An alternative method to measure non-registered economic activity in Mexico using satellite nightlights.

Erick Rangel-Gonzalez[a] and Irving Llamosas-Rosas[b]

October 2019

Abstract

We use satellite nightlights to measure economic activity in México and discrepancies between estimated an official GDP in order to identify the non-registered economy. Our work is related to Ghosh, et al. (2009) that uses measures of nightlights in Mexico to estimate informal economic activity but with significant differences: i) we use an updated version of nighttime light data based on VIIRS satellite, which offers a higher resolution; ii) we use the State of Nuevo Leon (that has the lower levels of informality in the country) as a benchmark economy besides the US economy; iii) we also incorporate into the analysis the contribution of different sectors to the state's GDP, since different economic activity will have different nightlight elasticities (i.e. agricultural activities, services and manufacturing). Our estimates indicate that the size of the non-registered economy in Mexico is about 25.25% of the GDP when using US economy as reference, and 29.03% when using the state of Nuevo Leon as a benchmark economy.

Keywords: nightlights, non-registered GDP, satellite, shadow economy, informality.

Work in progress

a. *Dirección General de Investigación Económica, Banco de México. email: erick.rangel@banxico.org.mx.*
b. *Dirección General de Investigación Económica, Banco de México. email: illamosasr@banxico.org.mx.*

I. Introduction

There are many economic activities that are not registered in the Gross Domestic Product (GDP). For example, small business or shops that are founded in the streets that produce or sell goods and services to the population. A large number of these economic agents are not likely neither to be registered formally with the government nor to pay taxes. It is also possible that some registered establishments report a lower level of production and sales, and therefore there is part of the economic activity that is not measured by the official estimates. The non-registered economic activities, also known as shadow economy, are by nature difficult to measure, because the agents in this sector try to remain undetected for monetary, regulatory or institutional issues.

Some examples of monetary reasons is to avoid paying taxes and social security contributions. Regulatory reasons include to evade administrative transactions and permits of operation, government bureaucracy, or the burden of regulatory framework. Institutional reasons consider factors such as corruption, quality of political or governmental institutions and weak enforcement of law.

The estimation of the total economic activity, including official and non-registered production of goods and services, is important to determine the welfare level of the population in one country, it is also essential for designing economic policies to encourage development, or policies that respond to changes in the economic cycle. In addition, an estimation of the shadow economy is a fundamental input to estimate tax evasion and for design policies aimed for combating this problem.

There are several forms of measuring shadow economy which can be categorized in direct and indirect approaches (Medina and Schneider, 2018). The first category includes: i) estimates through discrepancy method that uses the System of National Accounts Statistics; ii) representative surveys for population that usually focus on labor informality; iii) surveys of company managers that try to capture misreported business income and misreported wages as percentage of GDP; iv) and the estimation of the consumption-income-gap of households. The second category includes: i) discrepancy between expenditure and income statistics; ii) difference between official and actual labor force; iii) electricity approach; iv) transaction approach; v) currency demand approach; and Multiple Indicator, Multiple Causes (MIMIC) approach.[1]

[1] A detailed explanation of these approaches can be found in Medina and Schneider (2018).

Previous works for Mexico have used different methods in order to obtain a measure of this unobserved economy, the MIMIC method has larger estimates compared against National Accounts Methods (Medina and Scheider, 2017). According to the last one, INEGI reports for México in 2017 an amount of informal economy of 22.7 percent of the economy, while MIMIC has an estimate of 31.7%. Monetary demand methods has estimations that goes from 24.46% for the period 1996-2000 (Chapa et al. 2008), and 22.95 % for 2012 (Santos-Pérez, 2016).

Other studies like Gyomai and van de Ven (2014) estimates that the size of underground economy in Mexico is 5.5%, and 10.4% for the Informal sector, for a total of non-observed economy of 15.9% (this calculation does not includes illegal activities). Using the predictive mean matching method between 1991 and 2015, they also found that in México the shadow economy is 24.8 percent (more in line with INEGI national accounts approach).

In this research we use an alternative recently developed approach that employs nightlight data observed from space to measure non-registered economic activity, or shadow economy, in Mexico. Nightlight data represent an attractive path to analyze the shadow economy because luminosity detected from the space captures all kind of economic activity: formal and informal. The idea is to use satellite images of nightlights, through which economic activity can be captured, to measure economic activities not recorded in GDP by the differential between the estimated output by brightness and the official data. This idea had already been exploited Ghosh et al. (2002), who used the U.S. economy as a reference, as it has very low levels of informality.

This research has several differences with respect to Ghosh et al. (2002): i) we use GDP instead of Gross National Income as a measure of economic activity, which we believe is more appropriate for the analysis since we are using luminosity inside the country ; (ii) we use a different econometric specification to estimate the correlation between luminosity and economic activity, our model accounts for the fact that the relationship between nightlights and economic activity differs by sector; iii) we also present estimations using a different reference economy, the state of Nuevo Leon that it is recognized as the federal entity with the lowest levels of informality in Mexico; iv) we use more recent data from the NASA VIIRS satellites, which have higher resolution on luminosity and has similar or even better results for estimating economic statistics than DMSP lights used in previous works (Chen and Nordhaus, 2019).

The main findings can be summarized as follows. The size of the non-registered economy in Mexico is about 25.25 % of the GDP when using the US economy as a reference, and 29.03% when using the state of Nuevo Leon as a benchmark economy. State level estimates should be taken with caution since the econometric model does not allow to considerate too many specific characteristics, other than mixture of primary, secondary and tertiary sector within every federal entity. However, they can be used as a starting point for ranking the size of the shadow economy.

The rest of the paper is structured as follows. Section II contains a literature review of papers using nightlights data for measuring economic activity. A description of luminosity data and economic data are presented is sections III and IV. The econometric methodology is explained in section V, while the results can be found in section VI. Finally, section VII concludes.

II. Literature on nighttime lights with economic activity

Donaldson and Storeygard (2016) provide a comprehensive outline of the use of satellite data in different areas of economics, including growth and development, regional, urban and environmental economics. Another research that provide an excellent review of possible applications in economics and other areas is Daren et al. (2016).

The first paper to analyze the relationship between nighttime lights and economic activity was done by Elvidge et al. (1997). They study the connection of luminosity with population, GDP, and electric power consumption for twenty-one countries, during the period 1994-1995, at different levels of economic development. Later, Doll et al. (2000) construct a GDP global map using nighttime satellite data and purchasing power parity (PPP), the map had a one degree spatial resolution. Other works related with mapping economic activity via luminosity observed from space are Sutton and Constanza (2002) and Doll et al. (2006).

Ebener et al. (2005) analyzed correlation between nighttime lights and per capita GDP at the national and sub-national level. They used different parameters to test the relationship and found that the total and mean frequency of lights to be better correlated with per-capita GDP than lit area. Their analysis also controlled for agricultural level and resulted in a good fit for GDP at the national level but did not provide consistent estimates at the sub-national level. Later Keola et al. (2015) demonstrated that nightlights alone might not generate reliable estimates of value added in the agricultural sector. On the topic of economic activity at the subnational level, Sutton et al. (2007) use the stable nighttime lights data of 1992-93 and 2000 and add population data to the model in order to obtain better estimations of sub-national GDP for India, China, Turkey and the United States.

Introducing a very comprehensive framework in economics, Henderson et al. (2012) used a panel data of GDP and the night light intensity between 1992-93 and 2002-03 to estimate income growth, and GDP at the national and subnational level. They use their proposed model to estimate the income growth for bad data low to middle-income countries; to compare the income growth on the coast versus the interior in sub-Saharan Africa; to estimate the growth prime cities versus hinterland and to measure the effects of the malaria on growth in sub-Saharan Africa. Using a similar framework Chen and Nordhaus (2011) proposed the use of satellite data to improve the quality of GDP mainly in countries with low quality statistical systems or with not current data. LLamosas et al. (2018) applied the Henderson´s et al. model to estimate economics growth in the main tourist beach destines in Mexico for the period for the period 1993-2017 by using panel state data.

Using alternative approaches, Hu and Yao (2019) employ recent measurement error models to identify the nonlinear relationship between nightlights from the space and per capita economic growth. Guerrero and Mendoza (2019) present a statistical model to estimate economic growth using satellite data by using a single country approach, the authors presents estimations for Mexico, China and Chile.

The nighttime lights data, in addition to population data have also been used to produce a poverty map (Elvidge et al. 2009). The map is useful to identify the areas of poverty in the world, where economic activity is limited. Another similar application is the construction of a Night Light Development Index (NLDI) to evaluate unbalanced development at the national and regional level.

All the previous works used data from the U.S. Air Force Defense Meteoroidal Satellite Program (DMSP) Operational Linescan System (OLS) that provides data from 1992 to 2013. Since 2011 another satellite system was deployed by the NASA and the NOAA; the Suomi National Polar Partnership (SNPP) with the Visible Infrared Radiometer Suite (VIIRS). Chen and Nordhaus (2019) compare both databases for the United Sates and conclude, among other things, that VIIRS nightlight data provide equal or better cross-sectional GDP estimations than DMSP-OLS.[2]

Beyer et al. (2018) use VIIRS data to present a measure of monthly economic activity at district level using a spatial approach to allocate GDP by dividing agricultural and nonagricultural activities.

Focusing on the topic of informality and satellite imagery, Ghosh et al. (2009) use regression models using nightlights and adjusted official gross state product for the USA, and use the calculated

[2] In the next section, we explain the database's details.

parameters to obtain Estimated Gross Domestic Income (EGDI) at the national level and at the state level. According to the authors, the difference between the EGDI and the official gross national income is an estimate of the informal economy and the inflow of remittances. The same exercise is also applied for the Indian economy by Gosh et al (2010). Harati and Hardy (2013) argue that since economic activity predicted by nighttime lights must capture both formal and informal activities the difference between light predicted and official GDP might be used to estimate informal activity. Brock et al. (2014) analyze the change in luminosity and gross country product between 2000 and 2006 for the state of Veracruz (Mexico) to identify the evolution of the formal and informal activity. Finally, Tanaka and Keola (2015) analyze the shadow economy in Cambodia by estimating the relationship between nightlights and sales across regions separately for both formal and informal firms using a question of the economic census that allows them to identify registration with administrative agencies. Their results indicate an increase in absolute and relative terms of the informal sector.

III. What is nightlight satellite data?

Since 1992, satellites' images have been systematically digitized at the NOAA National Geophysical Data Center (NGDC). After 1992, the NASA and the NOAA via the Suomi National Polar Partnership (SNPP) with the VIIRSS released data nightlight. Donaldson and Storeygard (2016) group the main advantages of such remote sensing data into three categories: 1) access to information difficult to obtain by other means; 2) unusually high spatial resolution; and 3) wide geographic coverage.

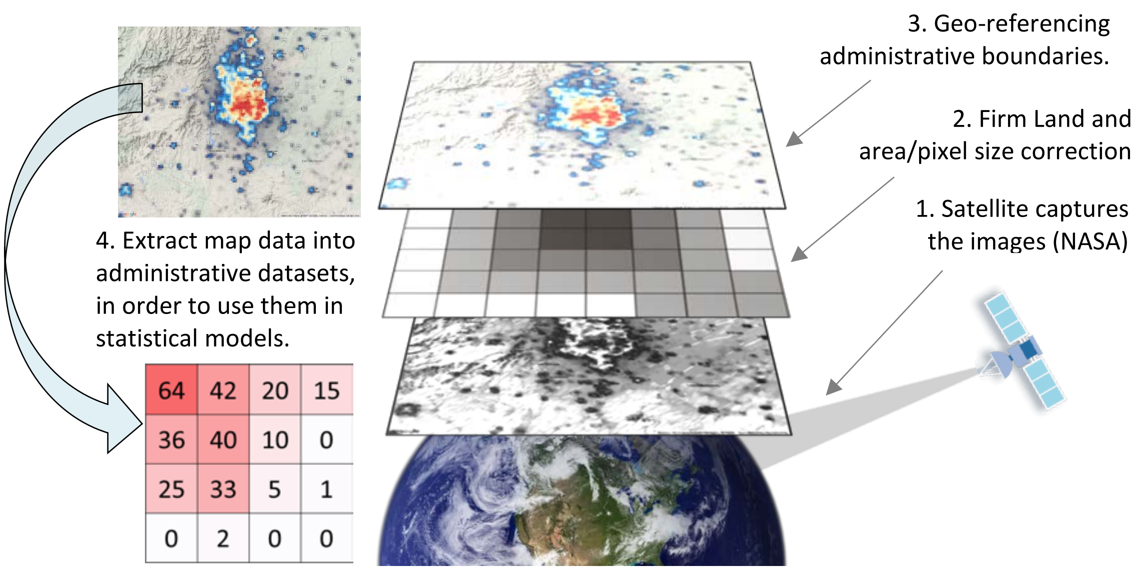
The NASA-NOAA data is available per section of the planet, the one corresponding to Mexico and the United States is Region 1 (North America), and each file (one month of information) is 2 gigabytes in size, mainly due to the largest accuracy of the satellites used (approximately 4 times more data per pixel than the DMSP-OLS).

The data observation unit is the pixel that corresponds to a geographic area equivalent to 0.44 square kilometers (15 arc-seconds grids). Within each pixel the brightness is measured on a scale from 0 to 4536.23 for the 2015 raster3. Subsequently, the data are reported according to a system of angular geographical coordinates (latitude and longitude) and paired with USA and Mexico States' administrative boundaries through GIS software. Figure 1 summarizes the processing of the data. We use the stable annual composite data produced by the Earth Observation Group (EOG) and the

[3] The DMSP-OLS pixels takes values on luminosity ranging from 0 to 63.

NOAA, since it is more stable than monthly data (Elvidge et al, 2017), this data is only available for 2015 and 2016 at present time.

Figure 1. Graphical Description of Data Processing



The satellite data was processed following the methodology proposed by Henderson, Storeygard and Weil (2012) to adjust the change in pixel size due to the curvature of the Planet, as well as take into account the reflections of the lights in water fields and Coastal.

It should be noted that the adjustment takes into account the rivers and lakes, by eliminating the pixels that are entirely within these spaces even if they register the presence of light. In addition, a correction is made to cases where the pixel is at the edges of the coast (for example, that part of the pixel is on the mainland and the other part in a lake, sea or river) multiplying the light recorded in that pixel by the area in kilometers stable two of the mainland registered in the Gridded Population of the World version 3 (GPWv3), which records the areas of the mainland in square kilometers for each pixel at a resolution of 15 arc-seconds.

The final adjustment is made by incorporating the maps of the administrative units (entity, municipality, any ad-hoc sub-region) to the adjustment of light and square kilometers of the mainland, adding all the pixels within the same unit.

We have to remark the differences between DMSP-OLS and Suomi NPP VIIRS satellite data, since the units measured by the former are relative, between 0 and 63 (64 levels in total) and radiances (nano Watts/cm2sr) (Bennett and Smith, 2017). This detail is important for the accuracy of the model, since several places (mainly cities) report the maximum value in the DMSP-OLS data, and VIIRS report the actual radiance.

IV. Economic Data

For the economic activity in the USA, we took the real GDP by state, in millions of chained 2012 dollars from the Bureau of Economic Activity, seasonally adjusted. Since different types of activities may have different impact in lights, we divided the economic activity in primary (agriculture, forestry, fishing and hunting), secondary (mining, quarrying, and oil and gas extraction, utilities, construction, manufacturing) and tertiary (rest) based on the North American Industry Classification System. In the case of Mexico, we have the real GDP by state from the National Institute of Statistics, Geography and Informatics (INEGI) divided by sector.[4]

For example, in the case of United States in Table 1 we can compare the data from California and Texas, and see that the former has a lower averaged sum of night time lights than the latter (211,909,232 vs 349,771,392) although the GDP is higher in California. This might be explained because there are different sector participations in each state, and the relationship between lights and GDP must be differentiated between sectors. In other words, the luminosity of tertiary sector, which is higher in California, has a higher impact on GDP than the influence of secondary sector, which is higher in Texas.

Table 1. Economic and Satellite Data for USA. 2015-2016 averages.

State	Sum of Night Time Lights (average 2015-2016)	GDP (in millions 2012 chained dollars, average 2015-2016)	Primary Sector participation	Secondary Sector participation	Tertiary Sector participation	Population (millions)
Alabama	66,004,088	189,165	2%	25%	74%	4.859
Arizona	62,037,312	284,059	1%	16%	83%	6.890
Arkansas	40,328,512	113,483	3%	23%	74%	2.984
California	211,909,232	2,462,491	2%	16%	82%	39.081
Colorado	48,578,768	311,668	1%	19%	80%	5.497
Connecticut	19,452,612	241,681	0%	15%	84%	3.583
Delaware	8,930,080	64,677	1%	11%	88%	0.945
District of Columbia	5,925,040	118,197	0%	2%	98%	0.681
Florida	172,619,936	849,979	1%	11%	88%	20.427
Georgia	110,816,000	487,234	1%	16%	83%	10.243

Continue next page…

[4] The sectors included in each of the primary, secondary and tertiary economic activities are the same across USA and México.

Table 1. (Cont.) Economic and Satellite Data for USA. 2015-2016 averages.

State	Sum of Night Time Lights (average 2015-2016)	GDP (in millions 2012 chained dollars, average 2015-2016)	Primary Sector participation	Secondary Sector participation	Tertiary Sector participation	Population (millions)
Idaho	12,719,692	64,387	7%	19%	73%	1.667
Illinois	144,613,184	742,771	1%	18%	81%	12.846
Indiana	74,937,680	312,504	2%	33%	65%	6.621
Iowa	40,851,216	168,581	7%	23%	70%	3.127
Kansas	34,011,848	146,590	4%	22%	74%	2.910
Kentucky	51,102,112	181,147	2%	25%	73%	4.432
Louisiana	77,966,992	229,926	1%	34%	65%	4.672
Maine	8,746,300	54,157	2%	14%	84%	1.330
Maryland	44,304,676	349,765	0%	11%	88%	5.996
Massachusetts	38,094,120	473,814	0%	15%	85%	6.811
Michigan	83,864,816	444,512	1%	24%	75%	9.942
Minnesota	58,227,464	314,477	2%	20%	77%	5.503
Mississippi	45,945,024	99,950	2%	23%	74%	2.988
Missouri	70,628,320	274,513	2%	18%	81%	6.079
Montana	11,673,602	44,469	6%	20%	74%	1.036
Nebraska	26,867,752	109,949	10%	15%	75%	1.899
Nevada	34,256,360	136,692	0%	12%	87%	2.894
New Hampshire	8,000,602	71,859	0%	16%	84%	1.339
New Jersey	56,717,524	535,382	0%	14%	86%	8.873
New Mexico	26,211,516	90,919	2%	22%	76%	2.091
New York	86,880,992	1,378,682	0%	9%	91%	19.652
North Carolina	98,659,328	470,342	1%	23%	75%	10.095
North Dakota	33,517,268	52,682	7%	28%	65%	0.754
Ohio	118,958,800	579,436	1%	24%	75%	11.626
Oklahoma	57,105,128	191,843	2%	35%	63%	3.918
Oregon	19,988,318	195,964	3%	20%	77%	4.054
Pennsylvania	88,494,704	681,667	1%	20%	79%	12.785
Rhode Island	6,160,310	52,994	0%	14%	86%	1.057
South Carolina	57,094,036	191,816	1%	24%	76%	4.925
South Dakota	13,627,812	45,345	11%	14%	75%	0.858
Tennessee	80,414,408	303,462	1%	19%	80%	6.618
Texas	349,771,392	1,591,553	1%	32%	67%	27.712
Utah	24,360,514	144,112	1%	21%	79%	3.013
Vermont	2,520,420	29,031	2%	16%	82%	0.624
Virginia	65,570,768	454,745	0%	14%	86%	8.387
Washington	38,484,664	451,719	2%	18%	80%	7.229
West Virginia	18,276,284	69,080	1%	31%	69%	1.836
Wisconsin	60,515,224	286,843	2%	24%	74%	5.767
Wyoming	10,525,880	38,822	2%	39%	58%	0.585

Source: NASA/NOAA , VIIRS-Night Light Satellite Data and Bureau of Economic Analysis.

Table 2. Economic and Satellite Data for USA. 2015-2016 averages.

State	Sum of Night Time Lights (average 2015-2016)	GDP (in millions 2013 pesos, average 2015,2016)	Primary Sector	Secondary Sector	Tertiary Sector	Population (millions)
Aguascalientes	4,189,581	207,198	4%	42%	54%	1.298
Baja California	11,431,235	516,979	3%	38%	59%	3.519
Baja California Sur	3,259,738	132,162	4%	28%	68%	0.778
Campeche	3,636,300	619,904	1%	87%	13%	0.917
Coahuila	12,925,819	578,354	2%	52%	46%	2.983
Colima	2,091,102	98,279	5%	26%	70%	0.731
Chiapas	11,075,630	290,555	7%	25%	68%	5.301
Chihuahua	15,793,577	527,817	6%	40%	54%	3.742
Ciudad de Mexico	16,672,784	2,899,069	0%	10%	90%	8.855
Durango	5,393,448	198,636	10%	30%	60%	1.779
Guanajuato	18,372,564	675,249	3%	38%	59%	5.852
Guerrero	4,492,022	234,483	5%	18%	76%	3.587
Hidalgo	8,488,594	258,869	4%	34%	62%	2.901
Jalisco	18,901,420	1,134,842	6%	32%	63%	7.992
Mexico	35,512,024	1,459,986	2%	26%	73%	17.051
Michoacan	6,835,790	399,968	12%	18%	70%	4.622
Morelos	4,667,478	189,409	3%	30%	67%	1.935
Nayarit	1,927,376	117,285	7%	21%	72%	1.239
Nuevo Leon	16,898,776	1,229,304	1%	36%	63%	5.131
Oaxaca	6,909,841	258,460	6%	27%	67%	4.034
Puebla	13,434,698	546,259	4%	33%	63%	6.237
Queretaro	7,956,289	377,770	2%	39%	58%	2.023
Quintana Roo	4,533,842	254,445	1%	12%	87%	1.602
San Luis Potosi	6,150,814	336,324	4%	38%	58%	2.771
Sinaloa	7,750,207	371,871	12%	20%	68%	3.007
Sonora	11,989,829	552,531	6%	46%	48%	2.961
Tabasco	14,558,578	544,485	1%	67%	31%	2.399
Tamaulipas	13,782,448	490,664	3%	38%	59%	3.572
Tlaxcala	3,472,484	96,917	3%	35%	61%	1.290
Veracruz	20,121,284	803,705	5%	33%	61%	8.088
Yucatan	6,832,301	237,317	4%	26%	70%	2.136
Zacatecas	5,010,151	157,911	8%	38%	54%	1.585

Source: NASA/NOAA , VIIRS-Night Light Satellite Data and INEGI.

In order to make plausible comparisons between GDP estimations in Mexico and USA, we use the Purchased Power Parity from the OECD. This parity is the rate of currency conversion between Mexico and USA that eliminates the differences in the price levels between the countries.

V. Model

Our specification to estimate the level of non-registered economy in Mexico is to identify the *corrected* elasticities of night time light on economic activity, since we assume that some of the activity goes non-registered; and the elasticities obtained could be biased because of that. In order to do so, we must rely on another country that is believed to have the lowest non-registered economy, in our case is the United States of America.

The Night Time Light – Economic Activity elasticity is not unique for all sectors, since night time light has a higher correlation with economic activity in secondary and tertiary sectors than in agriculture (Beyer et al., 2017). For this reason, we decompose the correlation with the participation rate of each sector for each state (Equation 1).

Also, we control for population and population squared for each economy, in order to capture agglomeration effects within states.

$$\ln\left(GDP_i^{US}\right) = \alpha + \beta_1^{US}\ln\left(SL_i^{US}\right) * P_i^{US} + \beta_2^{US}\ln\left(SL_i^{US}\right) * S_i^{US} + \beta_3^{US}\ln\left(SL_i^{US}\right) * T_i^{US} + \gamma_1^{US} Pop_i^{US} + \gamma_2^{US} {Pop^{US}}_i^2 + \varepsilon_i \quad (1)$$

Here, $\ln(GDP_i)$ is the natural logarithm of real Gross Domestic Product per capita for state i, $\ln(SL_i)$ is the natural logarithm of the sum of night time lights for each state; P_i, S_i, and T_i are the participation rate of primary, secondary and tertiary sector, respectively; and Pop_i is the population in millions for each state. Equation 1 is estimated by Ordinary Least Squares[5].

Based on Ghosh (2009), once we have the US coefficients from equation 1, we "predict" the real GDP for México using its luminosity data, (Equation 2).

$$\ln\left(\widehat{GDP_i^{MX}}\right) = \alpha + \beta_1^{US}\ln\left(SL_i^{MX}\right) * P_i^{MX} + \beta_2^{US}\ln\left(SL_i^{MX}\right) * S_i^{MX} + \beta_3^{US}\ln\left(SL_i^{MX}\right) * T_i^{MX} + \gamma_1^{US} Pop_i^{MX} + \gamma_2^{US} {Pop^{MX}}_i^2 \quad (2)$$

Therefore, the data from Mexico is used with the US elasticities in order to estimate the GDP in real US dollars (since the elasticities capture the relationship between lights and economic activity measured in millions of dollars). To capture the differences between registered and estimated GDP in México, we must change registered GDP in millions of pesos to dollars by using Power Purchased Parity adjusted rate.

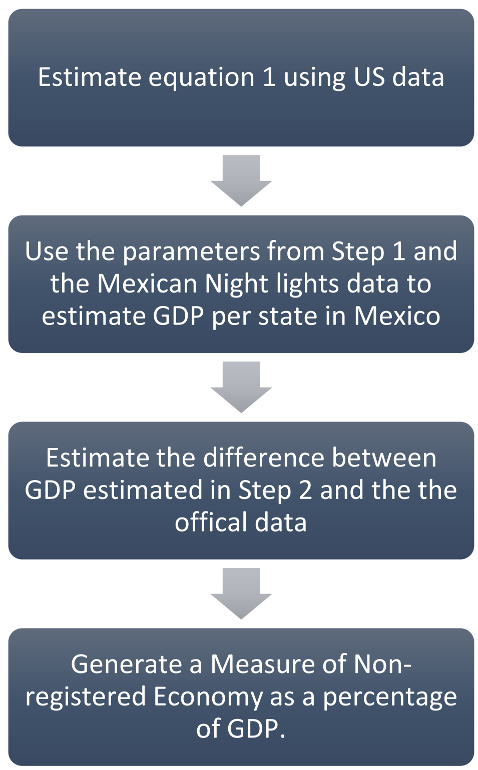
[5] We include the area of each state as a weight.

$$d_i = \widehat{GDP_i^{MX}} - \left[GDP_i^{INEGI} * PPP\left(\frac{US}{MX}\right)\right] \quad (3)$$

So, $d_i / \left[GDP_i^{INEGI} * PPP\left(\frac{US}{MX}\right)\right]$ is the percentage of the economic activity un-registered by Mexican authorities in the state i. In order to estimate the nonregistered economy at the national level, we have to add d_i for all states and divide it by the sum of all official GDPs per state.

Summarizing, the process for generate a measurement of non-registered economy, using US data, as a percentage of GDP is depicted in the following diagram.

Figure 2. Process of estimating non-registered economy using US data.



A possible criticism of the approach using USA elasticities over Mexican data rely that there should be differences between the two countries regarding the relationship between satellite captured nigh lights and economic activity, since there could be diminution in the rate of increase of lights as income rises (Henderson et al., 2012). This differences could potentially bias the results, since we are assuming a constant relationship between lights and GDP regardless the level of development of the country.

If we only have to rely on Mexican data, we must identify a state in which the registered GDP is close to the real economic activity. Gosh (2009) identify the state of Nuevo León with a positive residual between official GDP and modeled GDP, in other words, the state of Nuevo León has a higher economic activity (measured by Gosh methodology) than the official data. On top of that, official labor data from INEGI reports Nuevo León as the second state with lower labor informality, overcome only by Coahuila.

Taking Nuevo León as the benchmark state, and assuming that there is no non-registered economic activity in the state's GDP, we can identify the additional effect on GDP by the interaction of the logarithm of night light activity $\ln(SL_i^{MX})$ and a dummy variable of Nuevo León (equation 4).

$$\ln(GDP_i^{MX}) = \alpha + \beta_1^{MX} \ln(SL_i^{MX}) * P_i^{MX} + \beta_2^{MX} \ln(SL_i^{MX}) * S_i^{MX} + \beta_3^{MX} \ln(SL_i^{MX}) * T_i^{MX} + \delta \ln(SL_i^{MX}) * NL_i + \gamma_1^{MX} Pop_i^{MX} + \gamma_2^{MX} {Pop^{MX}}_i^2 + \varepsilon_i \quad (4)$$

Where NL_i is a dummy variable equal to one if the state i is Nuevo León, and zero elsewere.

Given that, the percentage effect of lights over GDP is given by:

$$\frac{\%\Delta(GDP_i^{MX})}{\%\Delta(SL_i^{MX})}_{NL=1} = \delta$$

Where δ is the marginal effect of Nuevo Leon on GDP. In other words, given that official GDP in Nuevo Leon closely captures the economic activity in the state, δ >0 will show the additional elasticity not obtained by the average state in México.

By assuming a constant elasticity between satellite night lights and GDP (only different by sector participation), the additional effect of the state of Nuevo Leon (δ) will capture the non-registered economic activity for each state.

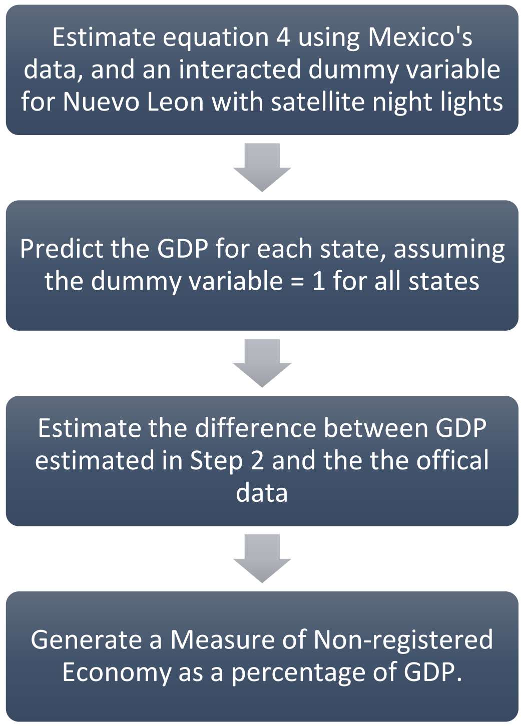
The estimated $\widehat{GDP_i^{MX}}_{NL=1}$ allowing $NL_i = 1$ for all i, gives us the GDP for each state with the measure of both registered and non-registered economic activity. Therefore, the difference between this estimation and official GDP gives us a measurement of the non-registered economic activity

$$d_i^{NL=1} = \widehat{GDP_i^{MX}}_{NL=1} - GDP_i^{MX} \quad (6)$$

So, $d_i^{NL=1}/GDP_i^{MX}$ is the percentage of the economic activity un-registered by Mexican official data. Same as before, we have to add d_i for all states and divide it by the sum of all official GDPs per state in order to estimate the nonregistered economy at the national level.

Summarizing, the process for generate a measurement of non-registered economy, using only Mexico's data, as a percentage of GDP is depicted in the following diagram.

Figure 3. Process of estimating non-registered economy using Mexico's data.



VI. Results.

VI.I Estimates using the US the Reference Economy

Table 3 presents the results of the equation 2 using state data for the US for the years 2015 and 2016. The estimations indicates, for both years, a higher elasticity of nightlights with respect to GDP for services with a coefficient that it is statistically significant at the 1 percent level. In contrast, the lowest estimated elasticity it is registered in the primary sector of the economy and it´s coefficient it is not statistically significant at any significance level.

The calculated elasticity for the manufacturing sector is in between the two other estimations and it is statistically significant at the 99 percent confidence level. Beyer et al (2018) estimated elasticities for agriculture, manufacturing and services for the world and for south Asia, their estimations for the primary sector were not statistically significant for south Asia, but they found an elasticity of 0.128 for the world regressions. In relation with the elasticities for manufacturing and services, their estimations indicated a higher elasticity for manufacturing than services when using the world data and the opposite outcome for the south Asia countries.

Table 3. Estimation Results of Equation 2, US Data, 2015 and 2016

Variables	2015		2016	
Ln(Light)*Primary Sector	0.228		0.165	
Ln(Light)*Secondary Sector	0.355	***	0.327	***
Ln(Light)*Tertiary Sector	0.465	***	0.453	***
Population	0.130	***	0.132	***
Population^2	-0.002	**	-0.002	**
Constant	3.866	*	4.189	**
Observations	49		49	
R^2	0.956		0.955	

Source: Own calculations using NASA/NOAA , VIIRS-Night Light Satellite Data and INEGI.

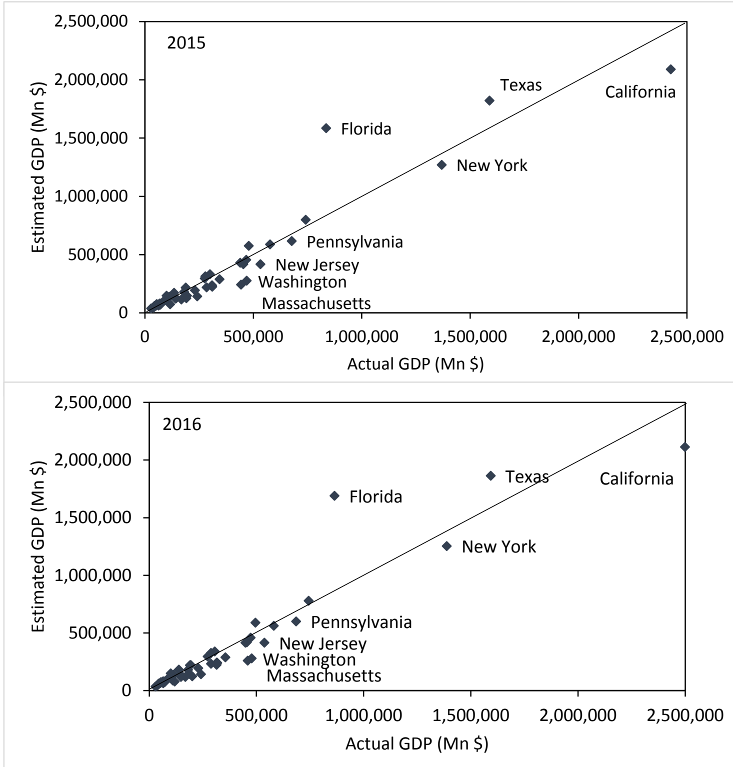
Results using area of each state as a weigth.

The regression analysis for both years also show that population have a statistically significant non-lineal effect over GDP per state. Finally, the regressions for both years present a good fit since the R^2 are 0.956 and 0.955 for the years 2015 and 2016 respectively.

Graph 1 presents the estimated GDP versus the actual GDP per state in the United States. The model presents on average a good fit for GDP estimations at the state level. The highest over estimations are presented for California and New York, while Florida and Texas register the highest underestimations. The US GDP estimated for 2015 by nighttime lights is 16,725,226.91 million dollars, which is obtained by adding the predicted GDP for every state. The official data for the same year is 17,048,048.58[6]. Therefore, there is a small underestimation of 1.89%. Similar results are obtained for the year 2016.

[6] We only include continental US in our calculations.

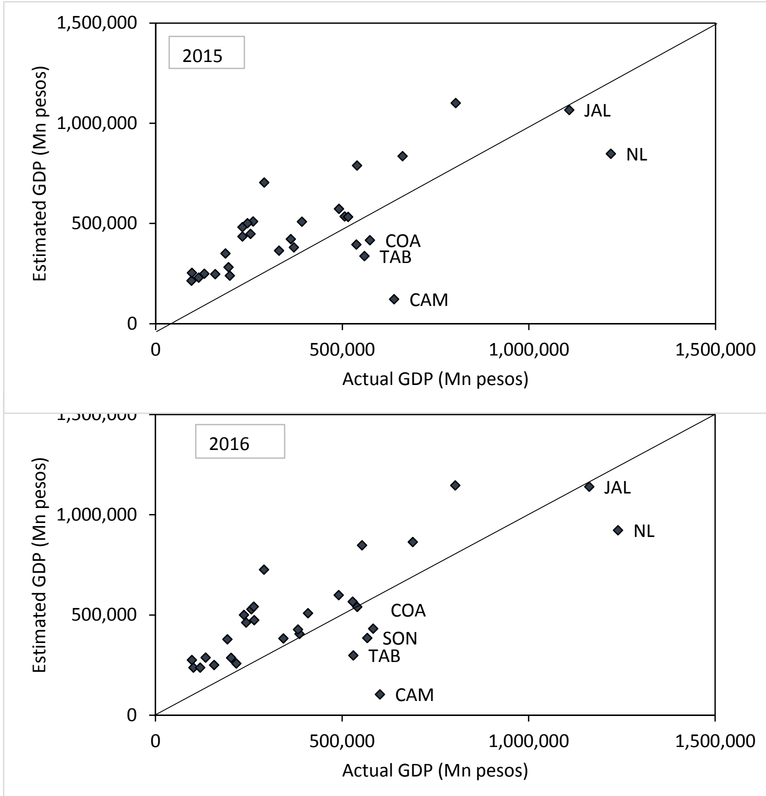
Graph 1. 1:1 Plot of the Actual versus Estimated GDP for the U.S. states



Source: Own calculations using NASA/NOAA , VIIRS-Night Light Satellite Data and INEGI.

Graph 2 presents the relationship between the official GDP per state in México and the estimation of GDP via luminosity registered in México and the US parameters. The graph shows that, for both years, GDP was overestimated for 26 states and underestimated for 6: Mexico City, Coahuila, Nuevo León, Sonora, Tabasco y Campeche. For the last two states, the difference might be explained for the luminosity in the sea generated by the oil platforms that has not been accounted in our estimation (these two federal entities are the biggest producers of oil in Mexico). For Mexico City, the overestimation could be explained because some firms have their corporative offices and are officially registered in that state but their production factories are located in other federal entities.

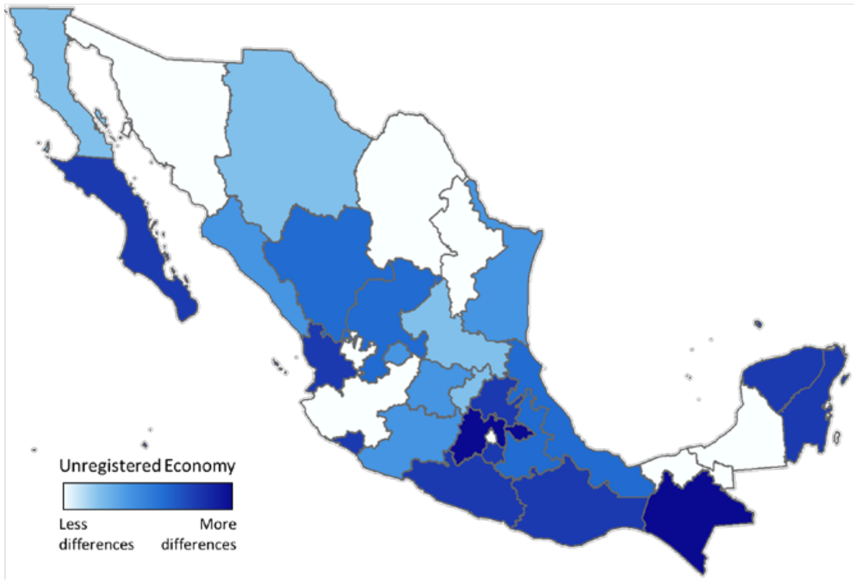
Graph 2. 1:1 Plot of the Actual versus Estimated GDP for the Mexican states



Source: Own calculations using NASA/NOAA , VIIRS-Night Light Satellite Data and INEGI.

Figure 4 contains the differentials between the estimated GDP by nightlights observed in Mexico and the US parameters for every state. It is possible to observe that the sates with the highest levels of non-registered activity (informality) are located in the south (with the exception of the mentioned oil producers). In contrast, the states located in the north of the country presents the lower levels of informality. Table A1 on the appendix contains the estimation of these differentials for every state and the ranking of the level of non-registered economic activity.

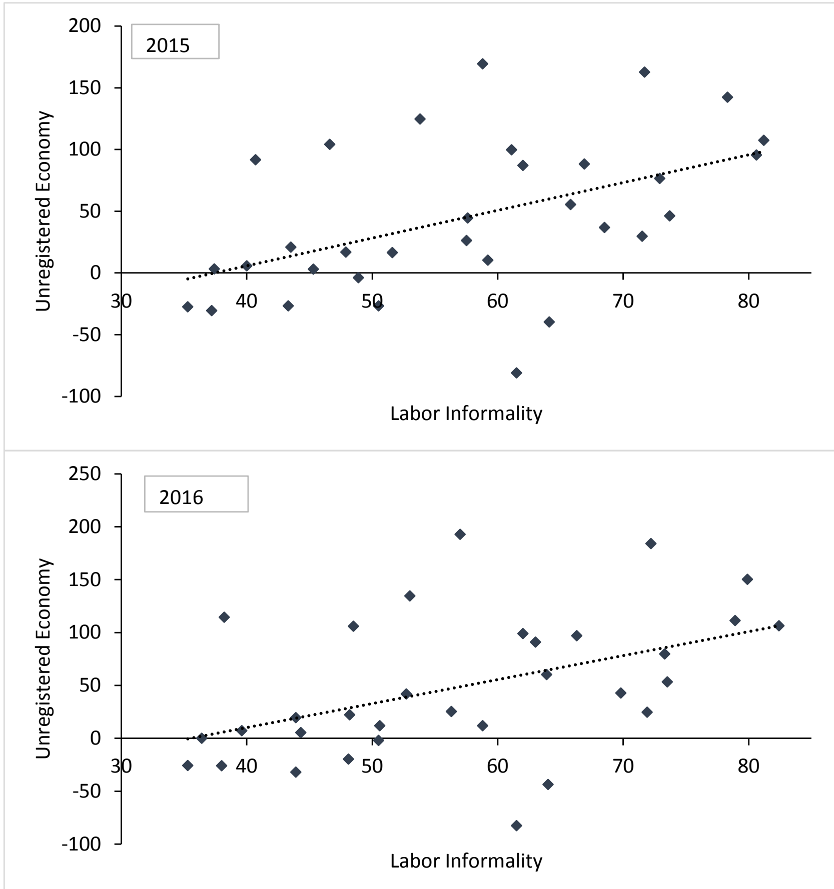
Figure 4. GDP Differentials using nightlights and US parameters (in percentage points)



Source: Source: Own calculations using NASA/NOAA , VIIRS-Night Light Satellite Data and INEGI.

INEGI has estimates of the informal labor force for every state in Mexico. In order to revise if our estimations of the non-registered activity follow a similar pattern of the official levels of labor informality we decide to plot our estimations versus the INEGI´s measure. Graph 3 shows a clear positive relationship between the two variables. This indicates that there is a positive correlation between the non-registered measure of economic activity and the official measure of labor informality (correlation of 0.235 and 0.212 for 2015 and 2016, respectively). Therefore, these differentials between the official GDP with respect the night lights estimates might represent a good way of ranking the levels of non-registered activity among states in Mexico.

Graph 3. Labor Informality vs Non-Registered Economy



Source: Own calculations using NASA/NOAA , VIIRS-Night Light Satellite Data and INEGI.

Finally, in order to estimate the non-registered economy, we add the estimated GDP for all states and subtract the sum of all official GDPs per state. The estimated difference is dived by the official measure and multiplied by a 100. The result indicates that the size of the non-registered economy is 22.72% and 27.78% of GDP for the years 2015 and 2016 respectively[7]. That is, our results indicates that an average of 25.25% of the economic activity in Mexico is not registered by the official numbers.

VI.II Estimates using the Nuevo Leon as the Reference Economy

Table 4 presents the results for equation 4, which was estimated using OLS and the labor informality measure of INEGI as weight. The estimation for both years indicates a higher elasticity of luminosity

[7] When performing the same exercise for the US economy the estimations are 1.89% for 2015 y 2.39% for 2016.

with respect to GDP for the manufacturing sector, the second highest elasticity is founded in the tertiary sector, and the lowest response between these two variables is registered, as expected, in the primary sector of the economy. Also, same as before, the coefficients are statistically significant for the tertiary and the secondary sector but not for the primary sector. It is also possible to observe quadratic effect of population over GDP. The coefficient of the interaction term between light and dummy for the state of Nuevo Leon is positive and statistically significant. Finally, the regressions for both years present a good fit since the R^2 is 0.817 for 2015 and 0.805 for 2016.

Table 4. Estimation Results of Equation 4, Mexico's Data, 2015 and 2016

Variables	(1) 2015		(2) 2016	
Ln(Light)*Primary Sector	0.029		0.002	
Ln(Light)*Secondary Sector	0.407	***	0.339	***
Ln(Light)*Tertiary Sector	0.336	***	0.255	***
Population	0.288	***	0.303	***
Population^2	-0.011	***	-0.011	***
NL*Light	0.023	***	0.027	***
Constant	6.433	***	7.591	***
Observations	1835		1828	
R^2	0.817		0.805	

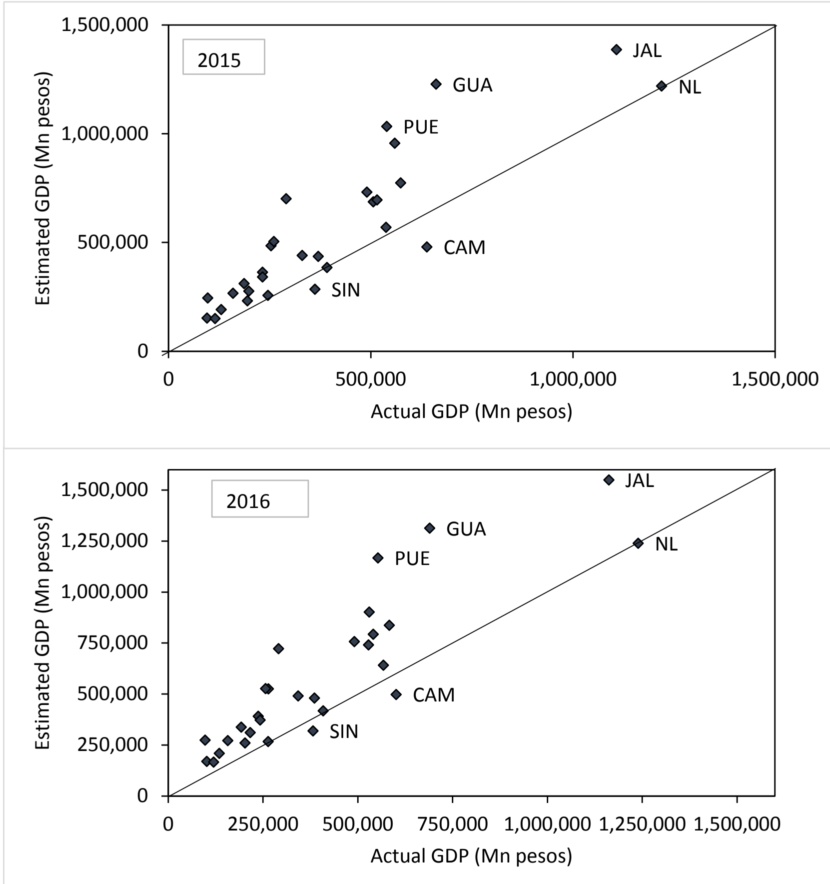
Source: Own calculations using NASA/NOAA , VIIRS-Night Light Satellite Data and INEGI.

Results using labor informality ratios as a weigths.

Graph 4 presents the relationship between the official GDP per state in México and the estimation of GDP via luminosity registered in México and using the parameter of luminosity that takes into account the effect of Nuevo Leon. The graph shows that, for both years, GDP was overestimated for 27 states and underestimated for 4: Mexico City, Campeche, Sinaloa and Michoacan, in this last state the differential is less than 2 percent.[8]

[8] For construction Nuevo León has a differential equal to zero.

Graph 4. 1:1 Plot of the Actual versus Estimated GDP for the Mexican states

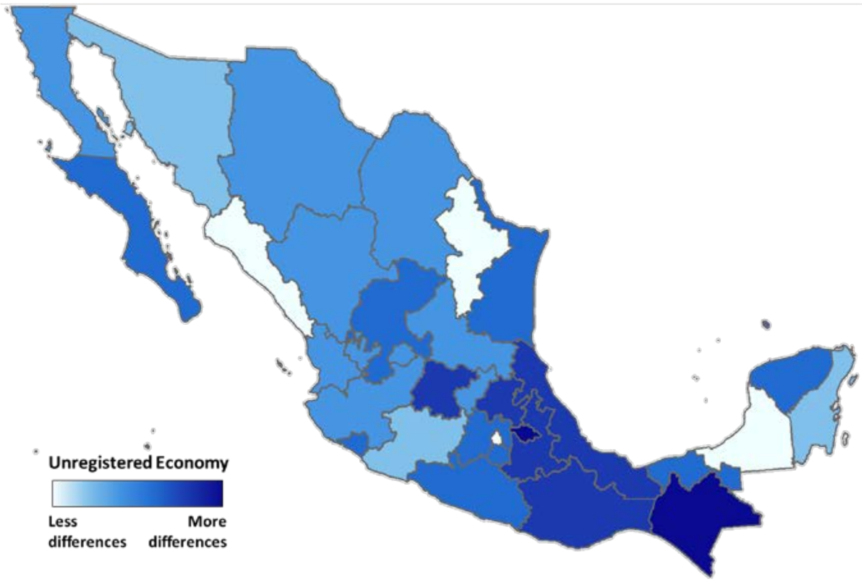


Source: Own calculations using NASA/NOAA , VIIRS-Night Light Satellite Data and INEGI.

Same as before, we construct a map showing the differentials, for every state, between the estimated GDP and the official GDP reported by INEGI (Figure 2). Although the differentials do not replicate the exact ordering of the previous estimation, they have the same general pattern with lower differentials in the north of the country and higher levels of non-registered economy in the south of Mexico. Table A2 on the appendix contains the estimation of these differentials for every state and the ranking of the level of non-registered economic activity.

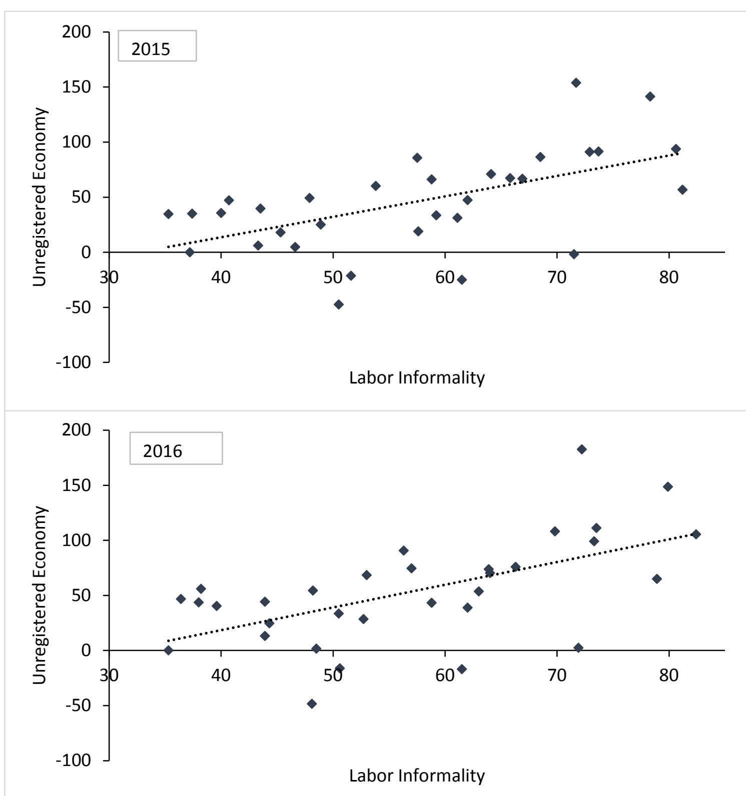
Graph 5 presents the relationship between the informal labor measure of INEGI and the calculated differentials between GDP estimated and official GDP per state. Once again, it is also possible to identify a positive correlation between these two measures of informality.

Figure 2. GDP Differentials using nightlights and NL as the Reference Economy



Source: Own calculations using NASA/NOAA , VIIRS-Night Light Satellite Data and INEGI.

Graph 5. Labor Informality vs Non-Registered Economy



Source: Own calculations using NASA/NOAA , VIIRS-Night Light Satellite Data and INEGI.

In order to estimate the non-registered economy, we add the estimated GDP for all states. Same as before, the estimated difference is divided by the official measure and multiplied by a 100. The result indicates that the size of the non-registered economy is 27.80% and 34.54% of GDP for the years 2015 and 2016 respectively. That is, our results indicates that an average of 31.17% of the economic activity in Mexico is not registered by the official numbers.

VII. Conclusions

The shadow economy in developing economies, such as México, represents a non-negligible size of the economic activity. Therefore estimation of the total economic activity, including official and non-registered production of goods and services, is important to determine the welfare level of the population in one country, it is also essential for designing economic policies to encourage development, or policies that respond to changes in the economic cycle. In addition, an estimation of the shadow economy is a fundamental input to estimate tax evasion and for design policies aimed for combating this problem.

In this research we use satellite images of nightlights, through which economic activity can be captured, to measure economic activities not recorded in GDP by the differential between the estimated output by brightness and the official data. The proposed econometric specification accounts for the fact that the relationship between nightlights and economic activity differs by sector when estimating the correlation between luminosity and economic activity. To derive our estimations we use data from the NASA/VIIRS satellites, which have higher resolution on luminosity and has similar or even better for estimating economic statistics than DMSP lights used in previous works.

Our estimates indicates that the size of the non-registered economy in Mexico is about 25.25 % of the GDP when using the US economy as a reference, and 31.17% when using the state of Nuevo Leon as a benchmark economy. State level estimates should be taken with caution since the econometric model does not allow to considerate to many specific characteristics, other than mixture of primary, secondary and tertiary sector within every federal entity. However, they can be used as a starting point for ranking the size of the shadow economy.

One limitation of the study is that we are assuming that the relationship between nightlight data an economic activity is the same for the United States and Mexico (in the first estimation) and for Nuevo Leon and the rest of the estates (in the second estimation), which might be not true. The use of the mixture of economic sectors (primary, secondary and tertiary) helps to attenuate the problem but not to eliminate it, since there might be differences in technology within every sector.

Finally, it is worthy to mention that this research is still in progress and therefore the estimates are preliminary and must be taken with caution.

Bibliography.

Bennett, M. and Smith, L. (2017). Advances in Using Multitemporal Night-Time Lights Satellite Imagery to Detect, Estimate, and Monitor Socioeconomic Dynamics. *Remote Sensing of Environment*, 192, pp. 176-197.

Beyer, R. C., Chhabra, E., Galdo, V., & Rama, M. (2018). Measuring Districts' Monthly Economic Activity from Outer Space. *Policy Research Working Paper 8523*.

Brock G., Tian J., and Yarbrough R. (2014). "The Informal Economy of Veracruz State During The Fox Administration". The Journal of Developing Areas, Vol. 48, No. 2 (Spring 2014), pp. 153-168.

Chapa, C.J. Flores D. Valero J. (2008). "La economía informal: estimaciones, comportamiento y potencial recaudatorio", Trillas.

Chen. X. and Nordhaus, William D. (2011). "Using Luminosity Data as Proxy for Economic Statistics". Proc. Natl. Acad. Sci. USA 2011, 108, 8589-8594. 12.

Chen, X., and Nordhaus, W. D. (2019). "VIIRS Nighttime Lights in the Estimation of Cross-Sectional and Time Series GDP". *Remote Sensing*, 1-11.

Donaldson D. and A. Storeygard (2016). "The View from Above: Applications of Satellite Data in Economics." Journal of Economic Perspectives, Vol.30, No.4, 171-198.

Deren Li, Zhao Xia and Li Xi. (2016). "Remote sensing of human beings – a perspective from nighttime light". Geo-spatial Information Science. 19, 69-79.

Doll, C. N.H., Muller, J-P., Morley, J. G. (2006). "Mapping regional economic activity from night-time light satellite imagery". Ecological Economics 57 (1), 75-92.

Doll. Christopher; Muller, Jan-Peter: and Elvidge, Christopher D. (2000). "Nighttime imagery as a tool for global mapping of socioeconomic parameters and greenhouse gas emissions". Ambio, Vol. 29, 157-162.

Ebener, Steeve, et al. (2005). "From Wealth to Health: Modeling the Distribution of Income per Capota at the Sub-national Level using Night-time Light Imagery". International Journal of Health Geographics.

Elvidge, C. D., Baugh, K. E., Kim, E. A., Kroehl, H. W., Davis, E. R., Davis, C. W. (1997). "Relation between satellite observed visible-near infrared emissions, population, economic activity and electric power consumption". International Journal of Remote Sensing 18 (6), 1373-1379.

Elvidge, Christopher D., et al. (2009) "A Global Poverty Map Derived from Satellite Data. Computers & Geosciences", 35(8), pp. 1652-1660.

Elvidge, CD.; Kimberly, B.; Mikhail, Z.; Feng, C.H.; Tilottama (2017), G. VIIRS night-time. Int. J. Remote Sens, 38, 5860-5879.

Gosh, T. Anderson S. Powerl R. Sutton, P. Elvidge C. (2009). Estimation of Mexico's Informal Economy and Remittances Using Nighttime Imagery. Remote Sensing, 418-444.

Gosh, T. Anderson S. Powerl R. Sutton, P. Elvidge C. (2010). Informal Economy and Remittance Estimates of India Using Nighttime Imagery. International Journal of Ecological Economics and Statistics. Vol. 17., 16-50.

Guerrero, V. M., & Mendoza, J. A. (2019). "On measuring economic growth from outer space: a single country approach". *Empirical Economics*, 971-990.

Gyomai, G. and van de Ven, P. (2014). The Non-Observed Economy in the System of National Accounts, OECD, Statistic Bief No. 18, June 2014.

Harati, R., Hardy, M., 2013. "Illuminating the shadow economy: using lights from space to estimate a new measure of informal economic activity". Centre d´Economie de la Sorbonne: mimeo.

Henderson J., Soreygard, A. Weil D. (2012). "Measuring Economic Growth from Outer Space", American Economic Review, Vol. 102, No. 2, April 2012, 994-1028.

Hu, Y., & Yao, J. (2019). Illuminating Economic Growth. *IMF Working Paper WP/19/77*.

Keola, S., Andersson, M., Hall, O., 2015. "Monitoring economic development from space: using nighttime light and land cover data to measure economic growth". World Development 66, 322-334.

Llamosas-Rosas, I., Rangel González, E., and Sandoval Bustos, M. (2018). "Medición de la actividad económica en las principales zonas turísticas de playa en México a través de la luminosidad fotografiada desde el espacio", Working Papers, No. 2018-10, Banco de México, Ciudad de México

Medina, L., & Schneider, F. (2018). Shadow Economies Around the World: What Did We Learn Over the Last 20 Years? IMF Working Paper WP/18/

Sutton, P.C.; and Costanza, R. (2002). "Global estimates of market and non-market values derived nighttime satellite imagery, landcover and ecosystem service valuation". Ecological Economics, Vol. 41, 509-527.

APPENDIX

Table A1. Ranking by state of Informal Labor Rate (INEGI) and Non-Registered Economy using US Data (own calculations) for 2015 and 2016.

Mexican State	Informal Labor Rate		Non-Registered Economy (from Eq. 2)			
					Levels	
	2015	2016	2015	2016	2015	2016
Aguascalientes	26	26	19	20	-20.96	-19.67
Baja California	29	28	23	23	-5.76	-7.35
Baja California Sur	28	29	9	5	-91.80	-114.55
Campeche	13	14	32	32	80.82	82.59
Coahuila de Zaragoza	32	30	29	29	27.35	25.82
Colima	19	18	4	4	-124.81	-134.60
Chiapas	3	2	3	3	-142.52	-150.21
Chihuahua	30	31	24	25	-3.38	-0.11
Ciudad de México	21	24	28	27	26.67	19.61
Durango	17	19	15	16	-44.53	-41.86
Guanajuato	18	17	18	17	-26.45	-25.45
Guerrero	1	3	5	6	-107.54	-111.34
Hidalgo	5	5	12	12	-76.64	-79.77
Jalisco	22	21	26	26	3.70	1.81
México	16	16	1	1	-169.57	-192.79
Michoacán de Ocampo	7	7	17	18	-29.87	-24.66
Morelos	9	9	10	10	-88.25	-97.05
Nayarit	14	13	7	9	-99.84	-99.00
Nuevo León	31	32	30	28	30.42	25.53
Oaxaca	2	1	8	7	-95.71	-106.35
Puebla	4	4	14	14	-46.24	-53.35
Querétaro	25	25	25	24	-3.14	-5.42
Quintana Roo	24	22	6	8	-104.29	-105.91
San Luis Potosí	15	15	22	21	-10.41	-12.06
Sinaloa	20	20	21	22	-16.55	-11.98
Sonora	27	27	27	30	26.56	32.01
Tabasco	11	10	31	31	39.58	43.56
Tamaulipas	23	23	20	19	-16.89	-22.22
Tlaxcala	6	6	2	2	-162.80	-184.15
Veracruz de Ignacio de la Llave	8	8	16	15	-36.96	-42.81
Yucatán	12	12	11	11	-87.24	-90.97
Zacatecas	10	11	13	13	-55.55	-60.31

Source: Informal Labor Rate: INEGI; Non-Registered Economy: Own calculations using data from INEGI, BEA and NASA/NOAA.

Table A2. Ranking by state of Informal Labor Rate (INEGI) and Non-Registered Economy using Mexico's Data (own calculations) for 2015 and 2016.

	Informal Labor Rate		Non-registered Economy (from Eq. 4)			
			Rankings		*Levels*	
	2015	2016	2015	2016	2015	2016
Aguascalientes	26	26	26	26	-39.69	-44.15
Baja California	29	28	29	28	-35.80	-40.43
Baja California Sur	28	29	28	29	-47.19	-55.95
Campeche	13	14	13	14	24.89	17.07
Coahuila de Zaragoza	32	30	32	30	-34.83	-43.73
Colima	19	18	19	18	-60.27	-68.41
Chiapas	3	2	3	2	-141.39	-148.67
Chihuahua	30	31	30	31	-35.08	-46.74
Ciudad de México	21	24	21	24	47.35	48.43
Durango	17	19	17	19	-19.09	-28.44
Guanajuato	18	17	18	17	-85.71	-90.63
Guerrero	1	3	1	3	-56.86	-64.97
Hidalgo	5	5	5	5	-91.14	-99.02
Jalisco	22	21	22	21	-25.23	-33.42
México	16	16	16	16	-66.26	-74.60
Michoacán de Ocampo	7	7	7	7	1.77	-2.39
Morelos	9	9	9	9	-66.89	-75.96
Nayarit	14	13	14	13	-31.22	-38.81
Nuevo León	31	32	31	32	0.00	0.00
Oaxaca	2	1	2	1	-93.79	-105.35
Puebla	4	4	4	4	-91.61	-111.18
Querétaro	25	25	25	25	-18.10	-24.61
Quintana Roo	24	22	24	22	-4.79	-1.55
San Luis Potosí	15	15	15	15	-33.57	-43.29
Sinaloa	20	20	20	20	21.23	16.26
Sonora	27	27	27	27	-6.13	-13.13
Tabasco	11	10	11	10	-71.04	-70.24
Tamaulipas	23	23	23	23	-49.26	-54.44
Tlaxcala	6	6	6	6	-153.98	-182.53
Veracruz de Ignacio de la Llave	8	8	8	8	-86.60	-108.03
Yucatán	12	12	12	12	-47.36	-53.55
Zacatecas	10	11	10	11	-67.46	-73.72

Source: Informal Labor Rate: INEGI; Non-registered Economy: Own calculations using data from INEGI and NASA/NOAA.