Contents

List of Tables

List of Figures

Abstract

This study analyses the utility of nighttime lights data (NLD) in estimating and upgrading the economic statistics of various sub-national units. Combined with regression based statistical methods, we use intensity of NLD to conduct a spatiotemporal analysis of urban sprawl and associated loss in cropland over a 12 year period for Punjab - the most populous province of Pakistan. We also estimate the elasticity between a luminosity metric (sum of lights – SOL) and the value of large-scale manufacturing for a time-series analysis of regional growth patterns over two decades in Pakistan. We find an elasticity of approximately 0.4 between large-scale manufacturing (LSM) production and SOL, and find a convergence in growth, albeit a slow one, between the richer and poorest provinces of the country; an estimated trend in annual expansion rate of two percent among the five largest urban areas of Punjab, and the highest degree of cannibalization of prime cropland by the fastest growing urban regions in the province. Our analysis then highlights the utility of using a unique data source like the NLD to help uncover spatial patterns and regional growth trends in countries that face considerable data gaps.

Keywords: monitoring urbanization, regional development, time series and spatial analysis

1. Introduction

Do economic statistics like Gross Domestic Product (GDP) indicate the level of economic development achieved at sub-national level? Moreover, even if regional GDP does show progress at lower administrative levels do we really need such information? The answer to the first question is in the negative whereas that to the second one is in positive. GDP is output produced by country's factors of production located domestically in any given year. When constructing this economic statistic, questions of whether or not such factors of production reside inside the geographical boundary of the country is a matter of importance. However, the way these factors are located within the country is not relevant. Furthermore, for analytical purposes, individual components of GDP are segregated in terms of contributions made by various sectors – manufacturing, agriculture, services etc. – and not by shares contributed by different regions of the country. Due to this practice, the resulting GDP and all other statistics derived from it reveal performance of the national economy and not regional economies as represented by various provinces or districts.

The absence of equivalent measures of GDP and other related economic statistics at the sub-national level poses a great problem when the goal is to gauge the level of development attained by a given geographical region situated inside the country. A single figure of GDP might be increasing over time indicating growing economy. Yet, which lower level administrative units are the biggest contributors or the largest beneficiaries of this growth – information that is particularly relevant from a policy perspective – is usually not readily available especially in the context of most developing economies.

In India, the Directorate of Economics Statistics of respective states in collaboration with subdivision(s) of 'Ministry of Statistics and Programme

Implementation' publishes Gross State Domestic Product (GSDP), which is GDP-equivalent figure for each state. Similarly, the Department of Commerce in U.S. through its subsidiary Bureau of Economic Analysis (BEA) reports GDP by both state and metropolitan area. Unfortunately, no official agency in Pakistan imparts any such statistical data for provinces or districts. Nevertheless, few non-governmental endeavours have been undertaken over the years. Bengali and Sadaqat (2005) applied national accounts methodology at regional level and produced GDP and it components for four provinces of Pakistan. In recent times, Beconhouse National University's research think tank[1] has also published similar data for Punjab province. However, these efforts are sporadic in nature and present data that is out-dated or is restricted only to a specific region.

Similarly, most developing countries are severely restricted in understanding patterns of urbanization, again due to the lack of appropriate data. Even seemingly simple operations such as maintaining detailed accounts of geographical boundaries of urban centres is largely absent. For example, in Punjab province of Pakistan, the 'Punjab Development Statistics' reports provide vital information regarding agriculture production, population, utilities etc. for both administratively (divisions, districts and tehsils) and statistically (urban and rural) defined regions. Nevertheless, the geographical boundaries are not available for either type of regions. The absence of this cartographical information of Punjab's urban areas – the most populous province of the country – in turn makes it extremely difficult to analyse the level of development among these regions across time and space.

---

[1] Burki, S. J., and A. G. Pasha. "State of the Economy." The Punjab Story. Lahore: Institute of Public Policy, Beaconhouse National University (2012).

Indeed, it is urbanization's direct effects on population densities, infrastructure planning and basic utilities (Ramachandra, Aithal, and Sanna 2012), along with its importance for formulating national and sub-national government polices as highlighted by Stow et al. (2016) that has motivated our focus on the documentation of urban densification and patterns for the five major cities of Punjab.

Yet, uncovering the geographical boundaries of urban areas is not the only objective of this study. We are also interested in the sustainability of urbanization. Studies quantifying urbanization and economic covariates in terms of GDP per capita, population density, electricity consumption etc. are numerous both at national and sub-national level. However, economic benefits are one facet of urban growth. Yet, how the process of urbanization affects ecological landscape and various environmental variables is an area of inquiry that has received considerably less attention. Keeping in view the necessity of studying urbanization within the context of sustainability and in the absence of sufficient knowledge in the case of Pakistan, we consider the relationship between urban growth and food supply as measured by cropland in the case of two largest urban areas of Punjab[2]. Here, we examine urban growth in the context of loss of land used for growing crops and other agricultural purposes. Studying urbanization and crop-agriculture production is important as the ever-growing population of urban areas require a continuous supply of food for the fulfilment of nutritious needs which is directly linked to labour productivity.

Our paper then has three objectives: (1) we look at urban sprawl and quantify its trend for the most populous province in Pakistan from 1992 to 2012; (2) we consider the link between urban growth and the loss of cropland from 2000 to 2012, and thereby

[2] Graphical analysis of Faisalabad, Gujranwala and Multan available on request

study the sustainability of urbanization for largest urban centres in Punjab; and (3) we analyse regional patterns of economic growth in Pakistan as a whole spanning two decades.

However, a meaningful achievement of these objective is severely hampered by the lack of adequate data. In the presence of these gaps in data, we combine existing data sources like the Census of Manufacturing Industries (CMI) with a new and exciting data resource – nighttime lights data – to fill the gap in regional growth and urban analysis in Pakistan.

In particular, we use the value of large-scale manufacturing (LSM) output as a measure of economic development. Not only does LSM have play a major role in GDP: LSM had 11 percent share in GDP of Pakistan for the year 2010-11[3], we also possess its regional growth figures. Next, we explore the relationship between our luminosity metric and LSM. As shown in Figure 1, growth rates of nighttime lights metric SOL (explained later) and LSM[4] follow a similar pattern in case of all four provinces of Pakistan.

[3] Punjab Development Statistics 2014

[4] Provincial LSM growth rates based on Arby (2008)

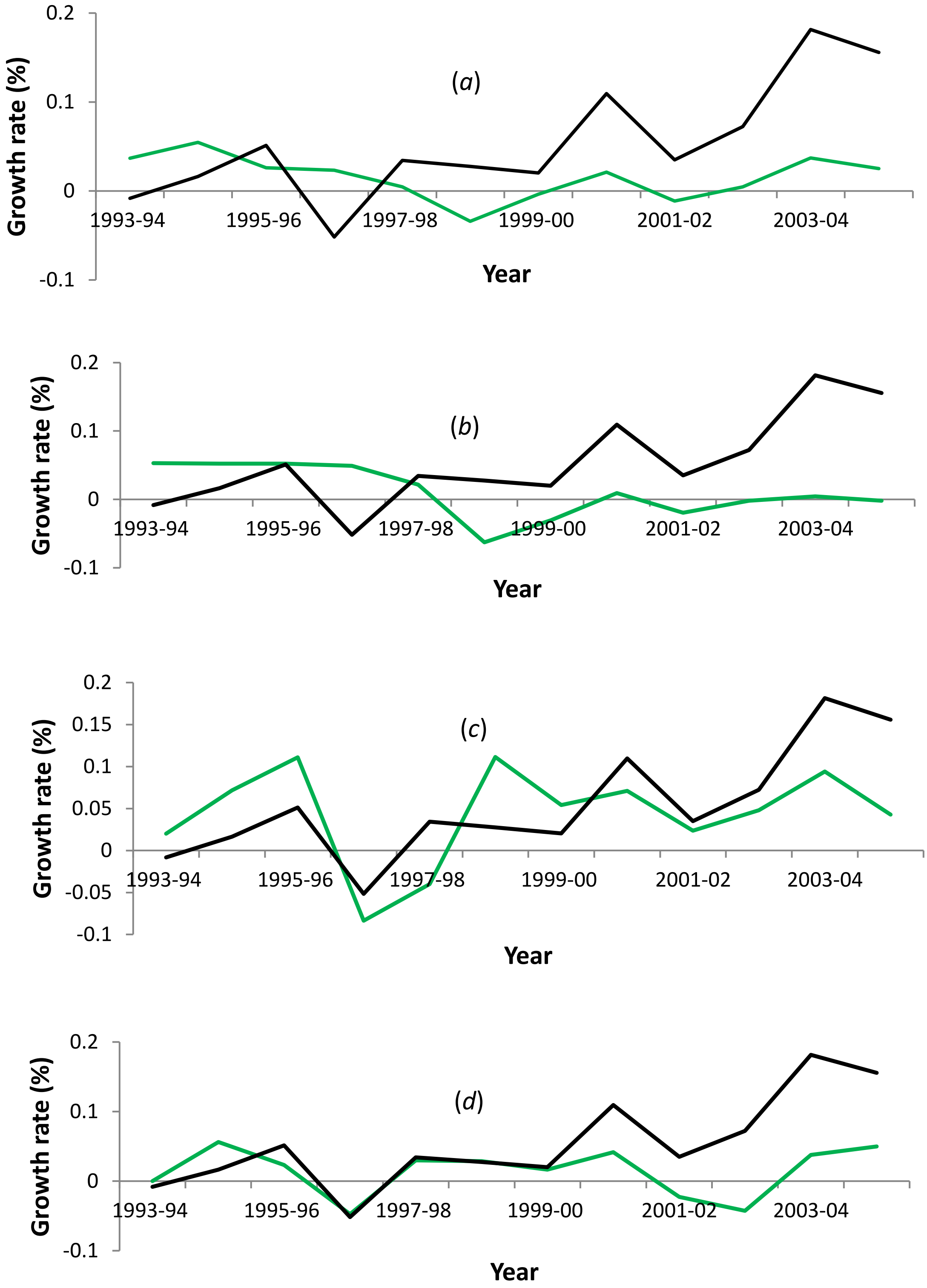


Figure 1: Growth rates of LSM (black) and SOL (green) for (a) Punjab, (b) Sindh, (c) Baluchistan and (d) Khyber Pakhtunkhwa or KPK

Given this high correlation, SOL then becomes a suitable proxy for economic development, which we then use to measure economic growth across regions and to highlight which regions in Pakistan have seen the greatest economic growth. Our analysis shows that publicly available data such as the Census of Manufacturing Industries (CMI) are overstating Pakistan's LSM growth rates. Moreover, we find that fluctuation in respective province's LSM growth rate is dependent on the initial values of GDP per capita and the region's share in SOL.

Our results vis-à-vis urbanization rates and associated cropland loss, we find that the fastest growing urban centre in Punjab – Islamabad/Rawalpindi – which is growing at an annual rate of four percent has the least sustainable pattern of urban growth as its urban boundary is expanding over largely more heavily cropland areas.

The remainder of this paper is structured as follows: immediately following the introduction, we review the literature in Section 2. Section 3 details the data and variable construction, while Section 4 discusses the econometric models and results. We conclude our study in Section 5.

2. Literature Review

2.1. Economics of Sub-national Regions and Phenomenon of Urbanization

There exists a wealth of literature that explores the nexus between GDP and various macro variables like interest rate, investment, money supply (Chang and Huang 2010; Hsiao and Hsiao 2006; Urbanovský 2016; Williams, Goodhart, and Gowland 1976). And although connections between these covariates have been theorized in detail and presented as empirical relationships in countless research papers, the association among the regional counterpart of these macro variables is a topic that has received less attention, despite its relevance in formulating meaningful national policies. The primary reason for this absence being the lack of data for sub-national units of countries.

Recent years however have marked a drastic change particularly when it comes to region-based work on India and China. For example, Sahoo and Acharya (2012) using Gross State Domestic Products (GSDP) of Indian states, constructed three indicators of GSDP per capita growth rate, price stability and fiscal balance, and used this information to rank states according to level of performance. Such state-level GDP values of India have also been utilized by Sehrawat and Giri (2015) in their paper which investigates the contribution of financial development to economic growth in 28 states.

Similarly, a good amount of research work has also emerged about China's provinces. Sheng, Shi, and Zhang (2014) focus on energy demand and economic growth and explore the relationship between the two by using dataset covering 27 Chinese provinces. After accounting for fixed effects and regional disparities as represented by varying urbanization and investment levels, the authors augment our knowledge by uncovering a positive effect of economic growth on demand for energy. Economic growth was also found by Yang and Xu (2006) to be the sole reason propelling the number of Small and Medium Size (SME) firms in Chinese provinces.

2.2. Night-time Lights for measuring Economic Output and Urban Sprawl

The use of nighttime lights data to explain various economic phenomena such as economic growth is more a recent development (see for example, Alesina et. al. 2012). Henderson, Storeygard, and Weil (2012) is by far the best example of seminal work in this department and is major source of inspiration for current work. Highlighting the need for improving economic statistics of developing countries, Henderson et al. (2012) estimated an elasticity of approximately 0.3 between GDP and luminosity measure derived from nighttime lights by using data on more than 150 countries and calculated revised estimates of GDP growth rates. Henderson’s et al. (2012) work established the link between nighttime lights and GDP at the multi-national level but did not explore this nexus at sub-national level.

However, (Bhandari and Roychowdhury 2011; Roychowdhury et al. 2012) verified the relationship established in Henderson et. al. (2012) at district(s) and village levels in the context of India with the help of simple regression techniques, and showed that sum of lights (SOL) along with its measures of central tendency did indeed explain alterations in GDP and its per-capita counterparts at smaller geographical regions. Yue, Gao, and Yang (2014) also confirmed the usefulness of night-time lights data at the sub-national level while employing it to divide GDP of China’s Zhejiang province among pixels of 250m resolution. They however also stressed the inadequacy of this data when used without the help of additional information e.g. land cover data.

Another major economic issue scrutinized with the help of nighttime lights in recent times relates to uncovering of the level and growth of urban areas. Keeping in view the varied nature of urbanization across countries, Zhang and Seto (2013) compared the estimates of urban growth using nighttime lights extracted via Google Earth for 240 different locations and affirmed the success rate at 93 percent. Similarly, Zhang and Seto (2011) centring the analysis on four countries, reaffirmed the ability of nighttime lights to capture the dynamics of urbanization. However, they also noted some reservations regarding oversaturation and non-appropriateness of data in the case of city-level geographical areas.

Turning to the use of nighttime lights data to understand urbanization patterns, we find several examples that focus on Chinese urban centres. Ma et al. (2015) divided nighttime lights images of various Chinese cities into five sections based on level of luminosity taking into account the differing socio-economic outlook within the city. They show an increase in lights intensity of all five sections of cities as urbanization has occurred over the years. Use of night-time lights in explaining variation in different economic statistics like population, GDP, electric consumption etc. and employment of different monotonic functions to establish empirical relationships between the variables was also justified by Ma et al. (2012) who researched on prefecture-level cities in the Chinese administrative structure. Ma et al. (2014) and Liu et al. (2012) also represent examples of capturing dynamics of China's urbanization with the help of nighttime lights bolstered by improved methods of inter-calibration and use of additional information in terms of geospatial topology.

2.3. Urban Sprawl and Agricultural Land

The dearth of data implies that the link of urban growth with crop production remains largely unstudied in the Pakistani context. However, there are a few examples in case of other countries that are worth mentioning. Yan et al. (2015) have explored the effects of urbanization on agricultural production in China by focusing on water use, and estimated a decrease of almost 0.5 percent in irrigated areas for all crops following a 1 percent increase in urbanization. Pandey and Seto (2015) also uncovered a negative relationship between agriculture land and urbanization in the case of India. Interestingly, they also show that the intensity of agricultural land loss is greater around smaller cities compared to that around larger ones. This finding is perhaps best explained by considering that smaller cities are more likely to be surrounded by richer cropland area, where intensive farming is likely to cause a much steeper drop in land productivity. Thus, we can expect that the geographical growth of smaller cities will likely have a greater impact on agricultural output. Colantoni et al. (2016) and Dadi et al. (2016) also highlight the harmful effects of urban expansion on cropland and agriculture production.

3. Data and Variables

3.1. Principal Dataset

Our primary dataset is "Version 4 DMSP-OLS Nighttime Lights Time Series" acquired from National Oceanic and Atmospheric Administration's (NOAA) National Centers for Environmental Information (NCEI) previously known as National Geophysical Data Center (NGDC). This dataset is freely accessible and available for download from http://www.ngdc.noaa.gov/eog/dmsp/downloadV4composites.html. The U.S. Air Force Defense Metrological Satellite Program (DMPS) Operation Linescan System (OLS) captures artificial lighting generated at the earth's exterior by piloting satellites that pass at night during 7pm to 9pm local time (Elvidge, Edward, et al. 2009) for the primary purpose of detecting cloud coverage.

There are two types of datasets under Version 4 of night-time lights time series: "Stable Lights" and "Average Lights x Pct". For our research, "Stable Lights" has been employed which covers the period from 1992 to 2013. Stable lights is probably the most widespread used type of nighttime lights data (hereafter, NLD) as it is available for greater number of years and is free from many complications.

The final dataset comprises of 34 composites or images gathered by six different satellites. Each composite relates to a specific calendar year and for some years, two composites are available as indicated in Table 1. The nature of NLD differs from data usually used by economists for analysis. Economic data generally involves amalgamation of rows and columns indicating numeric values for various variables or economic agents but in the case of NLD, information is visual in the sense that every year's data is submerged into an image that can be seen like any other picture. Images have a spatial resolution of 30-arc seconds meaning that each pixel represents approximately one km square of land area (Zhao, Zhou, and Samson 2015).

Quantitative information is attached with each pixel in the form of Digital Number (DN) that indicates intensity of light over the area covered by that pixel (Doll 2008). DN is numeric value that ranges from zero to 63; zero representing no light whereas 63 showing brightest light.

Year	**Satellite Names**					
	F10	F12	F14	F15	F16	F18
1992	✓					
1993	✓					
1994	✓	✓				
1995		✓				
1996		✓				
1997		✓	✓			
1998		✓	✓			
1999		✓	✓			
2000			✓	✓		
2001			✓	✓		
2002			✓	✓		
2003			✓	✓		
2004				✓	✓	
2005				✓	✓	
2006				✓	✓	
2007				✓	✓	
2008					✓	
2009					✓	
2010						✓
2011						✓
2012						✓
2013						✓

Table 1. Available composites of nighttime lights data.

3.1.1. Inter-calibration of NLD

Despite being attractive for economic analysis, the NLD loses much of its appeal when considered in raw form. As indicated by Zhao et al. (2015), use of NLD is limited when quantitative analysis covers more than one year. The major reason for this limitation is the absence of "radiometrically calibrated" images. The degradation of each satellite and its subsequent replacement along with the lack of an on-board calibration system in satellite(s) are major reasons preventing the comparison of images over time. Thus, the intensity of light as represented by DN values cannot be compared over time. For example, luminosity has not deemed to increase for particular pixel if it has a higher DN value in relation to DN value in preceding year. Elvidge, Daniel, et al. (2009) have produced a method that inter-calibrates the NLD thereby solving this issue. Almost every researcher using NLD for his/her study has adopted the Elvidge, Daniel, et. al, method or its variant[5]. Hence, we too use the inter-calibration as proposed in Elvidge,

---

[5] The Elvidge, Daniel, et al. (2009) inter-calibration is an experimental technique comprising of several steps. The first is to find the image (called reference image) with highest cumulative DN value. Second, we find the region or area (called reference area) in the chosen image that includes the full range of DN values from zero to 63 and changes very little over time. Finally, we plot the DN values of pixels comprising reference area for each satellite year in a scatter diagram against the DN values of same pixels in reference image and estimating inter-calibration coefficients using the following regression equation.

$$D_{new} = C_0 + C_1(D_{old}) + C_2\left(D_{old}^2\right) \quad (1)$$

where D_{old} refers to actual DN value attached with the image in raw form and DN_{new} is the inter-calibrated DN value whereas C_0, C_1 and C_2 are the coefficients. Based on this methodology, Elvidge, Daniel, et al. (2009) found satellite named F12's image of year 1999 and

Daniel, et al. (2009) and use the inter-calibration coefficients as calculated by Earth Observatory Group, NGDC, NOAA assistance - these coefficients are available for the time span considered in our analysis.[6]

The difference between the raw and inter-calibrated NLD values for Pakistan can be seen from the Figure 2. The unadjusted or raw DN sum for the whole region of Pakistan (including Azad Kashmir and Gilgit-Balistan) follows an upward trend but with frequent fluctuations in the series. On the other hand, the calibrated DN sum (relatively higher than the unadjusted one) improves in its smoothness.

However, even with inter-calibration of the NLD images, the data is still not completely useful for regression analysis. This is because of the difference in the format of years for which LSM and SOL information is available: The LSM value of production data is in the fiscal-year format whereas SOL information is available in calendar-year format. While regression analysis may still be feasible despite this irregularity, for consistency, ease of comparison, we average out the NLD images across and within years.

Sicily to be the reference year and region respectively that fulfilled the requirements set by the afore mentioned procedure.

[6] The authors would like to thank Tilottama Gosh for providing latest inter-calibration coefficients

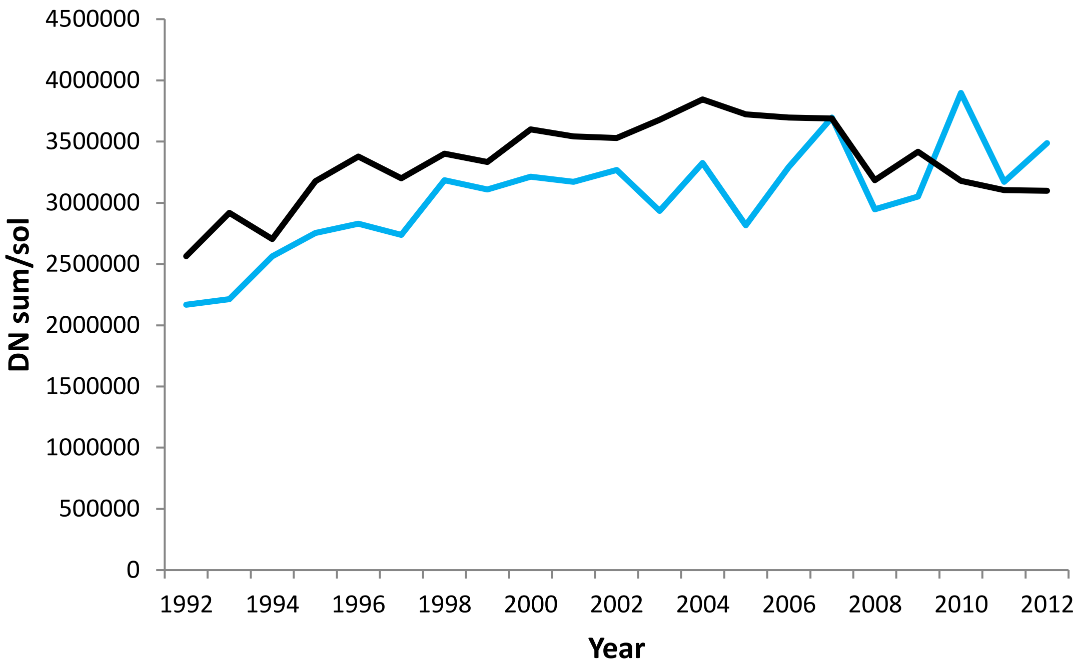


Figure 2: Time series of unadjusted (blue) and calibrated (black) DN sum for Pakistan.

3.2. Auxiliary Datasets

Apart from NLD, we use information on large-scale manufacturing by utilizing the Census of Large-Scale Manufacturing Industries (CMI) publications for the years 2005-06 and 2010-11, and provincial Development Statistics reports relating to years 1995-96 and 2000-01 to complete the empirical portion. CMI is a collaborative endeavour of Federal and Provincial bureaus of statistics that was undertaken in 1954 for the first time and since then has been conducted irregularly every five or so years. It seeks to collect information about employment, production value, raw material cost etc. of firms that employ ten or more employees and are involved in manufacturing process as defined in Pakistan's Factories Act 1934.

Table 2 gives an overview of the final district-level data used for regression analysis. Since NLD is available between 1992 and 2013, and many districts lack LSM information, an unbalanced panel data emerges. Furthermore, among the four provinces, Punjab has greater number of observations as it is divided into larger number of districts relative to other provinces, while Baluchistan's share of observations is the lowest. Tick marks in the year columns indicate presence of both NLD and CMI data for the respective districts.

		Fiscal Year			
Province	**No. of Districts**	**1995-96**	**2000-01**	**2005-06**	**2010-11**
Punjab	34	✓	✓	✓	✓
	1			✓	✓
Sindh	15	✓	✓	✓	
KPK	17			✓	
Baluchistan	2			✓	
TOTAL	69				

Table 2. Organization of district-level dataset for regression analysis.

For our analysis of urbanization and its sustainability, we make use of grid dataset complied by Ramankutty et al. (2008) that shows the global map of croplands for the year 2000. The spatial extent or the resolution of this dataset is 5-min that means that each grid cell equals approximately 10 km^2 of land area. Each pixel or grid cell of this dataset has been given a non-integer value that shows fraction (ranging from zero to one) of the physical area in the pixel covered by cropland (land used for the cultivation of food). For example, a pixel value of 0.3 means that 30 percent of the land area covered by the pixel is cropland.

4. Estimation

This section highlights the methodology followed, and the primary results for each of the objectives of this study.

4.1. Urban Sprawl in the Punjab

We conduct our analysis on five different regions of Punjab province. Four of these five regions correspond exactly with the four districts (Lahore, Faisalabad, Gujranwala and Multan) of the province whereas fifth region is an amalgamation of two areas – Islamabad Capital Territory (ICT) and Rawalpindi district. We have chosen these five regions for our analysis as these are the areas that have witnessed and are more likely to undergo further process of urbanization. Similarly, all of these five regions had population more than one million according the last census (1998) held in Pakistan.

For all the five regions, we choose the largest continuous urban or built-up area lying within these regions as our unit of analysis. However, the largest urban area need not be strictly residing inside the geographical boundary of the region/district. For example, the north-eastern part of largest urban area of Lahore district spreads into Sheikhupura district.

While there are varying techniques proposed with the literature when it comes to the extraction of urban or built-up areas from NLD or images, we make use of the 'thresholding technique'. The use of this technique requires that the researcher delineate the value of DN that effectively differentiates urban areas from other land areas. Although Imhoff et al. (1997) may be considered the pioneers of such work, more recent examples include Amaral et al. (2005) who have made use of the thresholding technique for extracting urban areas.

However, the primary drawback of using the thresholding technique is the choice of threshold (Gibson et al. 2015): if the threshold is set too low then we see an arbitrarily large urban area whereas setting it too high comes at the expense of the urban geographical area. In order to get around this issue we use several thresholds as suggested by the literature. This not only overcomes any concerns surrounding the arbitrariness of the threshold value but if different levels of thresholds result in the same pattern of urban growth over the years then we can assume that the final empirical growth rate of the respective urban areas is robust or closer to the true value.

Figure 3 explains the pre-processing executed on NLD for making different thresholds. Panel (a) of Figure 3 shows a portion of nighttime lights grid with each grid-cell of pixel attached with its respective DN value. Before choosing a threshold, all the pixels of nighttime image are divided by the maximum DN value that was assigned in the image. In this example, 63 is the maximum value. Panel (b) shows the resulting image after the division that is used for choosing different thresholds.

A threshold of 75 percent would mean that we are designating all those pixels with relative DN value of 0.75 and above as ‘urban or built-up’. Panel (c) of Figure 3 shows the extent of urban area (highlighted pixels) when 75 percent is used as the threshold whereas panel (d) refers to threshold of 70 percent.

(a)

45	48	51	63
40	46	47	62
28	36	35	50
22	36	36	50
19	21	31	49

(b)

0.71	0.76	0.81	1.00
0.63	0.73	0.75	0.98
0.44	0.57	0.56	0.79
0.35	0.57	0.57	0.79
0.30	0.33	0.49	0.78

(c)

0.71	0.76	0.81	1.00
0.63	0.73	0.75	0.98
0.44	0.57	0.56	0.79
0.35	0.57	0.57	0.79
0.30	0.33	0.49	0.78

(d)

0.71	0.76	0.81	1.00
0.63	0.73	0.75	0.98
0.44	0.57	0.56	0.79
0.35	0.57	0.57	0.79
0.30	0.33	0.49	0.78

Figure 3: Illustrative example of construction of threshold and size of urban area extracted.

The graphical representation in panel (c) and (d) of Figure 3 clearly shows that the size of extracted urban area depends on the level of threshold chosen – an issue raised earlier. Additionally we also estimate the boundaries of largest urban areas at 65 and 85 percent threshold levels – shown in Figure 4 in the case of Lahore.

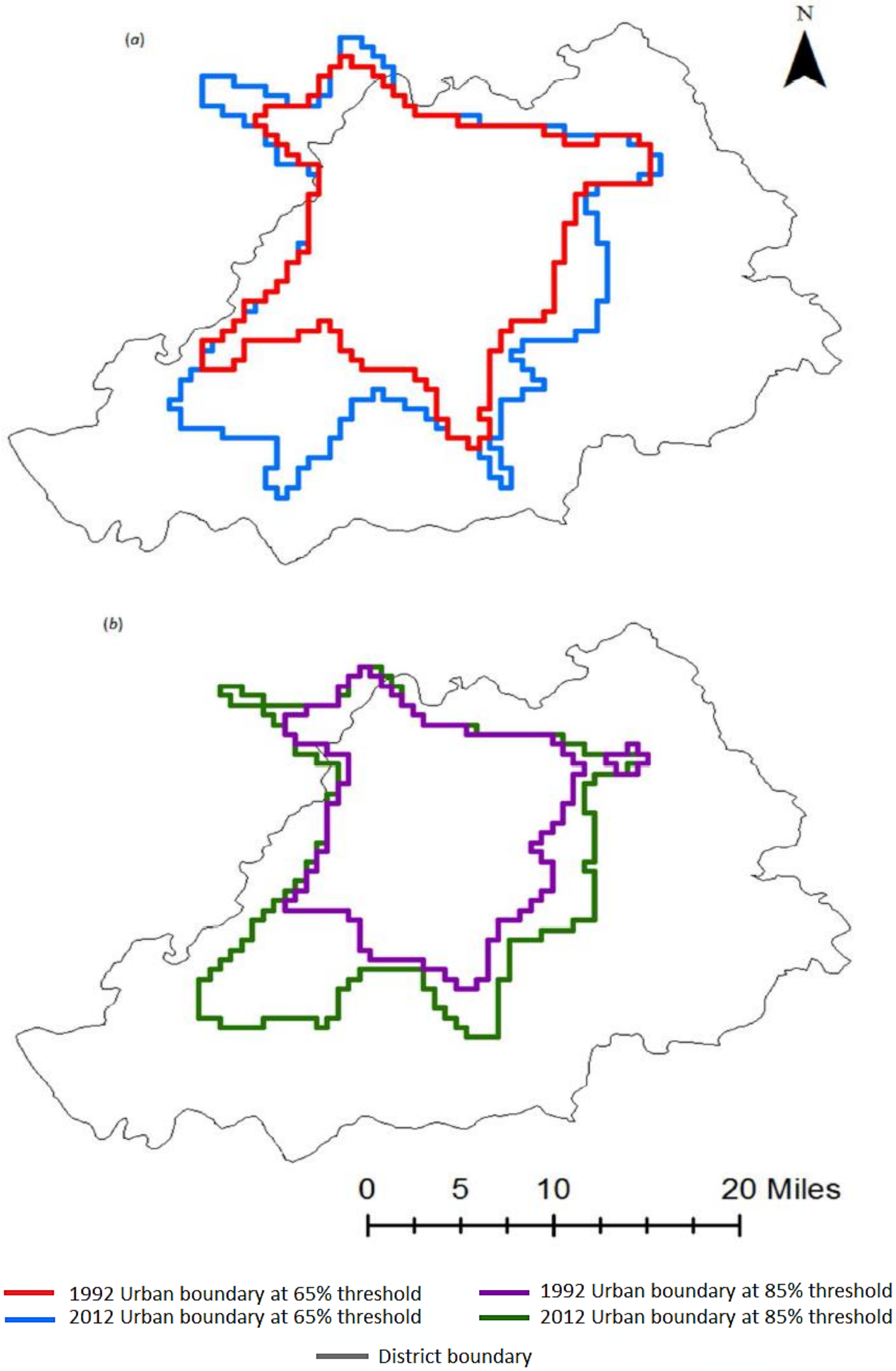
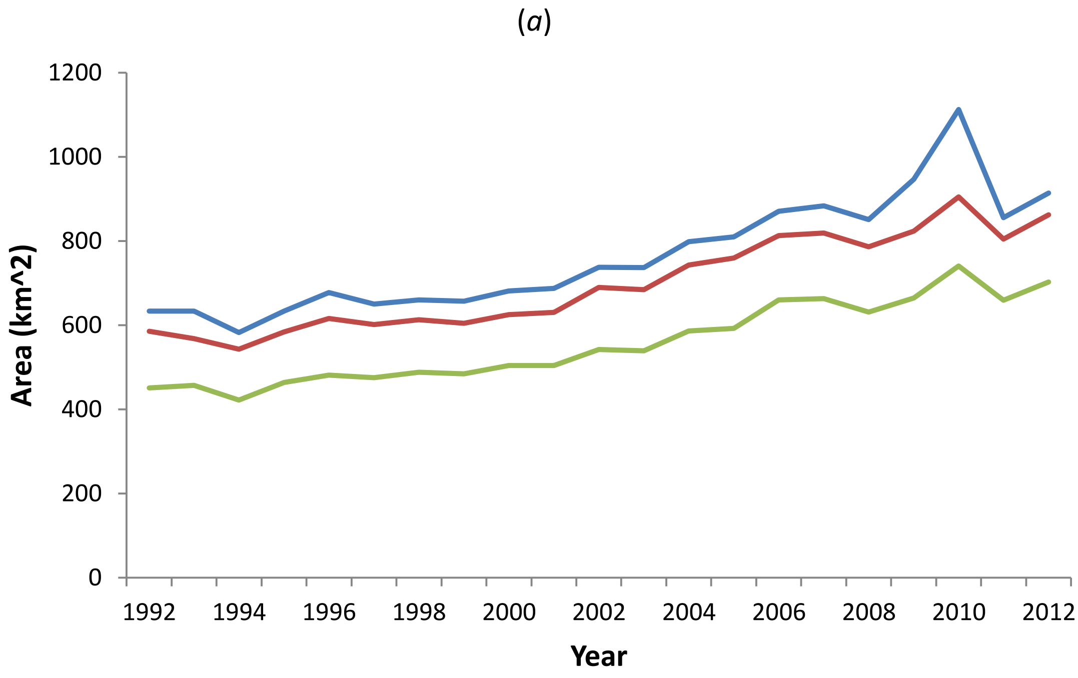


Figure 4: 1992 and 2012 boundaries of largest urban area of Lahore District with respect to (a) 65 percent and (b) 85 percent threshold levels.

Figure 5 (a) and (b) provide graphical representations of the annual trend for the largest urban areas – Lahore and Faisalabad districts – over the last two decades using three different threshold levels. The vertical axes of this figure represent area (in km^2) of the urban area extracted from NLD. For each year, the area of the urban centre is calculated by multiplying the number of pixels comprising urban area and size of nighttime lights pixel (i.e. 0.86 km^2 at the Equator).

As is visible from Figure 5 while the numerical size of urban area varies with the use of different thresholds, the trend followed by urban areas does not seem to depend on the choice of threshold. This conclusion is valid in case of all five urban areas of Punjab[7]. In case of Lahore (panel a), urban area is following an upward trend. This consistent increase in Lahore's sprawl is an expected outcome as this the most developed of all the cities of Punjab, typically offering the most employment opportunities along with other amenities. On the other hand, trend expansion of urban areas of Faisalabad can be bifurcated into two periods. Between 1992 and 2006, an increasing trend is visible from panel (b) whereas post 2006 period there is a decreasing trend.

[7] For brevity, Figure 5 does not show trends of Gujranwala, Islamabad/Rawalpindi and Multan districts.



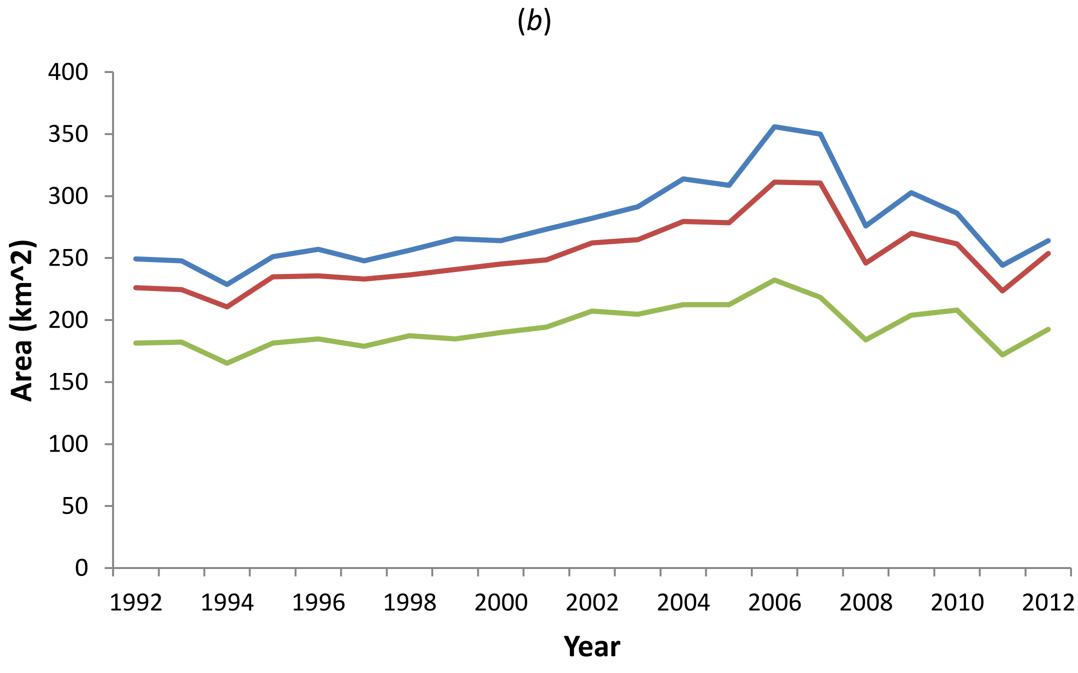


Figure 5: Time series of urban areas of (a) Lahore and (b) Faisalabad extracted corresponding to 65% (blue), 70% (red) and 85% (green) thresholds.

While such graphical analysis as shown in Figure 5 is informative as it gives a sense about the direction of trend followed by urban regions of Punjab province, quantifying urban expansion is also important from a policy-making perspective. To do this, we estimate average annual urban area growth rates for both regions as a whole and for individual urban centres. Equation (2) draws on the empirical procedure used by Gibson et al. (2015).

$$\ln A_{it} = B_0 + B_1 T + \rho_i (T \; x \; F_i) + B_3 F_i + \epsilon_{it} \qquad (2)$$

In the above equation, i and t indexes urban area and time respectively. A_{it} refers to geographical size of urban area in time t, T is the time trend, F_i is fixed effects for each urban area, and ϵ_{it} is the random error term. Equation (2) then is a Fixed Effects (FE) model that makes use of F_i term to take into account both observed and unobserved 'heterogeneity' or 'individual effect' of cross–sectional units (urban area in our case) (Greene 2003), where individual effects could be climate conditions, availability of natural resources, distance to river etc.

As explained by Gibson et al. (2015), $\hat{B}_1$represents the percentage change in urban area for one unit increase in T and indicates trend annual rate of expansion (after controlling for characteristics of each cross-section unit) in the case when regression is run without the interaction term $(T \; x \; F_i)$. On the other hand, $(\hat{B}_1 + \hat{\rho}_i)$ represents trend expansion rate for the i^{th} urban area.

Table 3 reports the results of estimating equation (2) at three different thresholds. Apart from interactions terms involving fixed affects for Faisalabad and Gujranwala districts, all coefficients are statistically significant at the one percent level at all three thresholds. We concentrate our analysis on the 85 percent threshold level as the regression run using this level produces the highest value of R^2. Since equation (2) is in log-level form, so the coefficients of explanatory variables can be interpreted in percentages.

(1)	(2)	(3)	(4)	(5)
VARIABLES	**65% Threshold** $\ln Area$	**70% Threshold** $\ln Area$	**85% Threshold** $\ln Area$	**85% Threshold** $\ln Area$
T	0.0118***	0.0132***	0.0138***	0.0199***
	(0.00298)	(0.00319)	(0.00302)	(0.00177)
$T\ x\ F_{lahore}$	0.0133***	0.0111**	0.0124***	
	(0.00422)	(0.00452)	(0.00427)	
$T\ x\ F_{islamabad}$	0.0260***	0.0252***	0.0260***	
	(0.00422)	(0.00452)	(0.00427)	
$T\ x\ F_{faisalabad}$	-0.00158	-0.00364	-0.00695	
	(0.00422)	(0.00452)	(0.00427)	
$T\ x\ F_{gujranwala}$	-0.00167	-0.000638	-0.000394	
	(0.00422)	(0.00452)	(0.00427)	
$Constant$	5.566***	5.467***	5.172***	5.171***
	(0.0167)	(0.0179)	(0.0170)	(0.0222)
Observations	105	105	105	105
R-squared	0.741	0.720	0.756	0.5626
Number of urban areas	5	5	5	5

Note: ***Significant at 1% level,
**Significant at 5% level,
*Significant at 10% level,
Standard errors in parenthesis

Table 3. Regression results for model or equation (2).

The coefficient of (T) variable in regression specification of column (5) of Table 3 gives the trend annual rate of expansion. According to our results, the five largest urban areas have expanded annually at the rate of two percent, on average, during the last 21 years ending on 2012.

The urban area-specific trend expansion rates can be found by summing the coefficient of T variable and the coefficient of respective urban area's interaction term (ρ_i). The respective trend expansion rates of all of the largest five urban areas of Punjab are provided in Table 4 – results are based on the regression estimates that were derived using threshold level of 85 percent used for detection of urban boundaries. The largest urban agglomerations of Lahore and Islamabad/Rawalpindi[8] are expanding annually at the rate of 2.6 percent and four percent respectively.

Region/City	Annual trend expansion rate
Lahore	2.6%
Islamabad/Rawalpindi	4.0%
Faisalabad	0.7%
Gujranwala	1.3%
Multan	1.4%

Table 4. Annual trend expansion rates of Punjab's largest cities.

4.2. Sustainability of Urban Growth

For analysing the sustainability of urban growth, we use a new methodological approach. We combine two pieces of information – unban boundaries as uncovered by NLD, and probabilistic knowledge about regions covered by cropland – in order to conduct trend analysis of urban growth and to its effect on cropland.

[8] Islamabad and Rawalpindi are twin cities: Islamabad is the capital of the country while Rawalpindi houses the general headquarters of the Pakistan army.

The first step involves creation of new raster[9] image for each year between 2000 and 2012 inclusive. This new image or urban mask is a binary representation of urban areas uncovered by NLD with number one attached to pixels identified as urban or built-up and zero otherwise. The second step involves the re-sampling of Ramankutty's et al. (2008) cropland dataset. Re-sampling is a procedure that is used when multiple raster datasets does not share same resolution. In our case, NLD are of 30-arc seconds whereas the global cropland image is of 5-min. Thus, we use ArcGIS 10.2.2 tool 'Resample' and convert the spatial resolution of the cropland image into that of night-time images. Thirdly, we divide each urban mask based on the cropland dataset. Pixel values of images produced after such division contains valuable information about how urban growth relates with the cropland area. As explained before, cropland or resampled-cropland dataset and urban masks contain pixel-values that range from zero to one. Thus, undefined pixels-values (i.e. one divided by zero) of post-division images represent part(s) of urban areas that are not covering any proportion of cropland whereas pixel-values greater than zero symbolize the spread of urban area over cropland.

As an example, consider the image shown in Figure 6 that is largest urban area of Lahore for the year 2000 as estimated using NLD at 85 percent threshold level. With the help of procedure described in the previous paragraph, we can divide this urban area with respect to categories of cropland prevailing in the year 2000. The grey coloured region in Figure 6 depicts part of urban/built-up area that is situated over land with zero percent cropland. Similarly, the red coloured region shows parts of the urban area that

---

[9] Raster image is rectangular array of pixels with each pixel attached some information (DN value in the case of NLD)

are situated over land with 62-80 percent cropland. Since cropland data is available only for 2000, we use this year as the base year and follow the changes in the sizes of the different sections of urban area with respect to different cropland-categories for years after 2000.

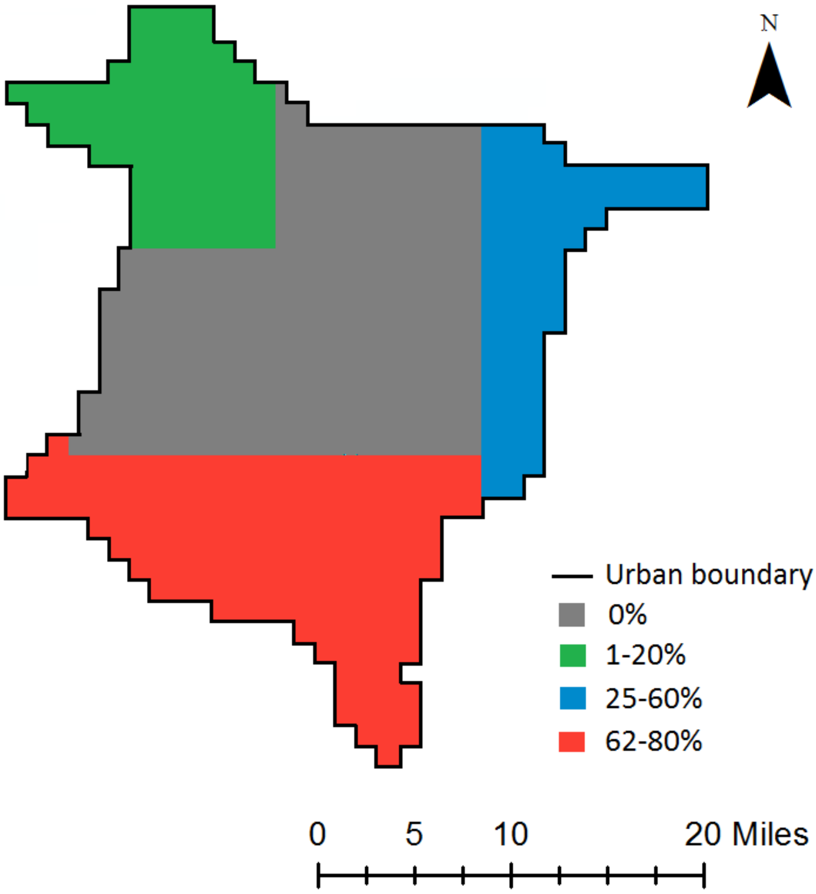
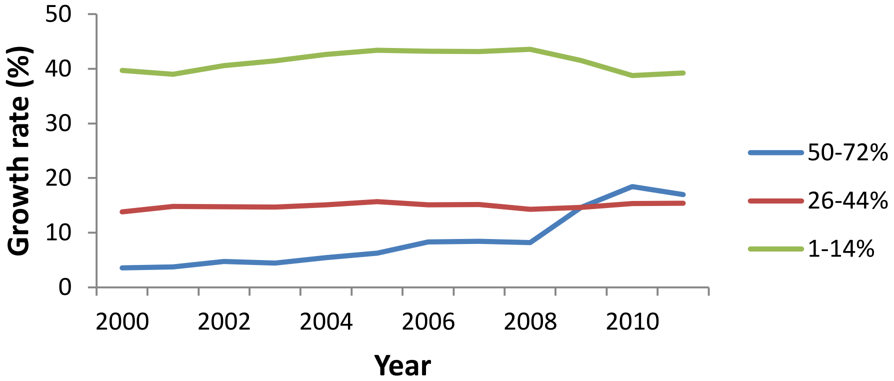


Figure 6: Illustrative example for understanding methodological approach of analyzing the sustainability of urban growth.

Figure 7, based on the geographical boundaries of largest urban areas of Lahore and Islamabad/Rawalpindi districts using 85 percent threshold, shows how the growth in these urban area has affected the composition of different categories of cropland areas over which the respective urban area was spread during the last decade. In the case of Lahore, the green line shows how the 'urban region' covering land areas containing croplands between one to 14 percent of the total urban area has changed. It can be seen from the graph that in both years 2000 and 2011, approximately 40 percent of the urbanized area was covering land area that contained cropland between one to 14 percent. On the other hand, the portion of urbanized area covering land containing cropland between 26 to 44 percent increased from 14 percent in 2000 to 15 percent in 2011 as shown by the red line. A sustainable urban growth would have resulted in a noticeable increase in the proportions represented by green and red lines. However, we notice an increasing trend for the blue line, which represents those parts of urbanized areas covering land areas containing cropland between 50 to 72 percent.

The urban growth of Islamabad/Rawalpindi – panel (b) of Figure 7 – too has performed poorly in the context of sustainability. From the year 2004 onward, proportions represented by red and blue lines follow an increasing trend. This indicates that the urban region has expanded on area containing a significant size of cropland. On the other hand, the proportion of 'urban region' covering land area containing cropland between one to 26 percent decreased instead of increasing over the years as highlighted by the green line.

(a)



(b)

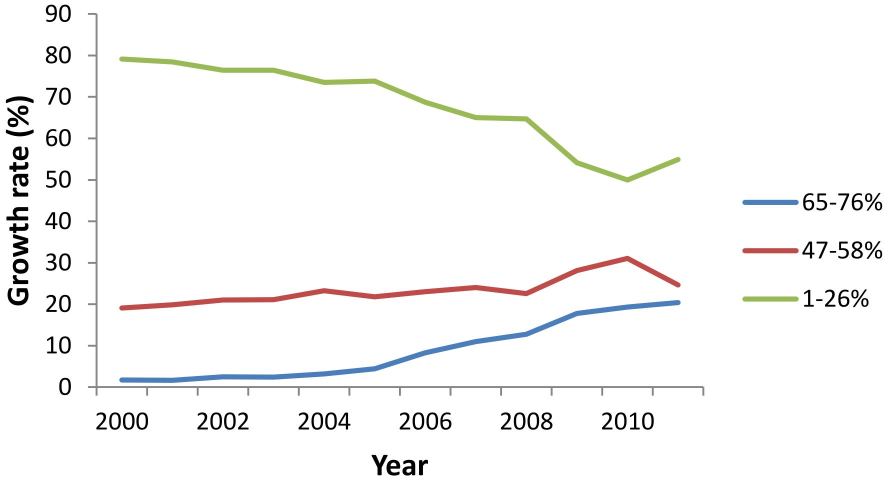


Figure 7: Time series of different proportions of cropland areas for (a) Lahore and (b) Islamabad/Rawalpindi.

4.3. Uncovering Sub-national Economic Growth

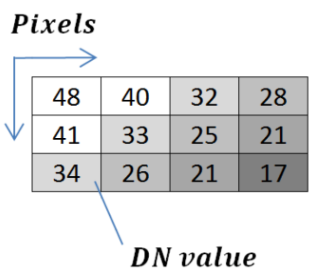
Given the dearth of literature relating NLD with manufacturing output in general, and NLD's relation with economic performance of Pakistan in particular, we begin our analysis of regional growth by estimating the elasticity between value of production and Sum of Lights (SOL):

$$\ln Y_{it} = \delta \ln S_{it} + \mathrm{X}_{\mathrm{it}}\beta + \alpha_i + u_{it} \tag{3}$$

where i and t indexes district and time respectively; X_{it} is matrix containing independent or control variables like employment cost, value of fixed assets at year-end, dummies controlling for province, capital city and time-effect; α_i is the unobserved time-invariant district-effect; u_{it} is disturbance term; $\ln Y_{it}$ refers to log of value of production; and $\ln S_{it}$ is the log of Sum of lights (SOL).

The SOL is derived from the NLD images. The mechanics of calculating SOL are simple: using the appropriate shapefile[10] and choosing the region of interest, the "Zonal Statistics as Table" tool in ArcGIS reports the sum of values attached to pixels comprising geographical area of the target region. The resulting sum is the SOL whose value varies over the years.

For a better understanding of SOL-derivation, Figure 8 shows a region j with simplified rectangular boundary comprising of $i = 12$ pixels. We can consider this region as one of the polygons or districts within Pakistan's shapefile. The right-hand side of Figure 9 shows the calculation of SOL for region j by adding all DN values attached to region j's pixels.



$$SOL_j = \sum_{i=1}^{n} DN_i = 48 + 40 + 32 + 28 + 41 + 33$$
$$25 + 21 + 34 + 26 + 21 + 17$$
$$= 366$$

Figure 8: Composition of pixels in night-time lights image and Illustrative representation of SOL metric derivation.

[10] Shapefile is a digital vector format that stores information about geometric location. Present study uses Pakistan's shapefile as used by Azar et al. (2013)

Equation (3) may be considered as an input-output model in which the dependent variable represents output from the manufacturing sector and explanatory variables as inputs used to produce that output. Thus, inclusion of employment cost and value of fixed assets as control variables constitute contributions from labour and capital – important determinants of manufacturing output (Dias 1991; Upendranadh, Vijayabaskar, and Vinod 1994). We also include regional and temporal dummy variables to control for the possible effects of the unequal distribution of manufacturing activities among the provinces and over different time spans. Table 5 provides details on these variables.

Variables	Definition
$\ln S$	Natural log of SOL
$\ln value\_prod$	Natural log of 'value of production' that comprises value of finished and semi-finished products, receipts for maintenance and work done, value of electricity and goods-for-resale sold, and value of work-in-progress and fixed assets produced as part of capital formation. Value of production is estimated at 2005-06 prices
$\ln emp\_cost$	Natural log of employment cost which includes wages and salaries with the addition of cash and non-cash benefits paid to workers
$\ln value\_fa$	Natural log of 'value of fixed assets at the end of year' that equals value of fixed assets at beginning of year plus investments and additions to fixed assets out of own production less sale of fixed assets during the year. Value of fixed assets at the end of year is estimated at 2005-06 prices
$Punjab$	Dummy variable equals 1 if region is Punjab province, 0 otherwise
$Sindh$	Dummy variable equals 1 if region is Sindh province, 0 otherwise
$Balochistan$	Dummy variable equals 1 if region is Baluchistan province, 0 otherwise
KPK	Dummy variable equals 1 if region is Khyber Pakhtonhawan province, 0 otherwise
$Time\_2001$	Dummy variable equals 1 if year of observation is 2000-01, 0 otherwise
$Time\_2006$	Dummy variable equals 1 if year of observation is 2005-06, 0 otherwise
$Time\_2011$	Dummy variable equals 1 if year of observation is 2010-11, 0 otherwise
$Capital\_city$	Dummy variable equals 1 if observed district contains provincial capital or Islamabad, 0 otherwise

Table 5. Definitions of variables used.

We estimate equation (3) by running a random effects (RE) model. Not only does the RE model not allow individual characteristics of cross-sectional units as represented by the varying intercept term to have any correlation with the regressors (Berndt 1991), but also fulfils our requirement of having a multi-level estimation technique given our data structure is hierarchical in nature i. e. cross-sectional units (districts) are lumped together at higher level (provinces). Our choice of the RE model is further confirmed by the Hausman test – results in Table 6.

Test name	*Null hypothesis*	*p-value*	*Decision*
Hausman Test	H_o: RE and FE don't differ significantly	0.381	Failed to reject H_o

Table 6: Result of Specification test.

The estimates for variants of equation (3) are provided in Table 7. The results are based on an unbalanced panel dataset that comprises 69 districts with unequal time-periods. Column 1 of the Table 7 presents results of the simple OLS without controlling for any explanatory variables apart from time effects. While the results are statistically significance at the one percent level, the coefficient value of SOL is quite high and cannot be considered as a reasonable value for elasticity. This is likely due to the exclusion of all other independent variables that affect large-scale value of production.

VARIABLES	(1) $\ln\_value\_pro$	(2) $\ln\_value\_pro$	(3) $\ln\_value\_pro$	(4) $\ln\_value\_pro$	(5) $\ln\_value\_pro$	(6) $\ln\_value\_pro$
$\ln\_S$	1.895*** (0.166)	1.634*** (0.224)	0.390*** (0.108)	0.414*** (0.103)	0.411*** (0.135)	0.397*** (0.136)
$\ln\_emp\_cos$			0.782*** (0.0315)	0.551*** (0.0538)	0.535*** (0.0532)	0.527*** (0.0545)
$\ln\_value\_fa$				0.219*** (0.0425)	0.224*** (0.0423)	0.228*** (0.0429)
KPK					-1.404*** (0.430)	-1.339*** (0.441)
$Punjab$					-1.245*** (0.440)	-1.153** (0.461)
$Sindh$					-1.263*** (0.441)	-1.176** (0.459)
$Capital\_city$						0.175 (0.266)
$Time\_2001$	-0.121 (0.265)	-0.113 (0.129)	0.0302 (0.0717)	-0.00323 (0.0679)	-0.00549 (0.0684)	-0.00612 (0.0686)
$Time\_2006$	0.550** (0.249)	0.417*** (0.127)	0.290*** (0.0700)	0.153** (0.0709)	0.151** (0.0727)	0.150** (0.0729)
$Time\_2011$	0.0250 (0.290)	0.100 (0.146)	0.204** (0.0807)	0.197*** (0.0762)	0.193** (0.0769)	0.191** (0.0773)
$Constant$	2.416 (1.802)	5.301** (2.361)	3.212*** (1.012)	2.822*** (0.970)	4.306*** (1.263)	4.418*** (1.272)
Observations	202	202	202	201	201	201
R-squared	0.399	0.196	0.749	0.778	0.780	0.780
Number of districts		69	69	69	69	69

Note: ***Significant at 1% level,
**Significant at 5% level,
*Significant at 10% level,
Standard errors in parenthesis

Table 7. Regression results for model or equation (3).

Columns (2) through (6) show the results of the RE model. Column (2) repeats the regression specification of equation (3) and reports the elasticity between value of production and SOL to be 1.634 – lower than estimated value using OLS technique. We can argue that this decline in the coefficient value between columns (1) and (2) is because the RE estimation controls for within and between-district effects. Column (3) takes the RE model further by including 'employment cost'. The coefficient of SOL retains its statistical significance at one percent but decreases in magnitude to 0.39. This decrease in coefficient value indicates that the effect of SOL was being over-estimated before. Column (4) takes one-step more and adds 'value of fixed asset at end-year' as additional regressor. The resulting coefficient of SOL decreases further indicative of fall in explanatory power of SOL. However, the decrease is minor. Columns (5) and (6) control for province and provincial capital city effects respectively. Inclusion of these dummy variables is a good way to check the robustness of regression results. While the parameter estimate of SOL experiences a fall in statistical significance and magnitude, the changes are not considerable enough to harm the interpretation of SOL's coefficient estimate.

The aim of empirical results of Table 7 was to uncover the relationship between LSM production value and NLD in terms of an elasticity that can be used to gauge the economic growth (indexed by LSM growth) at various sub-national levels of Pakistan. We quantify this elasticity by differentiating equation (3) with respect to $\ln SOL$ and equate it to the approximate value of δ – coefficient as estimated in Table 7. The targeted elasticity is as follows:

$$\frac{\partial \ln value\_prod}{\partial \ln SOL} \approx 0.4 \tag{4}$$

Figure 10 analyses the difference between annual growth rates of the value of production and manufacturing component of Gross National Product (GNP) and the SOL-estimated LSM value of production growth for Pakistan region as a whole. The dotted line in the graph is the 45^{o} line that indicates the locus of points at which both axis-values are same. As seen from the figure, a cluster of points is situated between the origin and vertical distance of seven showing similarity between two growth-rates of

LSM. However, it is also visible from the figure that GNP growth rates of LSM are consistently higher than the growth rates as estimated using NLD[11]. Thus, figure 10 presents evidence that Pakistan's official manufacturing growth rates may well be over-estimates particularly when we consider the high elasticity between SOL and the value of production and manufacturing.

Another important aspect of Figure 9 that demands attention is the identification of three outlier points that relate to fiscal years 2003-04, 2004-05 and 2005-06. GNP growth of LSM during these four years was in double digits. In year 2003-04, the LSM sector received considerable assistance from the banking system and registered the highest growth rate of real fixed investment. Keeping these facts in view and considering the corresponding lower values of satellite-estimated growth rates, we can conclude that the manufacturing growth that occurred during these three consecutive years was likely a one-time event as it was not associated with growth in infrastructure or other efficiency enhancing technology that would be captured using NLD.

[11] Growth rates of SOL of whole Pakistan region since 1992-93 to 2011-12 were multiplied by 0.4 (elasticity) to derive NLD-estimated LSM growth rates

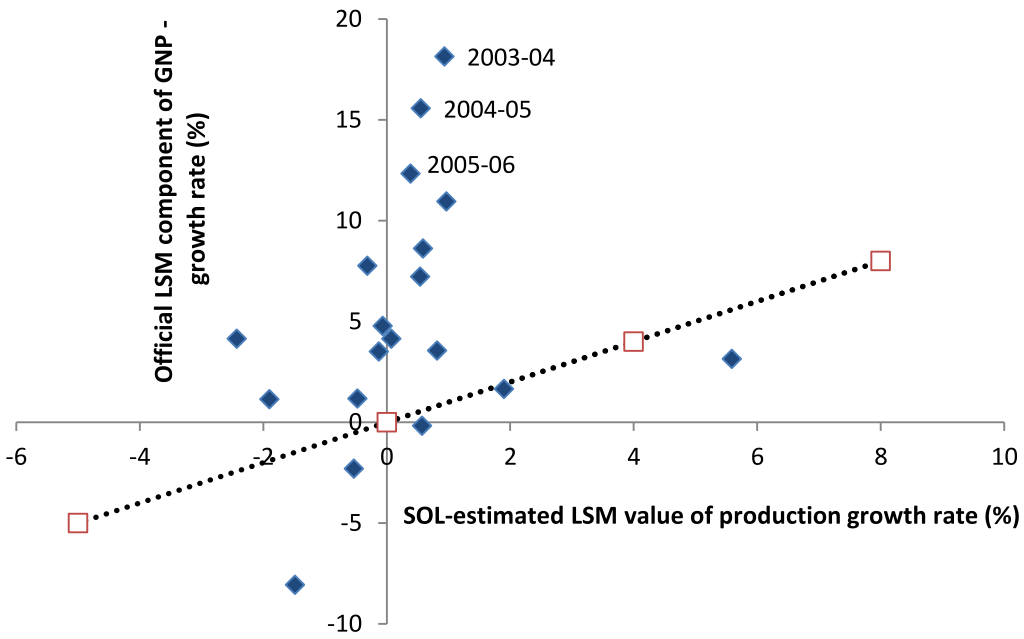
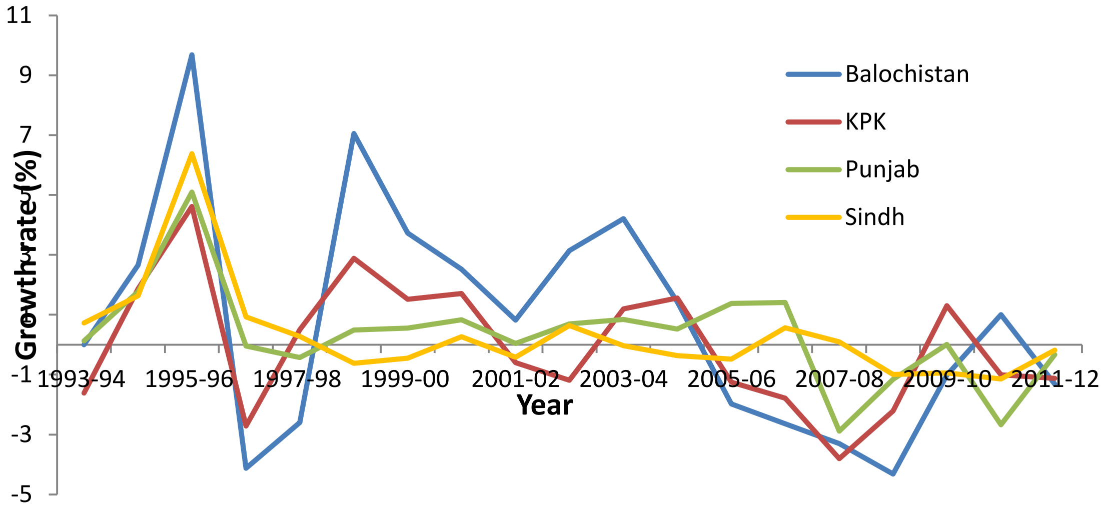


Figure 9: Comparison of GNP and NLD-measured growth rates of Pakistan's LSM.

The highest tier of administrative divisions of Pakistan is the province; we focus our regional growth analysis on this sub-national unit. Figure 10 sketches the annual growth rates of CMI value of production for the four provinces of Pakistan as estimated using SOL-elasticity and provides statistical information relating these growth rates. From the graph, it is clear that all four provinces are registering a decreasing trend over the 19-year period.

The tabular section of the Figure 10 sheds a little more light on the regional growth among the provinces. Baluchistan province has registered highest average annual growth rate of 0.78 percent for LSM production value whereas the most developed province of Punjab had a corresponding growth rate of 0.32 percent. This might seem confusing but an examination of initial values of LSM production value and SOL can help clarify this perplexity. At the start of 1990s, Baluchistan had only four and three percent shares in total SOL and CMI production value respectively. On the other hand, the corresponding shares for Punjab were 63 and 4 percent respectively. It appears then that convergence is at play here with poorer provinces catching-up with the more advanced. However, given the initial economic base of Baluchistan and the estimated growth rate, this catching-up will likely take up several years. Figure 9 also indicates the dependence of variability in CMI production value on the beginning size of economy.



	SOL_share (%) 1992-93	CMI_share (%) 1990-91	Average CMI growth rate	SD CMI growth rate
Punjab	63	46	0.32	1.7
Sindh	22	51	0.31	1.6
KPK	11	N/A	0	2.1
Balochistan	4	3	0.78	3.8

Figure 10: Time trend of NLD-measured growth rates of provincial LSM production value and corresponding summary statistics.

Yet another sub-national unit of Pakistan is the ‘district’ that forms the third tier of the administrative division. Figure 11 is a scatter diagram of the standard deviation of annual growth rates (NLD-estimated) of CMI of 44 districts of Pakistan against real GDP per capita of these districts in the last census year of 1998. For analytical purposes, we have used real GDP per capita as an indicator of the size of economy. The scatter diagram shows a very interesting trend: districts that reported higher initial values of real GDP per capita experienced lower variability in the annual growth rate of CMI over the 19 year period analysed here. If we consider level of variability in LSM production value as an indicator of efficiency then our current analysis suggests that income or the size of the (local) economy is one of the major determinant of efficiency in manufacturing sector at the district level in Pakistan.

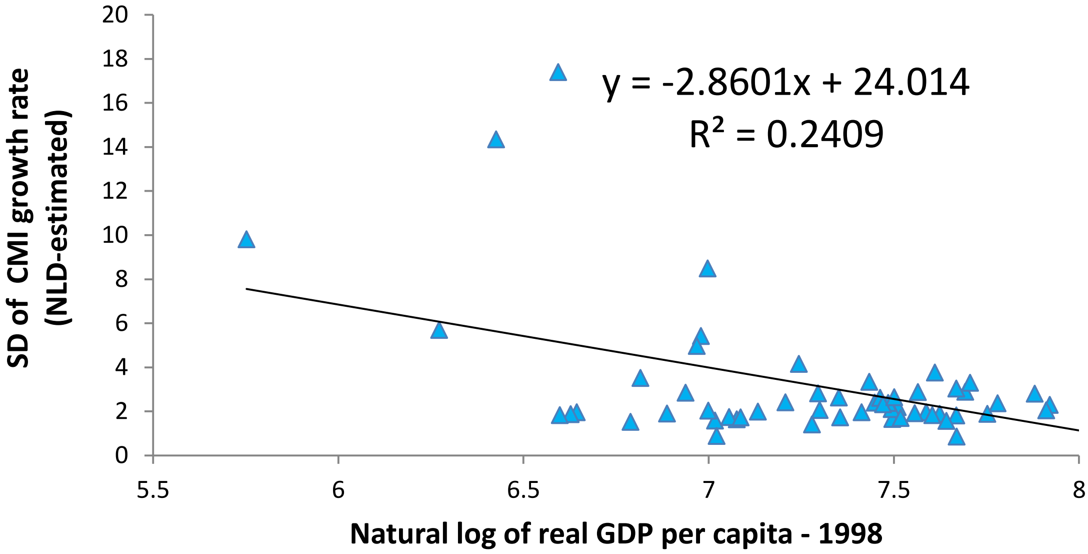


Figure 11: District-level scatterplot of dispersion in NLD-measured LSM growth rates against size of economy

5. Conclusion

This study was undertaken with the goal of analysing the utility of nighttime lights data in estimating and upgrading the economic statistics of various sub-national units. The use of regression model and statistical analysis has indeed shown that this unique data has the capacity to help improve the economic statistics of developing countries like Pakistan. Although nighttime lights correlate with numerous variables as indicated in the review of previous literature, we have explored its relationship in the context of large-scale manufacturing. Furthermore, we also make use of nighttime lights data to understand the dynamics of urban sprawl as well as the relationship of urbanization with cropland loss.

One of the main results of our analysis is that an increase in the luminosity metric of one percent is associated with an increase in LSM value of production increased of roughly 0.4 percent ceteris paribus. The underlying theory then is that geographic economic development is related with the output produced. Overall, the findings of this paper are consistent with Henderson et al., (2012) and, Bhandari and Roychowdhury (2011) as far as connection of nighttime lights and sub-national economic activity is concerned.

Turning to our work on urbanization, we extracted the geographical boundaries of the largest urban areas of Punjab province and with the help of regression techniques, concluded that the largest urban regions are growing at an average annual rate of approximately two percent at the 85 percent threshold used for detection of urban boundaries. While overall Pakistan then seems to be seeing an urban expansion rate that is lower than the rate found by Gibson et al. (2015) for the case of India – Gibson et. al. conclude that one million-plus populated agglomerations are expanding at an average annual rate of 2.4 percent, we did however find similar and higher expansion rates for specific regions.

The Islamabad/Rawalpindi urban area that has an estimated combined population of around 4.5 million was found to be growing at annual rate of four percent – highest among the largest urban areas of Punjab. The second highest annual expansion rate (2.6 percent) was registered by Lahore's urban region.

With respect to the sustainably of urban expansion, we analysed the growth of urban regions in the context of changes in the composition of different categories of cropland – a novel approach. Our pictorial investigation and numerical data indicated the urban growth of Islamabad/Rawalpindi to be the least sustainable.

The results of this study then not only indicate the various ways in which NLD may be used to fill data gaps and enrich the analysis of economic phenomena but it also motivates many new research agendas for fields that are hampered by significant data gaps. Some examples include gauging village electrification, or disaggregation of energy consumption. Furthermore, our analysis and the results therein carry with them some important notes from a policy perspective. The foremost amongst these is the cannibalization of cropland by urban sprawl that has significant welfare implications that must be addressed if the rapid expansion of these urban centres is to be sustained. Similarly, the differential growth rates across regions highlighted through our analysis merit a deeper study if policy is to have any significant and lasting impact on the slow rate of convergence found in our work.

Acknowledgements

The author would like to thank Mushtaq Khan and Abid Burki at Lahore University of Management Sciences (LUMS) for their useful comments and constructive suggestions throughout the time present research work was in process.

The author is also deeply grateful to Tilottama Ghosh – Research Associate at Cooperative Institute for Research in Environmental Sciences (CIRES) and Feng Chi Hsu – Associate Researcher at University of Colorado Boulder. Execution of procedures relating inter-calibration and other issues pertaining nighttime lights data would not have been possible without the help of these two persons.

All remaining errors are mine.

References

Amaral, Silvana, Gilberto Câmara, Antônio Miguel Vieira Monteiro, José Alberto Quintanilha, and Christopher D. Elvidge. "Estimating population and energy consumption in Brazilian Amazonia using DMSP night-time satellite data." *Computers, Environment and Urban Systems* 29, no. 2 (2005): 179-195.

Alesina, Alberto F., Stelios Michalopoulos, and Elias Papaioannou. "Ethnic inequality." *NBER Working Paper* w18512 (2012).

Arby, Muhammad Farooq. "Some issues in the national income accounts of Pakistan (rebasing, quarterly and provincial accounts and growth accounting)." (2008).

Azar, Derek, Ryan Engstrom, Jordan Graesser, and Joshua Comenetz. "Generation of fine-scale population layers using multi-resolution satellite imagery and geospatial data." *Remote Sensing of Environment* 130 (2013): 219-232.

Bell, Andrew, and Kelvyn Jones. "Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data." *Political Science Research and Methods* 3, no. 01 (2015): 133-153.

Bengali, Kaiser, and Mahpara Sadaqat. "Provincial Accounts of Pakistan: Methodology and Estimates 1973-2000." *SPDC WP* 5 (2005): 2005.

Berndt, Ernst R. "The Practice of Econometrics: Classic and Contemporary. 1991." *Reading: Addison-Wesley* 1.

Bhatta, Ba, S. Saraswati, and D. Bandyopadhyay. "Urban sprawl measurement from remote sensing data." *Applied geography* 30, no. 4 (2010): 731-740.

Bhandari, Laveesh, and Koel Roychowdhury. "Night lights and economic activity in India: A study using DMSP-OLS night time images." *Proceedings of the Asia-Pacific advanced network* 32 (2011): 218-236.

Cabral, Pedro, Gabriela Augusto, Mussie Tewolde, and Yikalo Araya. "Entropy in urban systems." *Entropy* 15, no. 12 (2013): 5223-5236.

Chang, Shu-Hwa, and Liang-Chou Huang. "The nexus of finance and GDP growth in Japan: Do real interest rates matter?." *Japan and the World Economy* 22, no. 4 (2010): 235-242.

Colantoni, Andrea, Efstathios Grigoriadis, Adele Sateriano, Giuseppe Venanzoni, and Luca Salvati. "Cities as selective land predators? A lesson on urban growth, deregulated planning and sprawl containment." *Science of the Total Environment* 545 (2016): 329-339.

Dias, Sriyani. "Factors affecting the productivity of manufacturing sector in Sri Lanka: a spatial analysis." *GeoJournal* 23, no. 2 (1991): 113-120.

Doll, Christopher NH. "CIESIN thematic guide to night-time light remote sensing and its applications." *Center for International Earth Science Information Network of Columbia University, Palisades, NY* (2008).

Econometrics, Introductory. "A Modern Approach." *Jeffrey M. Wooldridge/South-Western* (2003).

Ellis, Peter, and Mark Roberts. *Leveraging Urbanization in South Asia: Managing Spatial Transformation for Prosperity and Livability*. World Bank Publications, 2015.

Elvidge, Christopher D., Daniel Ziskin, Kimberly E. Baugh, Benjamin T. Tuttle, Tilottama Ghosh, Dee W. Pack, Edward H. Erwin, and Mikhail Zhizhin. "A fifteen year record of global natural gas flaring derived from satellite data." *Energies* 2, no. 3 (2009): 595-622.

Elvidge, Christopher D., Edward H. Erwin, Kimberly E. Baugh, Daniel Ziskin, Benjamin T. Tuttle, Tilottama Ghosh, and Paul C. Sutton. "Overview of DMSP nightime lights and future possibilities." In *2009 Joint Urban Remote Sensing Event*, pp. 1-5. IEEE, 2009.

Farooq, S., and S. Ahmad. "Urban sprawl development around Aligarh city: a study aided by satellite remote sensing and GIS." *Journal of the Indian Society of Remote Sensing* 36, no. 1 (2008): 77-88.

Firebaugh, Glenn. "Structural determinants of urbanization in Asia and Latin America, 1950-1970." *American Sociological Review* (1979): 199-215.

Wooldridge Jeffrey, M. "Econometric analysis of cross section and panel data." *Cambridge, MA: Massachusetts Institute of Technology* (2002).

Greene, William H. *Econometric analysis*. Pearson Education India, 2003.

Gibson, John, Geua Boe-Gibson, and Glen Stichbury. "Urban land expansion in India 1992–2012." *Food Policy* 56 (2015): 100-113.

Gujarati, D. N. "Basic Econometrics Fourth Edition McGraw Hill Gujarati, DN,(2003). Basic Econometrics." (2003).

Henderson, J. Vernon, Adam Storeygard, and David N. Weil. "Measuring economic growth from outer space." *The American Economic Review* 102, no. 2 (2012): 994-1028.

Hill, R. Carter, William E. Griffiths, and Guay C. Lim. *Principles of econometrics*. Vol. 5. Hoboken, NJ: Wiley, 2008.

Hsiao, Frank ST, and Mei-Chu W. Hsiao. "FDI, exports, and GDP in East and Southeast Asia—Panel data versus time-series causality analyses." *Journal of Asian Economics* 17, no. 6 (2006): 1082-1106.

Imhoff, Marc L., William T. Lawrence, David C. Stutzer, and Christopher D. Elvidge. "A technique for using composite DMSP/OLS “city lights” satellite data to map urban area." *Remote Sensing of Environment* 61, no. 3 (1997): 361-370.

Kasarda, John D., and Edward M. Crenshaw. "Third world urbanization: Dimensions, theories, and determinants." *Annual Review of Sociology*(1991): 467-501.

Khramov, Vadim, and John Ridings Lee. "The Economic Performance Index (EPI): an Intuitive Indicator for Assessing a Country's Economic Performance Dynamics in an Historical Perspective." (2013).

Levin, Noam, Kasper Johansen, Jorg M. Hacker, and Stuart Phinn. "A new source for high spatial resolution night time images—The EROS-B commercial satellite." *Remote Sensing of Environment* 149 (2014): 1-12.

Li, Xuecao, Peng Gong, and Lu Liang. "A 30-year (1984–2013) record of annual urban dynamics of Beijing City derived from Landsat data." *Remote Sensing of Environment* 166 (2015): 78-90.

Li, Ying-ying, Hao Zhang, and Wolfgang Kainz. "Monitoring patterns of urban heat islands of the fast-growing Shanghai metropolis, China: Using time-series of Landsat TM/ETM+ data." *International Journal of Applied Earth Observation and Geoinformation* 19 (2012): 127-138.

Liu, Zhifeng, Chunyang He, Qiaofeng Zhang, Qingxu Huang, and Yang Yang. "Extracting the dynamics of urban expansion in China using DMSP-OLS nighttime light data from 1992 to 2008." *Landscape and Urban Planning* 106, no. 1 (2012): 62-72.

Ma, Ting, Yuke Zhou, Chenghu Zhou, Susan Haynie, Tao Pei, and Tao Xu. "Night-time light derived estimation of spatio-temporal characteristics of urbanization dynamics using DMSP/OLS satellite data." *Remote Sensing of Environment* 158 (2015): 453-464.

Ma, Ting, Chenghu Zhou, Tao Pei, Susan Haynie, and Junfu Fan. "Quantitative estimation of urbanization dynamics using time series of DMSP/OLS nighttime light data: A comparative case study from China's cities." *Remote Sensing of Environment* 124 (2012): 99-107.

Ma, Qun, Chunyang He, Jianguo Wu, Zhifeng Liu, Qiaofeng Zhang, and Zexiang Sun. "Quantifying spatiotemporal patterns of urban impervious surfaces in China: An improved assessment using nighttime light data."*Landscape and Urban Planning* 130 (2014): 36-49.

Pandey, Bhartendu, and Karen C. Seto. "Urbanization and agricultural land loss in India: Comparing satellite estimates with census data." *Journal of environmental management* 148 (2015): 53-66.

Purohit, Brijesh C. "HEALTH CARE SYSTEM EFFICIENCY: A SUB-STATE LEVEL ANALYSIS FOR ORISSA (INDIA)." *Review of Urban & Regional Development Studies* 28, no. 1 (2016): 55-74.

Ramachandra, T. V., Bharath H. Aithal, and Durgappa D. Sanna. "Insights to urban dynamics through landscape spatial pattern analysis." *International Journal of Applied Earth Observation and Geoinformation* 18 (2012): 329-343.

Ramankutty, Navin, Amato T. Evan, Chad Monfreda, and Jonathan A. Foley. "Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000." *Global Biogeochemical Cycles* 22, no. 1 (2008).

Roychowdhury, Koel, S. J. Jones, C. Arrowsmith, and K. Reinke. "Night-Time Lights and Levels of Development: A Study Using DMSP-OLS Night-Time Images at the Sub-National Level." In *Proceedings of the XXII ISPRS Congress, Melbourne, VIC, Australia*, vol. 25, pp. 93-98. 2012.

Sahoo, Biresh K., and Debashis Acharya. "Constructing macroeconomic performance index of Indian states using DEA." *Journal of Economic Studies* 39, no. 1 (2012): 63-83.

Sehrawat, Madhu, and A. K. Giri. "The role of financial development in economic growth: empirical evidence from Indian states." *International Journal of Emerging Markets* 10, no. 4 (2015): 765-780.

Sheng, Yu, Xunpeng Shi, and Dandan Zhang. "Economic growth, regional disparities and energy demand in China." *Energy Policy* 71 (2014): 31-39.

Sridhar, Kala Seetharam. "Determinants of city growth and output in India."*Review of Urban & Regional Development Studies* 22, no. 1 (2010): 22-38.

Stow, Douglas A., John R. Weeks, Hsiao-chien Shih, Lloyd L. Coulter, Harry Johnson, Yu-Hsin Tsai, Andrew Kerr, Magdalena Benza, and Foster Mensah. "Inter-regional pattern of urbanization in southern Ghana in the first decade of the new millennium." *Applied Geography* 71 (2016): 32-43.

Suribabu, C. R., J. Bhaskar, and T. R. Neelakantan. "Land use/cover change detection of Tiruchirapalli City, India, using integrated remote sensing and GIS tools." *Journal of the Indian Society of Remote Sensing* 40, no. 4 (2012): 699-708.

Tripathi, Sabyasachi. "Do large agglomerations lead to economic growth? Evidence from urban India." *Review of Urban & Regional Development Studies* 25, no. 3 (2013): 176-200.

Upendranadh, C., M. Vijayabaskar, and Vinod Vyasulu. "Industrial growth and structure: analysis of manufacturing sector in Karnataka." *Economic and Political Weekly* (1994): M157-M164.

Urbanovský, Tomáš. "Interconnection of Interest Rate, Price Level, Money Supply and Real GDP: The Case of the Czech Republic." *Procedia-Social and Behavioral Sciences* 220 (2016): 531-540.

Wakode, Hemant Balwant, Klaus Baier, Ramakar Jha, and Raffig Azzam. "Analysis of urban growth using Landsat TM/ETM data and GIS—a case study of Hyderabad, India." *Arabian Journal of Geosciences* 7, no. 1 (2014): 109-121.

Williams, David, Charles AE Goodhart, and David H. Gowland. "Money, income, and causality: the UK experience." *The American Economic Review*66, no. 3 (1976): 417-423.

Wooldridge, Jeffrey M. *Econometric analysis of cross section and panel data*. MIT press, 2010.

Wu, Qiong, Hong-qing Li, Ru-song Wang, Juergen Paulussen, Yong He, Min Wang, Bi-hui Wang, and Zhen Wang. "Monitoring and predicting land use change in Beijing using remote sensing and GIS." *Landscape and urban planning* 78, no. 4 (2006): 322-333.

Yan, Tingting, Jinxia Wang, and Jikun Huang. "Urbanization, agricultural water use, and regional and national crop production in China." *Ecological Modelling* 318 (2015): 226-235.

Yang, Kaizhong, and Ying Xu. "Regional differences in the development of Chinese small and medium-sized enterprises." *Journal of Small Business and Enterprise Development* 13, no. 2 (2006): 174-184.

Yue, Wenze, Jiabin Gao, and Xuchao Yang. "Estimation of gross domestic product using multi-sensor remote sensing data: A case study in Zhejiang province, East China." *Remote Sensing* 6, no. 8 (2014): 7260-7275.

Zeng, Chen, Yaolin Liu, Alfred Stein, and Limin Jiao. "Characterization and spatial modeling of urban sprawl in the Wuhan Metropolitan Area, China."*International Journal of Applied Earth Observation and Geoinformation* 34 (2015): 10-24.

Zhang, Jialong, Margaret Kalacska, and Sarah Turner. "Using Landsat Thematic Mapper records to map land cover change and the impacts of reforestation programmes in the borderlands of southeast Yunnan, China: 1990–2010." *International Journal of Applied Earth Observation and Geoinformation* 31 (2014): 25-36.

Zhang, Qingling, and Karen C. Seto. "Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data." *Remote Sensing of Environment* 115, no. 9 (2011): 2320-2329.

Zhang, Qian, and Karen C. Seto. "Can night-time light data identify typologies of urbanization? A global assessment of successes and failures."*Remote Sensing* 5, no. 7 (2013): 3476-3494.

Zhao, Naizhuo, Yuyu Zhou, and Eric L. Samson. "Correcting incompatible DN values and geometric errors in nighttime lights time-series images."*IEEE Transactions on Geoscience and Remote Sensing* 53, no. 4 (2015): 2039-2049.