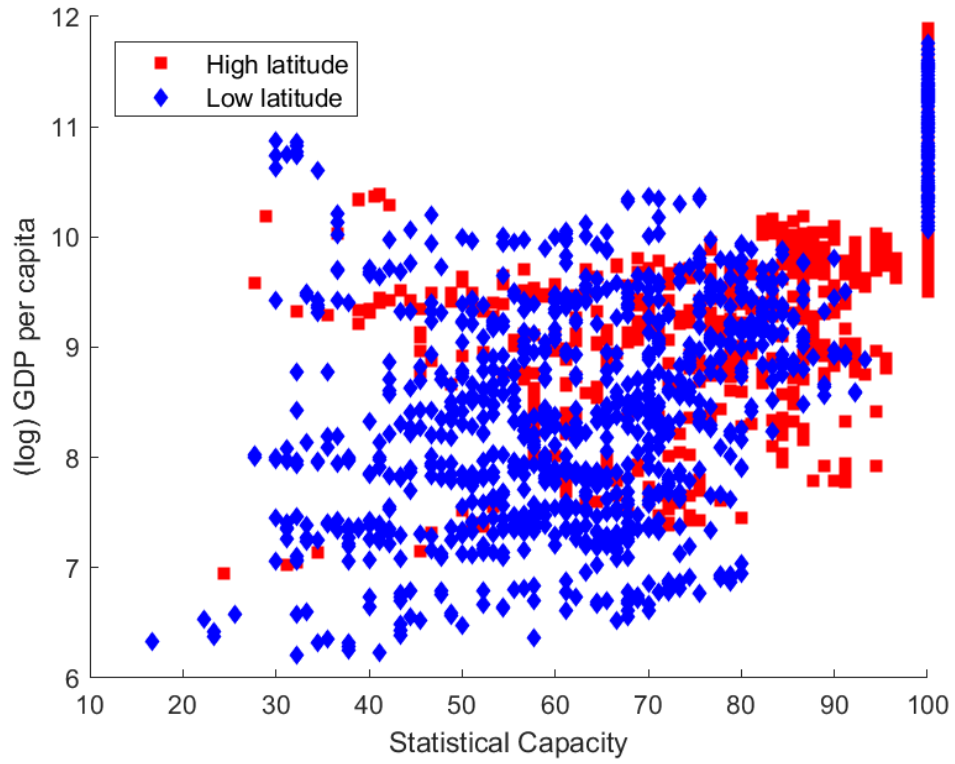
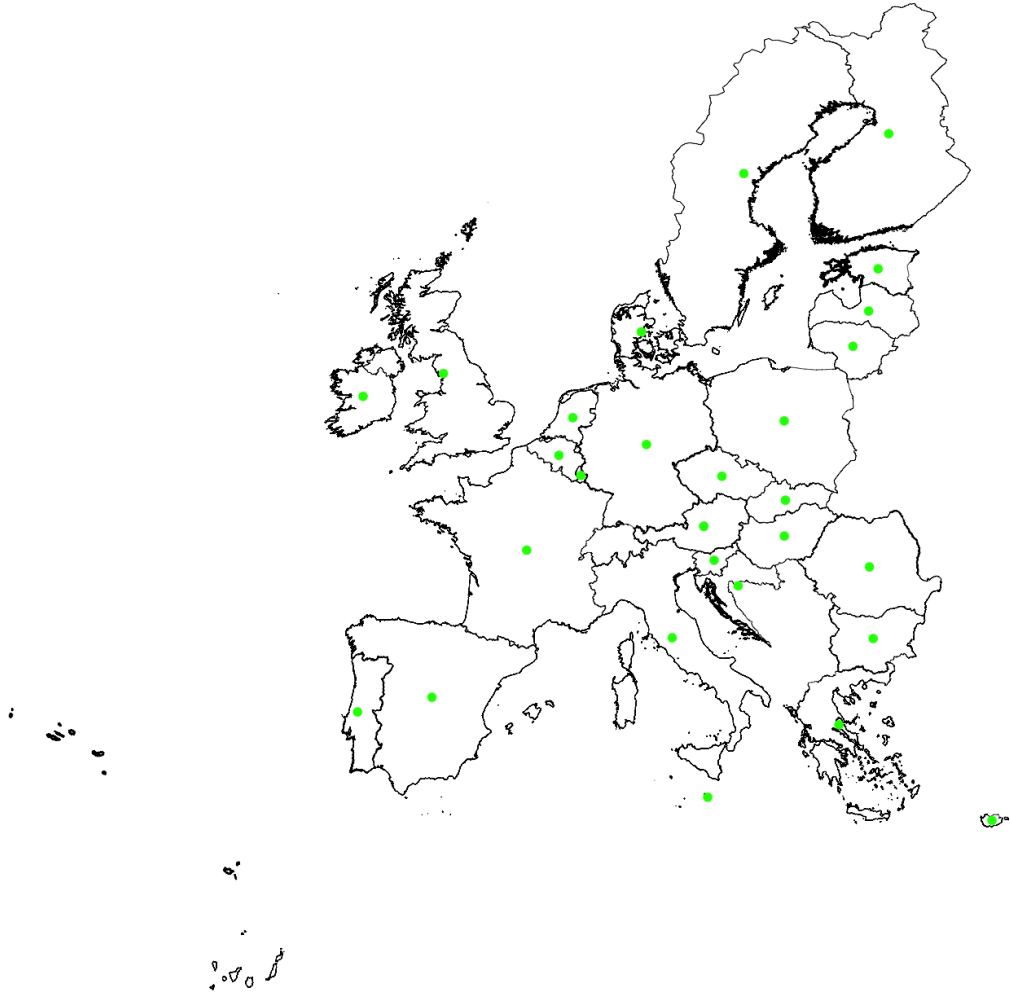


Figure 18. Centroids of Selected European Countries





**Figure 19. Real GDP per capita vs. Latitude by Statistical Capacity (1992-2013)**

