EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

University of Tübingen
Working Papers in
Business and Economics

No. 152

Re-evaluating RCTs with nightlights - an example from biometric smartcards in India

by

Merlin Stein

Faculty of Economics and Social Sciences
www.wiwi.uni-tuebingen.de

https://publikationen.uni-tuebingen.de/xmlui/handle/10900/95156

Re-evaluating RCTs with nightlights
- an example from biometric smartcards in India

Merlin Stein‡

November 2021

Abstract

Satellite data and randomized controlled trials (RCTs) are a powerful combination for analyzing causal effects beyond traditional survey-based indicators. The usage of remotely collected data for evaluating RCTs is cost-effective, objective and possible for anyone with treatment assignment data. By re-evaluating one of the largest RCTs - the smartcard intervention of Muralidharan et al. (2016) covering 20 million people - with Indian nighttime luminosity, this paper finds that nightlights as a specific type of satellite data likely often are too noisy to evaluate RCTs.

Building upon a post-treatment and a Difference-in-Differences approach, we do not find any statistically significant effects of the biometric smartcards on nightlights, contrasting Muralidharan et al. (2017)'s results of higher income level in treated areas. This can be mainly explained either with the noisiness-caused inability of nightlights to specifically capture economic effects or the absence of an increased economic activity due to a simple redistributive effect of the intervention. The former is more likely when looking at GDP implications of the noisiness in the luminosity data. Per head estimates, sensitivity checks for spillovers, subdistrict-level instead of village-level observations and different time-wise aggregations of nightlight data do not lead to changed results.

Although limited with nightlights, nonetheless, the potential for re-evaluating RCTs with satellite data in general is enormous in three ways: (1) For confirming claimed treatment effects, (2) to understand additional impacts and (3) for cost-effectively understanding long-term impacts of interventions. Using daytime imagery for analyzing RCTs is a promising direction for future research.

Keywords: RCT, randomized, nightlight, daylight, satellite, remote-sensing, nighttime luminosity, India, Census, Muralidharan, state capacity, GDP and nightlights
JEL classification: C33, C81, C93, E01, H53, H55, I32, I38, J65, O47, R12

‡Independent Researcher, Formerly University of Oxford and University of Tuebingen. stein.merlin@gmail.com. I am grateful for discussions with Martin Biewen and Pascal Erhardt, and comments from Stefan Glaisner, Lydia Hanlon, Michael Kaiser, Justin Sundermann and Julius Ulrich.

1 Introduction

In the last decades, the landscape of rigorous evaluations in the social sciences has been shaped strongly by randomized controlled trials (Jamison (2017)). Especially in Development Economics, RCTs are broadly deployed because of their high internal validity. In light of Duflo, Banerjee and Kremer's Nobel prize in 2019 for their pioneering work in this field, this paper argues that there is much potential in combining the power of RCTs for understanding causalities with low-cost, objective satellite data.

One of the standard forms of satellite imagery is the measurement of light intensity at night. Nightlights have been found to proxy economic activity pretty well (Henderson et al. (2012)). During the last ten years, an increasing amount of authors provided proof for the plausibility of nightlights as a proxy for GDP throughout the world. However, there is limited research dedicated to nightlights-based impact evaluations of policies.

To the best of the author's knowledge, the potential of RCTs combined with satellite data has not yet been used and analyzed, although both are a potent tool for the evaluation of interventions. RCTs are seen as one of the best methods to identify causality, as the technique outscores other approaches in unconfoundedness of treatment assignment, bearing in mind limitations, e.g., regarding external validity (Deaton and Cartwright (2018).

In comparison to GDP, nightlights have two core advantages. First, they can capture economic activity in regions with lower data quality and high unreported informal economic activity. Second, nightlights are available in higher granularity in real-time (publicly with some delays).

In this paper, nighttime luminosity data (in the following: nightlights or simply luminosity data) are used the first time to re-evaluate a randomized controlled trial, analyzing one of the most extensive randomized controlled policy interventions: The implementation of smartcards for the Indian national employment scheme (NREGS) and pension scheme (SSP) in the state of Andhra Pradesh between 2010 and 2012 by Muralidharan et al. (2016) (in the following "MNS").

The goal of this piece of research is to evaluate whether economic impacts indicated by MNS can be measured at the village or subdistrict level with nightlights, as GDP statistics do not exist beyond district level. This exemplary analysis of the smartcard RCT reveals whether nightlights as a specific form of satellite data enable valuable knowledge gains through impact measurements when combined with RCTs. Based on two main specifications, this paper does not find any statistically significant [1] effects of the smartcard intervention on nightlights, which does not change when including spillovers. Thus the suitability of nightlights for re-evaluating RCTs might be limited.

After a brief overview of related literature (Section 2), the smartcard RCT of MNS is examined with quarterly nightlight data (Section 3 and 4). Therefore, the post-treatment values of villages in control and treatment subdistricts are compared in a regression model that is as similar as possible to the original one used by MNS. An additional Difference-in-Differences estimate provides further discussion ground, while a useful interpretation of the effect size is made possible through a comparison with official GDP values. In the end, this paper summarizes critical learnings from the empirical analysis to finally answer the question of whether the combination of RCTs and nightlights as an more and more often-used type of satellite data is fruitful for evaluating stated effects of economic interventions(Section 5).

2 Literature overview

This chapter summarizes fundamental results from previous research on nightlights, and studied relationships between nightlights and GDP in the first part. In the second part, we

[1] From here on and in the following chapters, the term "significant" always refers to statistical significance, defined as a p-value below 5%.

provide an overview of relevant results of the examined smartcard RCT by MNS.

2.1 Nightlights in the literature

The use and study of satellite data in the social sciences have its roots in the early 2000s, with diverse applications across socio-economic indicators and analysis factors (e.g., Donaldson and Storeygard (2016)). The kickstart for nighttime luminosity as a widely used proxy for GDP is often contributed to Henderson et al. (2012), who reveal the potential of nightlights to augment GDP observations across the world. Nightlights have been sourced from different satellite programs, most of the studies in the following either use DMSP or VIIRS data.

Weidmann and Schutte (2017) confirm that the common assumption of electricity-driven night-time luminosity increases holds. However, they find nightlights to proxy wealth as a condition for investments in power generators or power grid connecting cables, and economic activity enabling this financial situation. Street lamps and emitted light of manufacturing, farming and domestic buildings are the main drivers of nightlights. As all are correlating with economic activity and wealth - no matter how the mechanism works in individual circumstances - nightlights are a reasonable proxy at least for those two indicators of well-being.[2] Hu and Yao (2019) and others especially investigate the fit of nightlights as a proxy of economic activity, concluding that it might even exceed the accuracy of GDP in regions of high informal activity and lower-income countries. Throughout the economic literature, nightlights are seen as a reasonable proxy for GDP, although there are ongoing discussions whether this is applicable for much smaller administrative units, such as the subdistrict and

[2]Beyond that, Michalopoulos and Papaioannou (2012), Ghosh et al. (2013) and others reveal significant correlations of nightlights and the Human Development Index. The direct connection between high life quality and the emission of light at night can be questioned; however, at least in regions with a relatively low living standard, access to electricity, and increased economic activity are often breaking the barriers towards higher life quality. This paper will mainly refer to economic activity proxied by nightlights, leaving it open to the reader whether to derive an indirect relationship to well-being.

village level as well (Bickenbach et al. (2016)). Additionally, the mapping of nightlights to GDP[3] is statistically different in urban and rural areas. Reasons for this are spillovers between very light urban areas in satellite measurements, and low-light agricultural activities in rural regions (Otchia and Asongu (2019); Wu et al. (2013)). The analysis used in this paper only looks at rural areas, so that the rural-urban bias, that is typical for nightlights, does not play a role.

In the Indian context, Bhandari and Roychowdhury (2011) and Prakash et al. (2019) conclude that nighttime luminosity is a reasonable indicator for economic activity and strongly correlates with state- and district-level GDP, bearing the potential to bridge the gap of missing subdistrict and lower level GDP estimates. Ghosh et al. (2010) harness remotely sensed light data to state that India's informal economy and remittances are very underestimated in official GDP estimates.[4]

Research applying nightlights for the analysis of large-scale governmental and non-governmental interventions is available in limited quantity when compared to other research using nightlight data. Nonetheless, several non-randomized studies have been conducted.[5]

[3]Note that in this paper, for the sake of simplicity, the term GDP is used for official data of economic activity aggregates on non-domestic levels too, which is then specified.

[4]Note: In contrast to many studies of other countries (e.g., Bundervoet et al. (2015)), Bhandari and Roychowdhury (2011) find nightlight-based estimates for GDP in Indian agricultural areas to be positively biased compared to urban areas.

[5]For instance, Mitnik et al. (2018) quantify the effects of transport infrastructure on local GDP in Haiti by exploiting the differential timing of rehabilitation projects and deploying a Difference-in-Differences panel fixed effects model but lacking randomization. Corral et al. (2018) use nightlights in several impact evaluations of local interventions. In event studies of the 2015 earthquake in Nepal and the 2016 demonetization in India, Beyer et al. (2018) reveal the strong effect on regions with high informal activity. One of the available impact evaluations with nightlights in India is conducted by Asher and Novosad (2020), publishing that road construction does not necessarily lead to more economic opportunities for the rural population.

2.2 The smartcard RCT

In a seminal RCT, MNS investigate the randomized roll-out of biometric smartcards and reorganization of the payment process for the world's biggest rural employment program (the National Rural Employment Guarantee Scheme (NREGS)) and the pension scheme SSP in the Indian state Andhra Pradesh. NREGS assures every rural household 100 days of paid work[6]. Due to implementation and leakage issues, the government of Andhra Pradesh moved from a leakage prone cash-based money transfer through the different administrative levels to a more direct money transfer, with fingerprint identification for money collection using so-called smartcards. By randomizing the roll-out of the smartcard program at the subdistrict level, the researchers estimate the intent-to-treat effect (ITT) of the newly available money transfer option. ITT estimates are necessary as conversion took time, with 68% of villages using smartcard-enabled NREGS payments after two years, at the endline in 2012. To prevent access problems caused by technical issues, smartcards were not mandatory for collecting wages.[7]

The smartcard intervention introduced several fundamental changes to the payment process: An organization not connected to the NREGS work reports was responsible for the payment process, the pay-out-point was closer to the village and fingerprint identification was introduced for payment collection. This lead to three main impacts. First, they observe a faster and more predictable payment process for beneficiaries. It took 22 minutes less to collect payment (-20% relative to the control mean), and benefits were collected 5.8 to 10 days sooner (- 17-29%) compared to the control group. Second, the proportion of households that reported working on NREGS increased by 7.1 percentage points (+17%), profiting of the reduction of quasi-ghost beneficiaries.

[6] At appr. $1.5-2 per day; mostly for local infrastructure projects

[7] This was an important strategy to guarantee inclusion when compared to similar programs (Muralidharan et al. (2020)).

Third, leakage[8] along the payment chain from the state government to beneficiaries decreased by 12.7 percentage points (-41%). Thus, more money reached the beneficiaries, amounting to a 24% increase in NREGS-related household earnings, with similar results for SSP. This significant increase in household income holds for every income class registered for NREGS or SSP. Total leakage reduction estimates equal $41.7 million per year, mainly driven by exacerbating overreporting and underpayment through the new payment process and the smartcards. Even within the same subdistrict, the government spending for the programs remained constant (there is an exogenous budget cap for planning purposes). This suggests substantial benefits of implementation improvements. The cost of the smartcard roll-out and operation for both schemes amounts to $6.3 million, while time savings for beneficiaries aggregate to $4.5 million per year, based on assuming an always available reservation wage of roughly $1.7/day. Even though the latter assumption seems naive, the leakage reduction by far exceeds implementation and operation costs.

However, it is crucial to keep in mind that a leakage reduction is not merely an increase in received cash without a group losing parts of their income. Instead, the smartcard program led to a redistribution of transfers from corrupt officials to beneficiaries. Although knowledge on specific points of leakage is limited, the highest leakage is probably at the local village/close-by villages level. So-called Field Assistants usually record attendance of the village-level projects and formerly collected payments from a nearby post office [9] to then distribute it to workers. Overreporting work or inventing ghost-beneficiaries could easily allow Field Assistants to harvest payments meant for beneficiaries. With fingerprint authentication, Field Assistants were restricted in accessing payments, and thus needed to

[8]Leakage is defined as the difference between official government data and survey-based beneficiary reports.

[9]According to government data, there is roughly one post office per four villages in India. Assuming one Field Assistant per post office, or at most covering post offices in a radius of 20-40km (1 hour travel by car or motorcycle) seems reasonable.

employ more real villagers to compensate for then missing ghost beneficiaries.[10]

As officials usually are substantially wealthier than rural beneficiaries, the transferred income from "the rich" to "the poor" contributes positively to a utilitarian social welfare function with diminishing returns.

In their follow-up paper, Muralidharan et al. (2017) find large positive market wage developments[11], as a consequence of higher reservation wages through the improved implementation of NREGS. As workers commute to nearby villages (assuming a maximum of a two hour walk of 20 km), which might be in non-treated subdistricts, the wages increase in the surrounding areas as well - the spillover-effects. Including these effects on nearby villages, the total income increase amounted to 10.6%, while non-adjusted 6.5%. Consequently, spillovers are highly relevant when considering the economic effects of the smartcard program. Besides, private sector employment days rose by 20% (spillover adjusted), NREGS employment days by 29%. The former increase is as large as the effect of the initial roll-out of the NREGS program, evaluated by Imbert and Papp (2015). Higher market wages lead to another redistributive effect from the rich to the poor: Landlords seeking employees have little chance but using a higher budget for labor wages. Nevertheless, effects on overall economic activity remain uncertain and thus will be studied in this paper in the following sections.

Beyond, MNS do not report Difference-in-Differences estimates as they cannot find any significant differences in the dependent variables at the baseline at the village level. Their study design is limited to comparing the treatment and non-treatment villages,[12] which is used in the following sections as well. Moreover, this paper builds upon the availability of quarterly data, enabling the estimation of Difference-in-Differences model, taking into ac-

[10] Leakage at higher levels might play a role as well, but is assumed to be of smaller size. Offices like the subdistrict offices could potentially manipulate attendance reported by the Field Assistants and additionally invent workers themselves, but need to work together with officials on the way or with Field Assistants.

[11] Consumer goods prices did not significantly increase, suggesting real income gains.

[12] There are non-studied subdistricts working as a time buffer that received smartcards mostly in 2011

count historic trends reasonably. The quarterly data also allows addressing possible caveats of MNS's strong focus on the NREGS-heavy month of June, triggered by surveys purposely taking place directly afterwards.

In summary, this re-evaluation of the smartcard RCT will provide insights on how nightlight data might depict additional and verify stated effects. Specifically, the impact of the smartcard intervention on the overall economic activity beyond income in studied NREGS-heavy periods in treated villages or subdistricts is still unclear, and cannot be derived from GDP statistics as those are neither available at the subdistrict nor village level. This paper aims to fill this gap by setting an example of value additions possible with research using remotely sensed luminosity data.

3 Data

3.1 Smartcard RCT replication files

MNS published replication files of their RCT. The treatment status of specific subdistricts is clearly identifiable with provided census codes. All other RCT data are labeled with random identifiers to follow data protection guidelines, and are thus not connectable to external data like nighttime lights.

Of 157 studied subdistricts with roughly 3,500 villages covering 20 million people, villagers in 112 randomly selected subdistricts[13] got access to smartcards in October 2010. The remaining 45 control subdistricts got access to smartcards by October 2012.

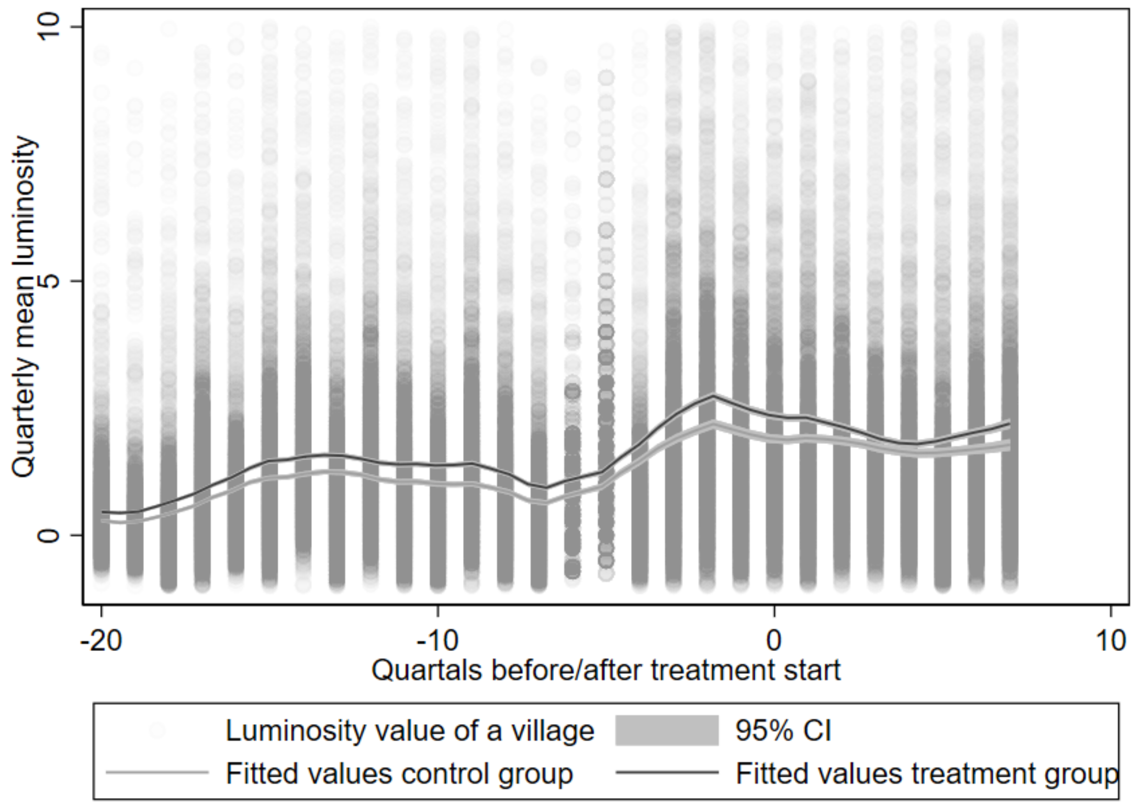
The program setup allowed the continued use of leakage-prone access without smartcards, to prevent IT-caused exclusion issues. Consequently, only 30% of the treated villages used smartcards after 12 months, 60% after 17 months. Thus, interpretation of the treatment

[13]Randomization followed a principal component for stratification to care about any imbalances. Additionally, the proportion of treated and control subdistricts per district was fixed.

effect needs to be adjusted - MNS, as well as this paper, focus on intent-to-treat effects.

3.2 Nightlights

Figure 1: Development of quarterly mean luminosity across the 3,582 studied villages (Q4 2005 - Q3 2012)



Notes: Quarterly mean luminosity is derived by taking the median of nighttime luminosity values observed by DMSP satellites within a month in an area of appr. one km^2 (30 arc-seconds) in which the village center lies, and then the average of three months. For the fitted values, a weighted local polynomial smoothing approach is used. The positive luminosity trend seems reasonable when looking at the economic development, increased wealth, and specifically a rise in the access and use of electricity. CI=Confidence Interval

A collaboration between Development Seed, the World Bank, and the University of Michigan led to *www.india.nightlights.io*, an open-source repository of nightlights with census identifiers, at the village level (Gaba et al. (2016)). Monthly median visibilities of weather-adjusted observations sourced from the Defense Meteorological Satellite Program (DMSP) are available. After they have processed the data further, the researchers matched DMSP raster images with a resolution of roughly one square kilometer to geocoded villages in India. Here, in this re-evaluation, only imagery from the satellites F16 and F18 are used, to prevent major inconsistencies between different satellites.

As a background, Table 3 (appendix) provides average village statistics, e.g., the average population per village amounts to appr. 2000, average literacy is 50%. In this paper, a quarterly mean of the available monthly village-level luminosity value is used as the dependent variable in order to smooth out some noise. This smoothing is usually done by using yearly aggregates (e.g., Prakash et al. (2019)), but because this method would be too rough for re-evaluating a two-year long RCT,[14] we chose the mean quarterly value. This value reflects the relative brightness of villages, interval scaled - negative values still reflect visible lights, but with relatively low luminosity. We use a panel of village-level luminosity values five years before the treatment (Q4 2005) up to the quartal before the treatment of the control group two years later (Q3 2012). Besides, areas classified as "urban" are excluded to account for the rural-urban bias of both, NREGS (which is only targeted at rural citizen) and luminosity measurements. 90% of the observations of luminosity lie between -1 and 6 (For illustration, see Figure 1 and descriptives in Table 3).

3.3 Indian Census

Every ten years, the government of India provides a village-level census. To estimate per head light intensities, the population values of the 2011 census are used. Besides, different subdistrict and village-level characteristics depicted in the 2001 and 2011 census are the basis for building a subdistrict-level principal component that captures differences in the characteristics of subdistricts. Longitude and latitude data of the villages enable the

[14] Nevertheless, we have tested the models with yearly means as well - Table 12 and 13 in the appendix show that they cannot find any significance of treatment and spillovers.

analysis of spillover effects between villages. Therefore, we calculate the fraction of treated villages within a radius of 20km.

4 Empirical analysis

In the following chapter, we first draw a theoretical hypothesis as a basis for explaining the two main estimation strategies to re-evaluate the smartcard RCT with the data described before. Afterwards, the regression results are analyzed, and enhanced with additional interpretations using alternative specifications and a comparison to official district GDP.

4.1 Theoretical hypothesis

Based on the economic literature proving the ability of nightlights to measure economic activity, the connection between the smartcard intervention and observable luminosity is assumed as follows. The smartcards lead to a redistribution of financial resources from officials and wealthy individuals to villagers in treated villages, raising the market wage with spillover effects to nearby villages. The higher income of villagers leads to higher consumption, boosting economic activity in the village and nearby villages. This economic activity can be observed as infrastructural projects like street lamps (Hodler and Raschky (2014)) and new buildings. In addition, more work is being done during the evening, which increases the luminosity seen from space. However, it is questionable whether the redistribution of the beforehand accumulated capital leads to significantly more observable investments. A redistribution within nearby villages is likely, as field officers usually live close to the villages for which they supervise NREGS activities. Landlords might react to higher labour cost in treated districts and thus focus their activities on non-treated regions, but MNS find positive employment effects for treated regions.

The randomized setup of the trial, which MNS carefully control for, implies a possible causal interpretation if we control for any other endogenous factors relevant for nightlights. Therefore, an analysis of electrification status by village using Census 2011 data showed that all studied villages have access to electrification (see Table 3). Politicians seeking elections might have used so-called freebies (e.g., lightbulbs) to convince voters, but state and national elections took place in 2009 and 2014 only, and smartcard roll-out is unlikely to correlate with freebies given away.

Additionally, there is no significant difference in the construction size of NREGS projects between treated and non-treated subdistricts, so that luminosity related to infrastructure is exogeneous. The Andhra Pradesh microfinance crisis starting in 2010/2011 (Mader (2013)) has likely affected the studied regions, but there is no evidence that this correlates in any form with the smartcard treatment.

4.2 Post-treatment estimate

The first of the two regression models deployed in this piece of research is built as close as possible to the models used by Muralidharan et al. (2017), looking at post-treatment differences at the endline in Q3 2012.

$$Y_{vmd} = \alpha + \beta_T T_{md} + \beta_N \tilde{N}^R_{vmd} + \gamma \overline{Y_{md}} + \lambda PC_{md} + \delta_d + \epsilon_{vmd} \quad (1)$$

Y_{vmd} is the mean of monthly median luminosity values for Q3 2012 for village v in mandal (=subdistrict) m in district d. β_T captures the effect of being in a treated mandal, and β_N the effect of having a higher fraction of neighboring villages who are in a treated mandal. $\overline{Y_{md}}$ is the mandal-level mean of the dependent variable at the baseline in Q3 2010. With λPC_{md} and δ_d we control fixed effects, the former is a principal component of key mandal characteristics from the 2011 census[15], the latter simply is a dummy for each of the seven districts in which the studied mandals are located.

[15] Total rural population, literacy rate, job cards per capita, percentage of citizen belonging to a scheduled caste or scheduled tribe, fraction of disabled citizen, fraction of old age citizen and pensions per capita

Table 1: Post-treatment estimates in the main village-level panel (Q3 2012)

VARIABLES	(1) Luminosity	(2) Luminosity	(3) Luminosity
Treatment	0.136	0.104	-0.0902
	(0.276)	(0.443)	(0.193)
Spillovers	-0.772	0.121	
	(0.544)	(0.824)	
Luminosity at baseline	0.808***		0.880***
	(0.0865)		(0.0900)
Constant	0.495	1.885***	0.0416
	(0.398)	(0.591)	(0.308)
Observations	3340	3340	3582
Mandal Principal Component	YES	YES	YES
District Fixed Effects	YES	YES	YES
Adjusted R-squared	0.279	0.126	0.279

Standard errors clustered at mandal level in parentheses.

*** p<0.01; ** p<0.05; * p<0.1

Notes: The third specification bears more observations as 242 villages were not included in accessible location datasets which were the basis for the calculation of spillovers. With statistical insignificance of the treatment coefficient, its negative sign in (3) is not of a concern.

Although the coefficient of the baseline mean of the light intensity per mandal is significantly different from zero, we can also reject that it is equal to one (p<5%, see Table 1 (1)), like MNS. In difference to their simple escape from a Difference-in-Differences (DiD) approach based on this finding, we argue that a DiD estimation is highly important to understand time-dependent differences between treatment and control group (see in the next subsection).

We do not find any significance in the treatment nor the spillover variables as an influencing factor for village luminosity, even after leaving out the baseline luminosity value. Either the overall economic activity proxied by nightlights did not change significantly because of the treatment, or the luminosity observations are too inaccurate to depict economic influences which appears likely when looking at the strong economic effects that MNS found.

4.3 Difference-in-Differences estimate

The following DiD estimate nurtures the richness of the luminosity panel. We look at the differences between control and treatment group, and for both at the differences in the years before and after the treatment. In order to understand the (intent-to-treat)-effects of the smartcard intervention, we include luminosity values from five years before the baseline for every studied village, until the endline in Q3 2012.

$$Y_{vmdt} = \alpha + \beta_T T_{mdt} + \beta_N \tilde{N}^R_{vmdt} + \sum_{j=-8}^{8} \tau T_{md(t+j)} + \delta_{md} + \theta_t + \epsilon_{vmdt} \qquad (2)$$

Y_{vmdt} is the quarterly mean of monthly median luminosity values for village v in mandal m in district d in the respective quarter t. Similar to above, β_T and β_N capture the treatment and spillover effects. In contrast to the cross-section post-treatment model, this model builds upon a panel with quarterly values between 2005 and 2012 across 3,582 villages. Furthermore, leads and lags of the treatment capturing +/- 2 years (+/- 8 quarters) are included, as well as a mandal dummy δ_{md} and a dummy θ_t for every quartal of the 28 covered within the seven years time period. The standard errors ϵ_{vmdt} are clustered at mandal level to prevent heteroscedasticity, like in the post-treatment estimate.

Table 2 shows that the effect of the smartcard intervention (through treatment or spillovers) on nighttime luminosity is insignificant, once time fixed effects are controlled for. This does not change if we control for subdistrict or village fixed effects when including leads and lags of the treatment variable to account for announcement effects and a delayed roll-out. Interestingly, the spillover variable takes all significance and effect size from the treatment

Table 2: DiD estimates in the main quarterly village-level panel

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Luminosity	Luminosity	Luminosity	Luminosity	Luminosity
Treatment	0.722***	-0.173	-0.162	-0.218	-0.219
	(0.147)	(0.455)	(0.465)	(0.241)	(0.239)
Spillovers		1.039**	0.931	0.0475	0.128
		(0.464)	(0.777)	(0.411)	(0.392)
Constant	1.351***	1.307***	0.545***	-0.969***	-1.007***
	(0.111)	(0.112)	(0.0607)	(0.106)	(0.109)
Observations	99640	97704	97704	97704	97704
Quarterly Fixed Effects	NO	NO	YES	YES	YES
Mandal Fixed Effects	NO	NO	NO	YES	NO
Village Fixed Effects	NO	NO	NO	NO	YES
Leads and Lags	NO	NO	NO	YES	YES
Adjusted R-squared	0.0111	0.0126	0.0725	0.310	0.641

Standard errors clustered at mandal level in parentheses.

*** p<0.01; ** p<0.05; * p<0.1

Notes: The treatment is insignificant once spillovers are included, and even those become insignificant once time fixed effects are included. Leads and Lags, mandal FE and village FE do not change this result. Some observations are lost once spillovers are included, as geocoding is not available for some villages.

variable, and even outgrows its effect size, but does not stay significantly different from zero once fixed effects are included. Two major implications can be drawn, like above: Either the smartcard program did not have a strong impact on economic activity, or the nightlight data cannot accurately depict the economic activity increase caused by the intervention.

4.4 Alternative specifications

This subsection describes how we test different alternative specifications (results in Table 6 to Table 16 in the appendix) with particular rationales in mind in order to validate our findings.

First, the dependent variable only reflects total village observations, it does not depict per head observations. Those are calculated based on the 2011 census figures, thus insensitive to larger migration in the considered time frame 2005-2012. Nevertheless, the results are similar to our previous estimations; there is no significant observable impact of the treatment and spillovers (Table 6 and 7). This insignificance is not surprising as population distribution is not tremendously different in control (mean village size = 2078, sd = 2900) and treatment group (mean village size = 1992, sd = 2514).

Second, the spillover effects are checked for sensitivity, expanding the radius for capturing spillovers from 20km to 40km. Again, the results are not affected (Table 8 and 9). If we use the initial monthly luminosity values (Table 10 and 11), we see a higher standard deviation of luminosity, and do not see a change in the results. P-values are higher for the monthly compared to the quarterly values for example for luminosity at the baseline, and spillovers in (2) of the DiD model, which is possibly indicating the imprecision of monthly values. A yearly instead of a quarterly mean is unable to capture any possible effects, with insignificant treatment coefficients for post-treatment and DiD estimates. (Table 12 and 13).

As the smartcard program mainly leads to a redistribution to villagers from officials who most likely operate within nearby villages but possibly also just in the same mandal, we create a new mandal-level panel aggregated from the village-level luminosity values. Neither the per head nor the median estimate are significant, once again proving the initial estimates. (Table 14 and 15).

Lastly, we specify a lagged regression equation. MNS replication files indicate that five quarters post-treatment, the village conversion rate surpasses 50%. Thus, the treatment and the spillover variable are lagged by five quarters. DiD estimates show no significance of the lagged treatment or spillovers for luminosity values (Table 16).

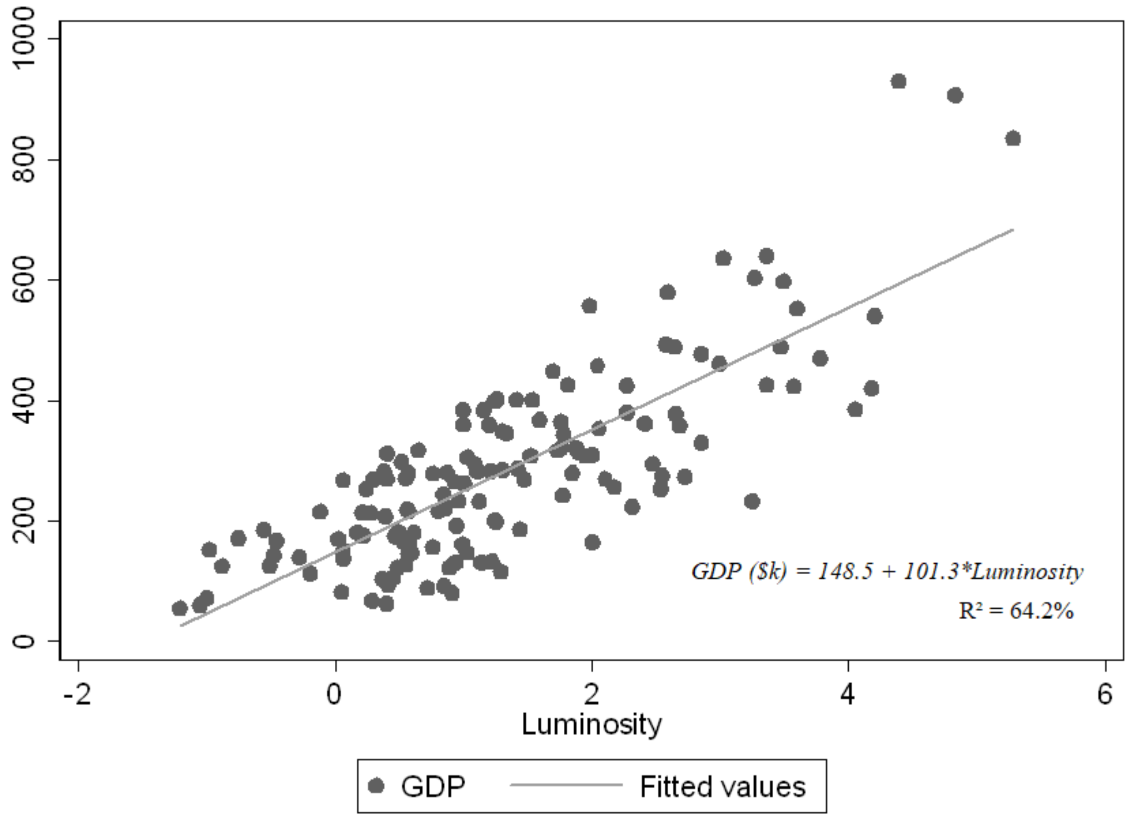
4.5 GDP and nightlights

Although the literature has proven the suitability of nightlights as a proxy for GDP, the following subsection shortly looks into this correlation for the studied districts. It is important to keep in mind that GDP estimates are only available at the district level and can be of lower accuracy for depicting economic activity than luminosity values. This short comparison is relevant in order to understand treatment sizes and be able to interpret the main regressions meaningfully.

64.7% of the variation in official GDP values in all districts of Andhra Pradesh in the period 2005-2012 can be explained with variation in luminosity levels. For every district, we divided the official GDP by the number of villages and transformed it into a monthly value to be able to interpret the estimates of the post-treatment and DiD model directly.[16] There is

[16] Of course, this approach bears inaccuracies, but difficult to improve given the availability of official GDP data only at the district level. An alternative approach would be to use the sum of nightlights per district and compare it to the district GDP values, but we purposely stay on the village-level for more direct interpretations. In the appendix, in Table 4, an alternative per head estimate of GDP and luminosity is provided.

Figure 2: District-wise GDP correlations with luminosity



Notes: This is based on a panel of all studied districts. We divide the official yearly GDP value for each district by the number of villages per district, and by twelve, to finally be able to compare it in absolute terms to the 3-months mean luminosity. Includes one value for every year 2005-2012 for every district in Andhra Pradesh. As k USD are used, the y-axis reflects values between 0 and 1 million USD.

some degree of variation in the estimated coefficient of luminosity on GDP depending on whether we only look at studied or all districts in Andhra Pradesh. For every extra point in luminosity, the absolute GDP per month in a village is higher by roughly 100,000 USD. [17] Using this as a reference while looking at village-wise luminosity developments, the enormous variations (standard deviation of 2.773, around the mean of 1.501) are a big concern. For example, when looking at the three months means of the luminosity of a specific village in the data, it is not uncommon to observe luminosity values at 1 in one quarter and 2.5

[17] Specifically, for every extra point in luminosity, 80,110 USD extra in studied districts, 101,300 when looking at all AP districts. Similarly, for every tenth of a point in luminosity per head, GDP per head is higher by 6,752 USD in studied districts, and 7,942 USD in all AP districts (see Table 4 and 5).

in the next quarter, then going down to 0.5 straight afterward. This would mean that the village GDP of appr. 100,000 USD (roughly 20-50$ on average per villager) more than doubled to 250,000 USD within one quarter, just to then drop to a fifth of that - to 50,000 USD. This reveals how noisy nightlight data is, and that yearly aggregates should be used if possible. However, for a two year long RCT with slow conversion and treatment of the control group at the end, this does not make much sense from a data point of view.

If applied to the estimated treatment size of the post-treatment model (Table 1 (1)), treatment was correlated with a higher village GDP of roughly 13,600 USD per month (sd = 27,600 USD) - though not significant. The - again insignificant - DiD estimate (Table 2 (4)) suggests a lower village GDP of 21,800 USD per month (sd = 24,100 USD), with a size of the spillovers of 4,750 USD per month (sd = 41,100 USD). Thus based on one standard deviation, a change of +/- 50,000 USD is estimated. Given the mean luminosity of 1.5 (Table 3), which equals 150,000 USD per month, this would mean a change by one third. Building upon these thoughts, the imprecision of nightlights is a barrier towards its usage for re-evaluations. Accurate satellite data with low variations across time, resulting in a village-wise consistent development, could be fruitfully used for estimating potential economic gains through evaluations of different policies. However, this is not the case for the Indian nightlight data used here, and to a vast extent, for micro-level high frequency nightlight data in general. More accurate forms of satellite data are needed to depict small changes induced by subdistrict- or village-level randomized controlled trials.

5 Limitations and potential

The exemplary analysis of the smartcard intervention shows that the combination of RCTs and nightlights is prone to several inaccuracies limiting the potential applicability.

Nighttime luminosity is found to be generally noisy, and prone to sensitivity differences in satellite sensors. Thus, aggregates like annual composites (as released by NOAA) are needed. However, RCTs do not usually span across multiple years, so that at least quarterly observations are necessary to understand impacts precisely. As seen in the analysis beforehand, these quarterly means are still prone to noisiness (although less noisy than monthly values), and on cloudy days, such as monsoon time in India, quarterly aggregates are based on just a few observations or in some cases, not even available (e.g., Q2 2009).

For more accuracy, the NASA-NOAA has launched the visible infrared imaging radiometer suite (VIIRS) in late 2011, which is found to be slightly less noisy (Addison and Stewart (2015). The absence of light at night does not necessarily indicate lower economic activity in the short term. Weather, cultural or political shifts of economic activity toward daytime or more unlit activities will continue to affect nightlights as an economic proxy, no matter how sophisticated the sensor systems are.

Beside aggregates across time, it is questionable in what manner nightlights or satellite data in general can depict economic activity at the granular levels on which economists typically randomize. Usually, randomization is done on the household or village-level. The use of satellite data seems unreasonable for many cases of the former. With the example at hand randomizing at an exceptionally high level - at subdistricts - it is possible to gain an understanding of the area granularity issue, by comparing the results of both, subdistrict-level and village-level estimates. We do not find any effect at either level so that we at least cannot falsify the hypothesis that village-level luminosity values are reasonable to use. Research by Dugoua et al. (2018) and others point in a similar direction.

MNS conducted on-the-ground interviews to explicitly look at the most impacted month of June. If nightlights are only valuable in quarterly or yearly aggregates, such specific outcome measurements are challenging to obtain. Nonetheless, the monthly estimates presented here show similar results, although noisier. One of the biggest potentials of satellite data lies in understanding the impacts of randomized interventions beyond the endline, especially when on-the-ground measurements are not possible or simply too costly. For this, even nightlights as yearly aggregates can be an option. In cases where the control group can access treatment right after the endline, like in the smartcard example, the potential of long-term evaluations is somewhat limited. But for hundreds of already conducted RCTs, with no treatment of the control group straight after the endline, satellite data provide an excellent, easy-to-harness potential. With remotely sensed data, researcher can deeply understand effects beyond a typical RCT evaluation setup looking at a couple of years between baseline and endline. Generally, the high inter-comparability of satellite data will help to gain much more objective data, e.g., the effects of a similar RCT conducted on different continents can be measured with the same satellite.

Furthermore, it is crucial to understand which variables of interest can be proxied with remotely sensed luminosity values. The most prominent measure is economic activity, the most straight-forward measure electricity use. The re-evaluation of the large-scale smartcard RCT has shown that roughly 50% of the variation of GDP for a seven year time span can be explained with luminosity values in the studied districts, even though not all subdistricts per district have been treated or were in the control group. The most often used measures in Development Economics randomized controlled trials are in the areas of economic benefit (income, goods, and other), health and education. Only for a small part of the former - mainly economic activity and electricity access - nightlights are widely found to be useful. [18] Although still slightly inappropriate for the latter two, daytime satellite data can depict far more economic indicators, e.g., roof conditions (Varshney et al. (2015). Beyond, remotely sensed data are compelling for ecological indicators like tree cover, vegetation and air pollution (Fowlie et al. (2019)), which could be an exciting addition to many

[18] Ghosh et al. (2013) are one of a few authors looking at nightlight predictions of human well-being and poverty, but it is once again GDP that is found to be most accurately predicted by nightlights.

randomized controlled impact measurements. Daytime imagery exceeds the capabilities of nightlights when looking at the spatial resolution, possible measurements of economic indicators and noisiness - although it lacks the simplicity of nightlights and usually requires advanced geomapping tools.Jain (2020) provide a useful overview of typical measurement errors in satellite data that are important for deriving causal inference.
For satellite imagery in general, as well as for nightlights, the critical question is: Can remotely sensed data capture small, but significant effects caused by randomized interventions? The spatial resolution, noisiness and usability are key features. As shown in the empirical analysis, nightlight data likely are rather unable to capture those small effects because of their noisiness. Daylight imagery might find significant positive effects. The combination of more sophisticated satellite data and RCTs is a key direction for future research.

However, the other side of this powerful combination needs to be kept in mind as well: The randomized controlled trial itself. Besides general issues like limited external validity, RCTs still need to be conducted on the ground, involving high setup costs. Although satellite data can help to reduce the burden of data collection, valuable microdata, e.g., related to personal feelings, health issues, and education can only be gathered through on-the-ground surveys. Satellite data can enable a cost-efficient analysis of large-scale RCTs targeted at economic activity, with low-cost measurements during the trial and beyond.

6 Conclusion

Satellite data and randomized controlled trials (RCTs) are a powerful combination for analyzing causal effects beyond traditional survey-based indicators. The usage of remotely collected data for evaluating RCTs is cost-effective, objective, open for anyone with treatment assignment data and possible beyond the endline. By re-evaluating one of the largest RCTs ever conducted in the context of development economics - the smartcard intervention of Muralidharan et al. (2016) covering nearly 20 million people - with Indian nighttime luminosity, this paper finds that nightlights as a specific type of satellite data are too noisy to evaluate short-term RCTs. Other satellite data, especially daytime imagery will probably outscore luminosity-based approaches.

In the analysis, there was no statistically significant effect found of the randomized smartcard intervention on nighttime luminosity, neither with a post-treatment nor a Difference-in-Differences estimate. Differently aggregated specifications confirm the initial absence of statistically significant effects of the smartcard treatment on economic activity proxied by nighttime luminosity. A look at the correlation with GDP to meaningfully interpret the noisiness of nightlight data leads to the conclusion that this kind of satellite data has a limited suitability for the analysis of RCTs.

Nonetheless, the potential for re-evaluating RCTs with other satellite data than nightlights is enormous in three ways: First, for confirming claimed treatment effects. Second, to understand further impacts, e.g., on the ecological or comprehensive economic side at the endline. Third, for understanding long-term impacts of randomized interventions. In the research, we have struggled to receive access just to the treatment assignment information for RCTs. There is a clear trade-off between data protection and usability for re-evaluation that is a barrier for harnessing the massive potential that the combination of RCT's treatment assignment data and satellite data brings. Moreover, large-scale RCTs can be evaluated much more cost-effectively with the support of satellite data. On-the-ground analysis will likely always remain relevant, while satellite data becomes a more and more critical addition with increased accuracy and granularity that new technical solutions like daytime imagery processing bring. Nightlights might remain valuable for other purposes, but likely not for re-evaluating RCTs, maybe except for some long-term analysis using yearly aggregates.

7 Appendix

Table 3: Descriptive statistics of key variables in the main quarterly village-level panel

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
Luminosity	99,640	1.501	2.773	-3.417	56.79
L per head	92,925	0.00636	0.160	-0.211	14.30
Treatment	99,640	0.208	0.406	0	1
Spillovers	97,704	0.196	0.347	0	1
Access to electricity	99,640	1	0	1	1
Total population	98,863	2,016	2,628	0	36,031
Literacy rate	92,925	0.501	0.101	0	1

Notes: This table reflects the full panel, with quarterly values between Q4 2005 and Q3 2012. The number of observations for per head luminosity and spillovers is lower because of missing population data or values of zero for some villages and missing geocoding for others. Because processed data is used, luminosity values do not represent any physical unit, and can only be interpreted relatively on an interval scale. Spillovers have been created by simply looking at the fraction of treated villages in a 20km radius using longitude and latitude data - in contrast to MNS, we do not exclude villages in the same mandal from this fraction. Although non-treated villages have spillover values greater than zero if they are close to treated villages, the mean of the treatment is still greater than the mean of the spillovers because treated villages in proximity to non-treated villages bear values below one. The mean of the treatment variable, 0.208, reflects a value of zero in the five years before the treatment starts, and a value of one for roughly 70% (112 out of 157 mandals were treated with similar number of villages) of the villages during the two year treatment period. For more specific comparisons of village and mandal characteristics between treatment and control group, MNS provide insightful tables, suggesting high similarity.

Table 4: District-wise GDP correlations with Luminosity (all AP districts)

VARIABLES	(1) GDP per head	(2) GDP per head	(3) GDP	(4) GDP
Luminosity per head	79.42***	18.68**		
	(6.820)	(7.160)		
Luminosity			101.3***	89.23***
			(6.321)	(8.512)
Constant	0.0701***	0.0498***	148.5***	128.6***
	(0.00297)	(0.00459)	(11.58)	(19.53)
Observations	144	144	144	144
District Fixed Effects	NO	YES	NO	YES
Year Fixed Effects	NO	YES	NO	YES
Adjusted R-squared	0.485	0.922	0.642	0.947

Standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: This is based on a panel of all studied districts. We divide the official yearly GDP value for each district by the number of villages per district, and by twelve, to finally be able to compare it in absolute terms to the 3-months mean luminosity. Includes one value for every year 2005-2012 for every district in Andhra Pradesh. Converted INR to USD with year-specific exchange rates from the Worldbank. Per head estimates are slightly skewed for values below zero, as luminosity values have no defined zero point. The dependent variable GDP is in k USD.

Table 5: District-wise GDP correlations with Luminosity (studied districts)

	(1)	(2)	(3)	(4)
VARIABLES	GDP per head	GDP per head	GDP	GDP
Luminosity per head	67.52***	6.319		
	(10.27)	(6.160)		
Luminosity			80.11***	44.91***
			(12.26)	(13.96)
Constant	0.0715***	0.0504***	156.4***	84.29***
	(0.00355)	(0.00298)	(13.46)	(19.31)
Observations	48	48	48	48
District Fixed Effects	NO	YES	NO	YES
Year Fixed Effects	NO	YES	NO	YES
Adjusted R-squared	0.473	0.975	0.470	0.950

Standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: Same procedure as above, but only includes studied districts. Keep in mind that for every studied district, there were treatment, control and non-studied buffer subdistricts. The dependent variable GDP is in k USD.

Table 6: Per head post-treatment estimates in the main quarterly village-level panel (Q3 2012)

	(1)	(2)	(3)
VARIABLES	L per head	L per head	L per head
Treatment	0.000659	-0.00536	-0.00139
	(0.000830)	(0.00433)	(0.00309)
Spillovers	0.000864	0.0121*	
	(0.00114)	(0.00704)	
L per head at baseline	0.361***		
	(0.0600)		
Constant	0.000518	0.00166	0.00841*
	(0.000883)	(0.00239)	(0.00484)
Observations	3318	3318	3340
Mandal Principal Component	YES	YES	YES
District Fixed Effects	YES	YES	YES
Adjusted R-squared	0.916	-3.10e-05	-0.000690

Standard errors clustered at mandal level in parentheses. L=Luminosity

*** $p<0.01$; ** $p<0.05$; * $p<0.1$

Note: Per head estimates have been created by dividing quarterly luminosity by the population of a village. Because the population is only available for the 2011 census, this per head estimate is slightly inaccurate, and not sensitive towards migration. Additionally, per head values for negative luminosity values are biased as well, because of the interval scale of luminosity, and thus only be an approximation. We do not exclude values below zero, as they are still relevant, and this bias because of the division by population does not seem like a strong concern as.

Table 7: Per head DiD estimates in the main quarterly village-level panel

VARIABLES	(1) L per head	(2) L per head	(3) L per head	(4) L per head	(5) L per head
Treatment	-0.000705	-0.00559	-0.00665	0.00801	0.00670
	(0.00169)	(0.00525)	(0.00534)	(0.00562)	(0.00580)
Spillovers		0.00543	0.0169	-0.00393	-0.00353
		(0.00507)	(0.0102)	(0.00526)	(0.00271)
Constant	0.00651**	0.00605**	0.00233**	-0.00409*	-0.00487**
	(0.00296)	(0.00283)	(0.00114)	(0.00230)	(0.00212)
Observations	92925	92749	92749	92749	92749
Quarterly Fixed Effects	NO	NO	YES	YES	YES
Mandal Fixed Effects	NO	NO	NO	YES	NO
Village Fixed Effects	NO	NO	NO	NO	YES
Leads and Lags	NO	NO	NO	YES	YES
Adjusted R-squared	-7.59e-06	1.96e-05	0.000196	0.0355	0.711

Standard errors clustered at mandal level in parentheses. L=Luminosity

*** $p<0.01$; ** $p<0.05$; * $p<0.1$

See above. Treatment and spillovers are insignificant also in this specification, even before time fixed effects are introduced. This might be due to the inaccuracies involved with calculating the per head estimate.

Table 8: Post-treatment estimates in the main quarterly village-level panel (Q3 2012), with 40km spillovers

	(1)	(2)
VARIABLES	Luminosity	Luminosity
Treatment	-0.117	0.0108
	(0.209)	(0.307)
Spillovers (40km)	-0.304	1.246
	(0.825)	(1.166)
Luminosity at baseline	0.799***	
	(0.0878)	
Constant	0.370	1.160
	(0.606)	(0.868)
Observations	3340	3340
Mandal Principal Component	YES	YES
District Fixed Effects	YES	YES
Adjusted R-squared	0.277	0.129

Standard errors clustered at mandal level in parentheses.

*** p<0.01; ** p<0.05; * p<0.1

Spillovers have been created by simply looking at the fraction of treated villages in a 40km radius using longitude and latitude data - in contrast to MNS, we do not exclude villages in the same mandal from this fraction.

Table 9: DiD estimates in the main quarterly village-level panel, with 40km spillovers

	(1)	(2)	(3)
VARIABLES	Luminosity	Luminosity	Luminosity
Treatment	0.722***	0.0364	-0.192
	(0.147)	(0.351)	(0.209)
Spillovers (40km)		0.869**	-0.0689
		(0.356)	(0.515)
Constant	1.351***	1.298***	-0.987***
	(0.111)	(0.114)	(0.141)
Observations	99640	97704	97704
Quarterly Fixed Effects	NO	NO	YES
Mandal Fixed Effects	NO	NO	YES
Village Fixed Effects	NO	NO	NO
Leads and Lags	NO	NO	YES
Adjusted R-squared	0.0111	0.0126	0.310

Standard errors clustered at mandal level in parentheses.

*** p<0.01; ** p<0.05; * p<0.1

Spillovers have been created by simply looking at the fraction of treated villages in a 40km radius using longitude and latitude data - in contrast to MNS, we do not exclude villages in the same mandal from this fraction. Of course, spillovers are zero for any village before treatment starts.

Table 10: Post-treatment estimates monthly, at the village level (07-09/2012)

	(1)	(2)
VARIABLES	Luminosity	Luminosity
Treatment	0.144	0.127
	(0.388)	(0.446)
Spillovers	-0.336	-0.0293
	(0.758)	(0.834)
Luminosity at the baseline	0.290***	
	(0.0554)	
Constant	2.188***	1.960***
	(0.557)	(0.611)
Observations	8427	9723
Mandal Principal Component	YES	YES
District Fixed Effects	YES	YES
Adjusted R-squared	0.136	0.0920

Standard errors clustered at mandal level in parentheses.

*** $p<0.01$; ** $p<0.05$; * $p<0.1$

Notes: For easier relative interpretations, the mean of Luminosity is 1.507 (sd=3.079).

Table 11: DiD estimates monthly, at the village level

VARIABLES	(1) Luminosity	(2) Luminosity	(3) Luminosity	(4) Luminosity	(5) Luminosity
Treatment	0.732***	-0.172	-0.159	-0.207	-0.196
	(0.148)	(0.456)	(0.466)	(0.236)	(0.232)
Spillovers		1.057**	0.915	0.0604	0.139
		(0.467)	(0.778)	(0.420)	(0.397)
Constant	1.343***	1.293***	0.993***	-0.506***	-0.530***
	(0.111)	(0.113)	(0.0747)	(0.107)	(0.109)
Observations	275604	269829	269829	269829	269829
Monthly Fixed Effects	NO	NO	YES	YES	YES
Mandal Fixed Effects	NO	NO	NO	YES	NO
Village Fixed Effects	NO	NO	NO	NO	YES
Leads and Lags	NO	NO	NO	NO	NO
Adjusted R-squared	0.00983	0.0112	0.0984	0.294	0.569

Standard errors clustered at mandal level in parentheses.

*** $p<0.01$; ** $p<0.05$; * $p<0.1$

Notes: For easier relative interpretations, the mean of Luminosity is 1.507 (sd=3.079). In the monthly panel, only every third lead and lag has been included.

Table 12: Post-treatment estimates yearly, at the village level (2012)

	(1)	(2)
VARIABLES	Luminosity	Luminosity
Treatment	-0.0801	-0.0541
	(0.235)	(0.440)
Spillovers	-0.391	0.362
	(0.453)	(0.845)
Luminosity at the baseline	0.812***	
	(0.0665)	
Constant	0.0733	2.598***
	(0.308)	(0.620)
Observations	3340	3340
Mandal Principal Component	YES	YES
District Fixed Effects	YES	YES
Adjusted R-squared	0.338	0.123

Standard errors clustered at mandal level in parentheses.

*** p<0.01; ** p<0.05; * p<0.1

Notes: For easier relative interpretations, the mean of Luminosity is 1.54 (sd=2.59), the mean of Luminosity per head is 0.00657 (sd=0.152). The post-treatment effects are statistically insignificant - which is not surprising, as we can only observe three values for the years 2010 (baseline), 2011 (excluded here) and 2012 (endline) so that yearly means are too rough for RCTs spanning across two years or less.

Table 13: DiD estimates yearly, at the village level

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Luminosity	Luminosity	Luminosity	Luminosity	Luminosity
Treatment	0.696***	-0.201	-0.194	-0.382	-0.380
	(0.153)	(0.475)	(0.485)	(0.239)	(0.242)
Spillovers		1.038**	0.968	0.0604	0.143
		(0.484)	(0.804)	(0.438)	(0.444)
Constant	1.398***	1.355***	0.657***	-0.882***	-0.942***
	(0.116)	(0.118)	(0.0767)	(0.0943)	(0.103)
Observations	25074	24590	24590	24590	24590
Yearly Fixed Effects	NO	NO	YES	YES	YES
Mandal Fixed Effects	NO	NO	NO	YES	NO
Village Fixed Effects	NO	NO	NO	NO	YES
Leads and Lags	NO	NO	NO	YES	YES
Adjusted R-squared	0.0118	0.0134	0.0637	0.355	0.742

Standard errors clustered at mandal level in parentheses.

*** p<0.01; ** p<0.05; * p<0.1

Notes: For easier relative interpretations, the mean of Luminosity is 1.54 (sd=2.59), the mean of Luminosity per head is 0.00657 (sd=0.152).

Table 14: Post-treatment estimates quarterly, at the subdistrict level (Q3 2012)

	(1)	(2)	(3)	(4)
VARIABLES	L per head	L per head	L-median of villages	L-median of villages
Treatment	-1.91e-05	0.000182	-0.0736	0.0761
	(0.000112)	(0.000151)	(0.176)	(0.208)
L per head at baseline	0.828***			
	(0.0781)			
L-median at baseline			0.647***	
			(0.0843)	
Constant	2.95e-05	0.00132***	0.649*	1.762***
	(0.000232)	(0.000266)	(0.329)	(0.349)
Observations	157	157	157	157
Mandal Principal Component	YES	YES	YES	YES
District Fixed Effects	YES	YES	YES	YES
Adjusted R-squared	0.644	0.285	0.564	0.317

Standard errors clustered at mandal level in parentheses. L=Luminosity

*** $p<0.01$; ** $p<0.05$; * $p<0.1$

Notes: For easier relative interpretations, the mean of Luminosity is 1.219 (sd=1.331), the mean of Luminosity per head is 0.000765 (sd=0.000968). As we look at the rough subdistrict-level panel here, we do not include spillovers between bordering villages.

Table 15: DiD estimates quarterly, at the subdistrict level

	(1)	(2)	(3)	(4)
VARIABLES	L per head	L per head	L-median of villages	L-median of villages
Treatment	0.000362***	1.14e-06	0.575***	-0.105
	(6.73e-05)	(0.000116)	(0.0925)	(0.217)
Constant	0.000691***	-0.000501***	1.101***	-0.907***
	(5.71e-05)	(4.32e-05)	(0.0731)	(0.0624)
Observations	4375	4375	4375	4375
Quarterly Fixed Effects	NO	YES	NO	YES
Mandal Fixed Effects	NO	YES	NO	YES
Leads and Lags	NO	YES	NO	YES
Adjusted R-squared	0.0226	0.788	0.0301	0.763

Standard errors clustered at mandal level in parentheses. L=Luminosity

*** p<0.01; ** p<0.05; * p<0.1

Notes: For easier relative interpretations, the mean of Luminosity is 1.219 (sd=1.331), the mean of Luminosity per head is 0.000765 (sd=0.000968). As we look at the rough subdistrict-level panel here, we do not include spillovers between bordering villages.

Table 16: Lagged DiD estimates in the main quarterly village-level panel

VARIABLES	(1) Luminosity	(2) Luminosity	(3) Luminosity	(4) Luminosity	(5) Luminosity
Treatment (5 quartals lagged)	0.564***	-0.118	-0.135	-0.0979	-0.0843
	(0.157)	(0.515)	(0.525)	(0.287)	(0.287)
Spillovers (5 quartals lagged)		0.724	0.903	0.00661	0.0728
		(0.518)	(0.864)	(0.465)	(0.450)
Constant	1.457***	1.447***	0.545***	-1.000***	-1.039***
	(0.119)	(0.120)	(0.0607)	(0.0820)	(0.0843)
Observations	99640	98914	98914	98914	98914
Quarterly Fixed Effects	NO	NO	YES	YES	YES
Mandal Fixed Effects	NO	NO	NO	YES	NO
Village Fixed Effects	NO	NO	NO	NO	YES
Leads and Lags	NO	NO	NO	NO	NO
Adjusted R-squared	0.00296	0.00263	0.0695	0.309	0.635

Standard errors clustered at mandal level in parentheses.

*** p<0.01; ** p<0.05; * p<0.1

In this analysis, both treatment and spillovers have been lagged 5 quarters. According to conversion data from MNS, in December 2011 and then constantly in all months of Q1 2012, and 2012 in general, the fraction of converted villages picks up and reaches values beyond 60% . Q1 2012 is five quarters after the start of the treatment in Q4 2010, so that values are lagged 5 quarters.

References

Addison, D. and Stewart, B. *Nighttime Lights Revisited: The Use of Nighttime Lights Data as a Proxy for Economic Variables*. Policy Research Working Papers. The World Bank, November 2015. doi: 10.1596/1813-9450-7496.

Asher, S. and Novosad, P. Rural Roads and Local Economic Development. *American Economic Review*, 110(3):797–823, March 2020. ISSN 0002-8282. doi: 10.1257/aer.20180268.

Beyer, R. C. M., Chhabra, E., Galdo, V., and Rama, M. *Measuring Districts' Monthly Economic Activity from Outer Space*. Policy Research Working Papers. The World Bank, July 2018. doi: 10.1596/1813-9450-8523.

Bhandari, L. and Roychowdhury, K. Night Lights and Economic Activity in India: A study using DMSP-OLS night time images. *Proceedings of the Asia-Pacific Advanced Network*, 32(0):218, December 2011. ISSN 2227-3026. doi: 10.7125/APAN.32.24.

Bickenbach, F., Bode, E., Nunnenkamp, P., and Söder, M. Night lights and regional GDP. *Review of World Economics (Weltwirtschaftliches Archiv)*, 152(2):425–447, 2016.

Bundervoet, T., Maiyo, L., and Sanghi, A. *Bright Lights, Big Cities: Measuring National and Subnational Economic Growth in Africa from Outer Space, with an Application to Kenya and Rwanda*. Policy Research Working Papers. The World Bank, October 2015. doi: 10.1596/1813-9450-7461.

Corral, L. R., Schling, M., and Montiel, C. The economic and ecological impact of natural resource extraction: The case of the Camisea Gas Project in Peru. Working Paper IDB-WP-934, IDB Working Paper Series, 2018.

Deaton, A. and Cartwright, N. Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine*, 210:2–21, August 2018. ISSN 0277-9536. doi: 10.1016/j.socscimed.2017.12.005.

Donaldson, D. and Storeygard, A. The View from Above: Applications of Satellite Data in Economics. *Journal of Economic Perspectives*, 30(4):171–198, November 2016. ISSN 0895-3309. doi: 10.1257/jep.30.4.171.

Dugoua, E., Kennedy, R., and Urpelainen, J. Satellite data for the social sciences: Measuring rural electrification with night-time lights. *International Journal of Remote Sensing*, 39(9):2690–2701, May 2018. ISSN 0143-1161. doi: 10.1080/01431161.2017.1420936. _eprint: https://doi.org/10.1080/01431161.2017.1420936.

Fowlie, M., Rubin, E., and Walker, R. Bringing Satellite-Based Air Quality Estimates Down to Earth. *AEA Papers and Proceedings*, 109:283–288, May 2019. ISSN 2574-0768, 2574-0776. doi: 10.1257/pandp.20191064.

Gaba, K. M., Min, B., Thakker, A., and Elvidge, C. India Lights. http://india.nightlights.io/, 2016.

Ghosh, T., Powell, R. L., Anderson, S., Sutton, P. C., and Elvidge, C. D. Informal Economy and Remittance Estimates of India Using Nighttime Imagery. *International Journal of Ecological Economics and Statistics™*, 17(P10):16–50, March 2010. ISSN 0973-7537.

Ghosh, T., Anderson, S. J., Elvidge, C. D., and Sutton, P. C. Using Nighttime Satellite Imagery as a Proxy Measure of Human Well-Being. *Sustainability*, 5(12):4988–5019, December 2013. doi: 10.3390/su5124988.

Henderson, J. V., Storeygard, A., and Weil, D. N. Measuring Economic Growth from Outer Space. *American Economic Review*, 102(2):994–1028, April 2012. ISSN 0002-8282. doi: 10.1257/aer.102.2.994.

Hodler, R. and Raschky, P. Regional Favoritism. *The Quarterly Journal of Economics*, 129 (2):995–1033, 2014.

Hu, Y. and Yao, J. Illuminating Economic Growth. page 64, 2019.

Imbert, C. and Papp, J. Labor Market Effects of Social Programs: Evidence from India's Employment Guarantee. *American Economic Journal: Applied Economics*, 7(2):233–263, April 2015. ISSN 1945-7782. doi: 10.1257/app.20130401.

Jain, M. The Benefits and Pitfalls of Using Satellite Data for Causal Inference. *Review of Environmental Economics and Policy*, 14(1):157–169, January 2020. ISSN 1750-6816. doi: 10.1093/reep/rez023.

Jamison, J. C. The Entry of Randomized Assignment into the Social Sciences. SSRN Scholarly Paper ID 2969643, Social Science Research Network, Rochester, NY, May 2017.

Mader, P. Rise and Fall of Microfinance in India: The Andhra Pradesh Crisis in Perspective. *Strategic Change*, 22(1-2):47–66, 2013. ISSN 1057-9265, 1086-1718, 1099-1697. doi: 10.1002/jsc.1921.

Michalopoulos, S. and Papaioannou, E. National Institutions and Subnational Development in Africa. Working Paper 18275, National Bureau of Economic Research, August 2012.

Mitnik, O. A., Sanchez, R., and Yañez, P. Bright Investments: Measuring the Impact of Transport Infrastructure Using Luminosity Data in Haiti. Technical report, Inter-American Development Bank, December 2018.

Muralidharan, K., Niehaus, P., and Sukhtankar, S. Building State Capacity: Evidence from Biometric Smartcards in India. *American Economic Review*, 106(10):2895–2929, October 2016. ISSN 0002-8282. doi: 10.1257/aer.20141346.

Muralidharan, K., Niehaus, P., and Sukhtankar, S. General Equilibrium Effects of (Improving) Public Employment Programs: Experimental Evidence from India. Technical Report w23838, National Bureau of Economic Research, Cambridge, MA, September 2017.

Muralidharan, K., Niehaus, P., and Sukhtankar, S. Identity Verification Standards in Welfare Programs: Experimental Evidence from India. SSRN Scholarly Paper ID 3535334, Social Science Research Network, Rochester, NY, February 2020.

Otchia, C. and Asongu, S. Industrial Growth in Sub-Saharan Africa: Evidence from Machine Learning with Insights from Nightlight Satellite Images. *SSRN Electronic Journal*, 2019. ISSN 1556-5068. doi: 10.2139/ssrn.3436950.

Prakash, A., Shukla, A. K., Bhowmick, C., and Beyer, R. C. M. Night-time Luminosity: Does it Brighten Understanding of Economic Activity in India? *Reserve Bank of India Occasional Papers*, 40(1), 2019.

Varshney, K. R., Chen, G. H., Abelson, B., Nowocin, K., Sakhrani, V., Xu, L., and Spatocco, B. L. Targeting Villages for Rural Development Using Satellite Image Analysis. *Big Data*, 3(1):41–53, March 2015. ISSN 2167-6461, 2167-647X. doi: 10.1089/big.2014.0061.

Weidmann, N. B. and Schutte, S. Using night light emissions for the prediction of local wealth. *Journal of Peace Research*, 54(2):125–140, March 2017. ISSN 0022-3433. doi: 10.1177/0022343316630359.

Wu, J., Wang, Z., Li, W., and Peng, J. Exploring factors affecting the relationship between light consumption and GDP based on DMSP/OLS nighttime satellite imagery. *Remote Sensing of Environment*, 134:111–119, July 2013. ISSN 0034-4257. doi: 10.1016/j.rse.2013.03.001.