Foreign Aid and Growth: A Sp P-VAR Analysis Using Satellite Sub-National Data for Uganda

Andrea Civelli* — *University of Arkansas*

Andrew Horowitz† — *University of Arkansas*

Arilton Teixeira‡ — *Ibmec*

Abstract

We develop a measurement strategy for the impact of foreign aid based on a regional spatial panel vector-autoregressive model (Sp P-VAR). We illustrate the strategy using Ugandan districts. Data for the regional units (ADM2) is assembled combining satellite sources for socio-economic activity, geo-located aid disbursements, and traditional household surveys. We find statistically significant positive and persistent effects of aid shocks on nighttime luminosity. Mapping nightlights to economic activity, the results suggest that the economic magnitude of these effects is small, but significant – with a multiplier between 4 and 5 in the long-run. The VAR addresses endogeneity concerns associated with non-random aid assignment.

Keywords: Official Development Assistance, Aid Impact, Regional Economic Growth and Development, Satellite Nighttime Luminosity, Spatial Panel VAR.
JEL Classification: R11, R2, O11, O12, O22, O55.

*Corresponding author. Sam M. Walton College of Business, Department of Economics, Business Building 402, Fayetteville, AR 72701. Email: andrea.civelli@gmail.com.

†Sam M. Walton College of Business, Department of Economics, Business Building 402, Fayetteville, AR 72701. Email: AHorowitz@walton.uark.edu.

‡Ibmec, Belo Horizonte (Brazil). Email: arilton8821@gmail.com.

This research is supported with a grant from AidData (CCN 0402-82221-21-0000). Critical assistance was also provided by the Center for Advanced Spatial Technologies (CAST) at the University of Arkansas, USA. We thank Christopher Kilby, Jack Cothren, the participants of the ERSA-2015, AidData/William and Mary Seminar, and SEA-2016 for helpful comments. Sadek Yousuf provided excellent research assistance.

1 Introduction

The effect of official development aid (ODA) on recipient countries' economic growth has been the subject of intense debate in the economic development literature. OECD countries have spent more than 3 trillion dollars on foreign aid since 1970, with the explicitly stated goals of economic development, growth, and poverty reduction in recipient countries. However, little consensus on the effect of aid on growth has been achieved.[1] One impediment to consensus is how to address donors' endogenous allocation of aid across recipients, which makes establishing causality between aid and its effects difficult. Recent work by Rajan and Subramanian (2008), Deaton (2010), Clemens, Radelet, Bhavnani, and Bazzi (2012), and Galiani, Knack, Xu, and Zou (2017) have argued convincingly that earlier estimates of ODA effects may be subject to endogeneity bias. The endogeneity issue is also complicated by the difficulty of clearly disentangling donors' motivations. Kilby and Dreher (2010), Dreher, Eichenauer, and Gehring (2014), and Civelli, Horowitz, and Teixeira (2016) argue that measuring aid impact requires consideration of donors' motive.[2] Similarly, trade or geopolitical motives (see respectively Berthélemy, 2006; Alesina and Dollar, 2000), could obfuscate estimates of the ODA effects.

A second major issue obfuscating aid effects is "over-aggregation," which could mask impacts of local treatments at the national level. Tierney (2011), Dreher and Lohmann (2015), and others, argue that a subnational level is best for measuring aid effects since multiple sources of noise accumulate in country level aggregation. Likewise, aggregation in other dimensions can obscure the impact of aid on growth. Clemens, Radelet, Bhavnani, and Bazzi (2012), for example, disaggregate aid into early and long impact varieties, finding significant growth effects only with the early impact category. Similarly, aid can be disaggregated by sector (health, education, irrigation, etc.). Recently, randomized field experiments (see for example Duflo, Glennerster, and Kremer, 2008) have provided true control treatment analyses of projects' local effects. At this level, many projects appear effective.

This gap in the apparent micro and macro effects of aid has been dubbed the micro-macro paradox (Dreher and Lohmann, 2015; Mosley, 1987). However, most of the micro-level project studies do not claim a linkage to economic growth or measurement of the full spillover effects. Yet, the micro and macro effects of aid projects are linked by definition; the total impact of foreign aid upon growth must be associated with the cumulative effect of the individual projects upon growth. Consequently, a growing literature focuses on improving sub-national ODA impact measures.

This paper proposes a subnational ODA impact measurement strategy that simultaneously tackles the endogeneity and aggregation problems by combining a spatial panel vector-autoregressive (Sp P-VAR) model and multiple sources of regional data. We assess the potential of this strategy, critically discuss its limitations, and provide a demonstration using regional data from Uganda.

Utilizing a VAR model is attractive in this context because it provides an intuitive way to impose identifying restrictions that can address the endogeneity. This solution is based

[1] On this point, see for instance Clemens, Radelet, Bhavnani, and Bazzi (2012). An introduction to the ODA literature can be found in Temple (2010), Addison and Tarp (2015), and Radelet (2006).

[2] Civelli, Horowitz, and Teixeira (2016) find evidence that that altruistic motivations, which could help illuminate reverse causality, explains only a minority of aid transactions.

on the recursive orthogonalization of the covariance matrix of the estimated residuals of the VAR model. This approach has been used extensively in the empirical macroeconomics literature since Sims (1980), but has seen only limited use in the development literature. The identification of ODA effects with this methodolgy is based on a scheme that isolates the exogenous shocks to ODA by removing the endogenous component. This is achieved by jointly modeling ODA and economic activity as a single system and by assuming that structural innovations to ODA can affect economic activity contemporaneously, while economic shocks are assumed to have an impact on ODA disbursements only with a temporal lag. This scheme is natural in this context since the aid disbursement process is lengthy, with a prior commitment phase that often precedes disbursement by more than a year (Kilby, 2013).

Prior applications of the VAR model to ODA include Lof, Mekasha, and Tarp (2015), who apply a co-integrated P-VAR to 59 countries, finding that the average long-run response of income is about $4.5 - 5$ larger than an initial increase in ODA disbursements. Similarly, Juselius, Moller, and Tarp (2014) estimate individual countries co-integrated VAR models for 36 sub-Saharan African countries, finding a positive long-run impact of ODA flows on the macroeconomy for most countries. Gillanders (2016) estimates a P-VAR model for a set of sub-Saharan African countries and reports a positive, but small, increase in economic growth following a fairly substantial aid shock. We follow Lof, Mekasha, and Tarp (2015) in the identification strategy, but we estimate a fixed-effects panel VAR as in Gillanders (2016). All of these papers focus on the dynamics of aid at the national level; importantly, our paper differs from these studies in adopting a sub-national perspective.

The use of disaggregated data brings the advantage of a more direct link between ODA and growth. However, the combination of the panel structure necessary to estimate the P-VAR and the regional disaggregation of the data introduces some unavoidable costs and limitations. First, the P-VAR requires a sufficiently balanced panel structure, with very limited missing observations and frequently sampled time series of the endogenous variables (annual observations, at least). Second, while the cross-section dimension helps increase the number of observations, the dynamic relation between endogenous variables is derived from the time dimension of the model, which requires a sufficiently large time sample. Third, it is difficult to obtain sub-national variables at annual frequency from official statistics to use in the VAR model. Finally, the panel estimation imposes the assumption of homogeneous effects across units, which implies the measurement of an average effect over the whole country.

We address these difficulties by combining information from multiple and distinct data sources. With regard to the main obstacle of data availability for P-VAR estimation in the sub-national low-income country context, we side-step the problem by utilizing nighttime luminance (nightlight) data as proxy for economic activity. ODA disbursements at the sub-national level are available from the AidData Consortium's geo-coded project mappings. We then map nightlight to more conventional economic measures, such as consumption expenditure per household, using geo-coded data from living standard measurement type surveys conducted by the Uganda Bureau of Statistics. We also rely on other geographic information system datasets to obtain data about rainfall, population dynamics, land use, and the surface size of the geographic units.

We choose Ugandan districts as the baseline cross-section units of the analysis in an attempt to achieve a sufficiently balanced panel, adopting AidData's districts geographic definitions. This yields 35 districts for which both nightlights and ODA disbursements are observed frequently enough over the sample $1996 - 2012$. As further discussed in Section 3, these 35 of Uganda's 112 districts contains about 50% of Ugandan population, though only 30% of Ugandan territory. The remaining districts show extremely discontinuous luminosity series, or no light at all. These non-luminous districts together account for only 1.2% of nightlight and 11% of total aid, on average. However, in order to preserve the information from these non-luminous districts, we aggregate them to create a synthetic district, which is added to the panel. We then standardize all observations by the district surface area to maintain comparability across units. Under the assumption of homogeneity of the effects of ODA across geographic locations, our strategy of synthetic district creation to preserve the information in the non-luminous districts has the cost of reducing the benefit of disaggregation for local effect measurement for that unit.

The use of nightlight satellite data to proxy income, both nationally and sub-nationally, has grown rapidly in the development literature (Chen and Nordhaus, 2011; Henderson, Storeygard, and Weil, 2012, 2011; Dreher and Lohmann, 2015). Nightlight data provides a means to side-step well-known problems associated with traditional income surveys in low-income countries, such as infrequent surveys, large informal sectors, recipient data-gathering capacity constraints, and recall errors in the absence of formal income records.

The theoretical causal mechanisms of ODA to nightlight are straightforward. As discussed in the seminal nightlight papers in the economics literature (see Henderson, Storeygard, and Weil, 2012, 2011; Chen and Nordhaus, 2011) nightlight, as measured by satellite, is highly correlated with income. Since the stated objective of ODA by the donors is income growth and development,[3] the theoretical causal chain from ODA to nightlight contains only one link. Of course, different types of aid can have different temporal lags and channels to growth, and different inherent effects on nightlight. For example, a bridge project that connects two areas with large potential economic synergies may have an immediate (within year) impact on income growth and nightlight. A "soft" aid project, such as one that improves the quality of primary education, might have multiple channels to growth and light. If the education project involves school construction or the hiring of new teachers, it may also generate income growth fairly quickly, assuming an output-gap exists. On the other hand, the income growth effect via the human capital formation channel will only be realized in the long-run. Finally, electrification projects could conceivably increase nightlight without a direct impact on income. We address these concerns by excluding power-supply projects in a robustness check and by providing some suggestive evidence on the differences in the responses to early-impact and late-impact aid disbursements.

While many papers have studied the effects of aid at sub-national levels, ours is the first (to our knowledge) to utilize Sp P-VAR estimation, nightlight data, and geo-located AidData to explore the impact of aid on sub-national growth.[4] Dreher and Lohmann (2015) analyze

[3]See the mission statement of OECD aid activities (http://www.oecd.org/about/) and note that the aid in our data set is all associated with OECD-affiliated donors.

[4]Among the others, examples of studies that have looked at the sub-national impact of aid projects on various outcomes are: Jablonski (2014) and Briggs (2012) for electoral outcomes, Crost, Felter, and Johnston (2014) and Findley, Powell, Strandow, and Tanner (2011) for conflict, and Hamilton and Stankwitz (2012) for deforestation.

ADM1 and ADM2 regions using data for World Bank projects for a large set of countries and an interacted instrumental variable approach, but they do not find significant causal effects of aid on growth. De and Becker (2015) use the AidData geo-coded ODA datasets and find a positive effect of health and water aid on the disease reduction in Malawi. Similarly, Dionne, Kramon, and Roberts (2013) study the effectiveness of sector-specific aid in Malawi within a classic two-stage allocation-impact framework. Our preference for analysis of Uganda is mostly driven by the reliability of the AidData data for this country, which was part of the very first wave of AidData releases and has undergone several updating and cleaning revisions. However, our methodology is highly scalable, with straightforward application to many low-income countries.

We find that an initial exogenous shock to ODA is associated with a statistically significantly positive impact response of nightlight, and the positive response persists for more than ten years. Mapping nightlight to economic activity, we find economic magnitudes in line with the mildly optimistic estimates usually reported at the national level. The results of our baseline specification suggest a cumulative ten-year multiplier of ODA on per capita expenditure, including the spatial component of the transmission channel, between 4 and 5. Similarly, the estimates of the overall short-run multiplier are around 1.5 for the response to a temporary shock. Finally, we quantify the rebounding effect due to the spatial structure of the model to account for about 20 to 50% of the total effect of ODA shocks.

The remainder of the paper is organized as follows. Section 2 introduces the empirical strategy and discuss some of its limitations. Section 3 describes the data. Section 4 presents the empirical results. Section 5 concludes and discusses extensions.

2 Empirical Strategy

The empirical strategy of this paper relies on the use of a spatial panel vector-autoregressive model (Sp P-VAR) for the analysis of sub-national effects of foreign aid on economic growth. This approach addresses two main concerns in the aid effectiveness literature. First, the VAR model allows us to address the endogeneity of aid disbursements by imposing some restrictions on the dynamics of the model based on intuitive economic considerations. Second, analysis at the sub-national level can help address the difficulties in the measurement of aid effectiveness due to over-aggregation of aid types with different characteristics.

The VAR is a linear model that requires suitably balanced panel structure for estimation. Although the cross-section dimension helpfully increases the number of observations for estimation, the dynamic relation between the endogenous variables of the model can only be inferred from a sufficiently large time sample. Specifically, the low frequency of survey data (e.g., LSMSs or expenditure surveys) in many low-income countries prevents us from directly estimating a Sp P-VAR model of sub-national economic growth using such data. Therefore, we follow a two-step procedure in order to exploit the identification advantages of the VAR as well as the advantages from ODA disaggregation. We first estimate a Sp P-VAR in nightlights and ODA with data at district level, exploiting the characteristic of nighlights data of being available at annual frequency for virtually any level of geographic disaggregation. We then map the effects of ODA via lights to economic activity by adapting the predictive equation of Henderson, Storeygard, and Weil (2012) to our context.

2.1 Specification of the Spatial P-VAR model

We start with the reduced-form representation of a fixed-effects Sp P-VAR model in equation (1), in which we include a time autoregressive lag, a spatial autoregressive component, and possible exogenous explanatory variables. This model can be generalized to higher orders of autoregressive lags on both the time and space dimension, but for sake of clarity we simply focus on the specification used in the analysis:

$$Y_{i,t} = A_1 Y_{i,t-1} + S\bar{Y}_{i,t} + BX_{i,t} + u_i + e_{i,t}. \tag{1}$$

In this equation, $Y_{i,t}$ is the vector of the n endogenous variables of the model, in which $i = 1 \dots N$ indicates the cross-sectional units of the panel and $t = 1 \dots T$ the time dimension; $\bar{Y}_{i,t}$ corresponds to the spatial autoregressive lag of the endogenous vector (fully defined below); and $X_{i,t}$ is the vector of exogenous variables. The vectors of panel fixed-effects and idiosyncratic errors are u_i and $e_{i,t}$ respectively, while time fixed-effects are embedded in the model by subtracting from each variable its cross-sectional mean before estimation.[5] Finally, A_1, S, and B are $(n \times n)$ coefficient matrices.

In our application, cross-section units correspond to Ugandan districts (including the synthetic district), while t is expressed in years. The endogenous vector includes the logs of the ratio of nightlight and aid disbursements to the district surface area, $light_{i,t}$ and $oda_{i,t}$ respectively

$$Y_{i,t} = \begin{bmatrix} oda_{i,t} \\ light_{i,t} \end{bmatrix}. \tag{2}$$

The choice of normalizing variables by district area follows the luminosity literature (see for example Henderson, Storeygard, and Weil, 2012, 2011; Chen and Nordhaus, 2011; Dreher and Lohmann, 2015). This is standard when we want to approximate income dynamics with luminosity for two reasons. First, light growth occurs both at the extensive and intensive margin: that is, dark areas transitioning to light as well as the nightlight signal becoming more intense. Since the measurement of light by the DSMP satellites is top-coded, significant upper bound truncation in urban areas is not unusual. Where truncation occurs, light growth can only occur at the land area extensive margin. A second reason to prefer measures per land-unit area is the public goods nature of nightlight in many settings. For example, in a typical low-income country, the nightlight emissions and capacity are likely insensitive to the number of members in the household. The model is specified in log-levels and it estimates the elasticity of the response of luminosity to aid disbursements.

The spatial lag for district i is defined as a weighted average of the values of its neighbor districts, $\bar{Y}_{i,t} = \sum_{j \neq i} w_{i,j} Y_{j,t}$. The neighbors of a district are those that share a common border with it. The weights $w_{i,j}$ for district i are defined by the row-normalized entries of the contiguity matrix of the district map. Entry (i, j) of the contiguity matrix is either 1 if districts i and j are neighbors or 0 if j is not a neighbor of i, with a diagonal of zeros by construction. We collect the spatial weights together in matrix W. The spatial lags of

[5] This demeaning of the data is equivalent to including period dummies in the exogenous vector X, but it has the advantage of restraining the number of parameters to estimate. The number of parameters would increase in the size of the time sample, and the advantage is particularly significant in small-size models like ours.

any other exogenous variable can be constructed following the same approach. This is very helpful because spatial instruments derived from exogenous variables enhance the reliability of the estimation of a dynamic model with spatial components. As an exogenous variable we include in $X_{i,t}$ the log of total annual rainfall at district level.[6,7]

2.2 Estimation strategy

The Sp P-VAR in (1) is estimated over the sample $1996 - 2012$ using a modified version of the Love-Zicchino Stata package (Love and Zicchino, 2006), with the modification being necessary to especially accommodate the spatial autoregressive component. The estimation methodology is an extension to the multi-equation case of the dynamic panel GMM approach of Arellano and Bond (1991a). This approach is suitable for small time dimension and large cross-section samples, and it assumes that the idiosyncratic error terms have finite moments and in particular $\mathbb{E}(e_{i,t}) = 0$, $\mathbb{E}(e_{i,t}e'_{i,t}) = \Sigma$ (constant covariance matrix across units), $\mathbb{E}(e'_{i,t}e_{i,s}) = 0$ for $t \neq s$ (no serial correlation over time), and $\mathbb{E}(e'_{i,t}e_{j,t}) = 0$ for $i \neq j$ (no cross-section dependence).

A within estimator that removes the fixed-effects in a dynamic panel setup such as (1) introduces an additional bias in the estimation of the coefficient matrix, due to the correlation between the transformed residuals and the transformed vector of endogenous variables. As normally done in the context of dynamic panels, instrumentation with lags of Y_i is used to deal with this problem. Since the estimation procedure removes the fixed-effects using a forward orthogonal transformation, the first lag of Y_i would be a suitable instrument if the residuals are not autocorrelated. However, the estimation procedure can be straightforwardly adapted to cases with short-term autocorrelation by distancing the lag of the instruments. Since we found some mild serial correlation of the first order in the residuals of our estimated models, the results we report are based on the use of the second and third-lags of Y_i which provides one over-identifying condition for the GMM estimation.[8]

The GMM estimates of the model coefficients are, however, robust to arbitrary patterns of heteroskedasticity and autocorrelation within individuals when the residuals covariance is modeled clustering by district and estimated by two-step GMM (see Roodman, 2009, for a discussion at the panel level).[9]

[6] Although other definitions of the weighting scheme are possible, using spatial matrices based on border contiguity is probably the most common approach in spatial analysis.

[7] Details regarding the construction of variables, data sources, and the full definition of the cross-sectional districts are provided in Section 3 and Appendix A.

[8] The relatively small size of our panel suggests particular caution with the proliferation of the instrument. Hence, a parsimonious instrument selection is preferable for the baseline model in the analysis. Increasing the number of over-identifying GMM conditions seems to quickly lead to over-fitting problems as reflected by the implausibly high p-values with which the the Hansen's test for the validity of over-identifying conditions is passed (Roodman, 2009). Also, notice that the removal of fixed-effects is independent of the ODA endogeneity issue and the instrumentation does not aim to solve this.

[9] The estimation of the residuals covariance matrix used to compute the impulse responses must reflect these assumptions. The original Love-Zicchino package works under the assumption of homoskedasticity only; our modification introduces clustering and within district autocorrelation in the computation of the covariance. However, we maintain the assumption of no cross-section dependence of the residuals in the computation of the structural shocks of the model, as further discussed in this section and, in part, justified by our analysis in Section 4.

On the contrary, cross-section dependence has to be explicitly accounted for. One relevant source of this dependence with contiguous geographic units is clearly spatial autocorrelation. The use of the spatial lags attenuates this concern. These spatial terms are not exogenous by construction because they are correlated with the current innovations to $Y_{i,t}$, but they can be handled in the GMM framework adopting the same instrumenting approach as for the VAR vector of endogenous variables. A suitable set of instruments includes the time-lags of the spatial lag itself, as well as the spatial lags of the exogenous variable. Further cross-section dependence due to common factors that affect all units simultaneously is controlled for by also including time fixed-effects in the estimation.[10]

A last, but important caveat, in the use of this type of pooled VAR model is that it relies on the assumption of homogeneous slopes across units. This choice is in part dictated by the short time sample of the data, which prevents any serious attempt of fitting heterogeneous models in which the parameters vary by district and/or by time period. However, due to the regional scale of the study, we believe it is fair to assume only moderate socio-economic heterogeneity across districts. Similarly, due to the limited degrees of freedom in the time dimension, we cannot reasonably test for poolability. In principle, however, this could be done with a test for parameter restrictions like the Wald test or the LR test where the unconstrained model allows for unit specific slopes and the restricted model imposes the common slope coefficients.[11] A somewhat easier approach would be to test for poolability at a individual equation level, and in this case a test based on the Mean Group estimator of Pesaran and Smith (1995) would be more feasible.

2.3 Structural form and impulse response analysis

Let us now consider the structural counterpart of model (1) in which the contemporaneous effects between the endogenous variables of district i are explicitly represented by

$$Y_{i,t} = \Gamma_0 Y_{i,t} + \Gamma_1 Y_{i,t-1} + S_0 \bar{Y}_{i,t} + B_0 X_{i,t} + \eta_i + v_{i,t} \tag{3}$$

where our attention is especially focused on $v_{i,t}$, the vector of orthogonal structural residuals of the model, while Γ_0, Γ_1, S_0, and B_0 are $(n \times n)$ matrices of structural coefficients, and η_i is the vector of fixed-effects. The reduced-form (1) is obtained from (3) by solving it for $Y_{i,t}$, after defining $A_0 = (I_n - \Gamma_0)^{-1}$ and setting $A_1 = A_0\Gamma_1$, $S = A_0 S_0$, $B = A_0 B_0$, $u_i = A_0 \eta_i$, and $e_{i,t} = A_0 v_{i,t}$.

A_0, hence, also defines the relation between structural and reduced-form residuals of the model. The VAR approach assumes that the reduced-form residuals are a linear combination

[10] Bouayad-Agha, Tutpin, and Védrine (2013) and Brady (2011) use the GMM instrumentation framework to estimate spatial effects in a spatial dynamic panel model. We extend their approach to the P-VAR model. Alternative methods for the estimation of spatial P-VAR models have been proposed in the recent literature. Beenstock and Felsenstein (2007) adopt an IV approach and a bias correction solution to the incidental parameter problem; Márquez, Ramajo, and Hewings (2015) add the spatial terms to the vector of endogenous variables of the VAR system and estimate individual models for each region; Badinger, Muller, and Tondl (2004) spatially filter their data before applying the GMM to a dynamic panel model. These are all valid alternative options, but we prefer to remain as close as possible to the well-known Arellano-Bond framework.

[11] Although these tests have known asymptotic distributions, their small-sample performance has to be usually explored empirically.

of the unobservable orthogonal structural innovations. As in the non-spatial case, knowledge of $v_{i,t}$ is necessary to give a structural interpretation to the impulse response functions (IRF) of the variables of the VAR to the shocks of the model. Even though it is not possible to fully recover the structural innovations from the reduced-form residuals, since equation (1) is structurally under-identified, the structural shocks can be reconstructed by means of an orthogonalization of the estimated vector of residuals.

Econometrically, the orthogonalization is achieved through a factorization of the estimated covariance matrix of the residuals Σ. However, since the factorization is not unique, it is necessary to adopt some selection criteria to choose a specific one. In practice, given the linear setup of the VAR model, a convenient approach to selecting an orthogonalizing scheme is to impose some restrictions on the contemporaneous relations between structural and reduced-form innovations in A_0, providing an economic justification in support of the restrictions. One of the most common of these schemes is known as the Cholesky identification approach because it exploits the Cholesky decomposition to factorize Σ; the Cholesky factor is then matched to A_0.

We can express the covariance matrix of $e_{i,t}$ in function of its lower-triangular Cholesky factor D as $\Sigma = DD'$, where D^{-1} is the factorizing matrix that orthogonalizes the covariance matrix of $e_{i,t}$. From $e_{i,t} = A_0 v_{i,t}$ and normalizing to one the variance of the structural shocks so that $\mathbb{E}(v_{i,t}v'_{i,t}) = I_n$, it is easy to see the Cholesky identification scheme imposes a set of zero-restrictions on the off-diagonal elements of A_0 (suitably re-ordered) to uniquely identify A_0 as a lower-triangular factorization matrix. With a vector of two endogenous variables as here, only one off-diagonal element exists and, hence, only one restriction is necessary. We postpone the economic discussion of the identifying assumptions to Section 2.4.

Given this structural identification of the innovations, model (1) allows us to compute a first type of IRF of the variables of a district to shocks to variables in the same district, in which time-only impulse responses are considered. Let $\Lambda_{j,h}$ indicate the vector of the n responses, then

$$\Lambda_{j,h} = A_1^h A_0 \lambda_j \tag{4}$$

where $j = 1 \dots n$ indicates the structural impulse with which the system is shocked, λ_j is a $(n \times 1)$ selection vector equal 1 at position j and zero elsewhere, and h tracks the horizon of the response function. The assumption of homogeneity across units implies a common shape of the responses across districts, whereas in a model with region-specific coefficients the IRF must be indexed by i. Although the estimates of A_1 and A_0 are obtained correctly controlling for the spatial structure of the model, $\Lambda_{j,h}$ corresponds to the time-only impulse response because it does not embed the space dynamics in the response to the shocks. In this respect, it must be interpreted as the simple and direct effect of a shock in a district on the endogenous variables of that same district.

In addition to the direct time IRF, however, it is interesting to also study the time-space IRF of the model, in which spatial spillovers allow the shocks in one district to affect the dynamics of the variables in other districts and the following rebounding effects are explicitly taken into account in computing the IRF. To this end, we rewrite the model by stacking (1) over the cross-sections

$$Y_t = \mathbb{A}_1 Y_{t-1} + \mathbb{SW} Y_t + \mathbb{B} X_t + u + e_t \tag{5}$$

where Y_t is the $(nN \times 1)$ vector of endogenous variables for all the districts stacked by i, and

X_t, u, and e_t are analogously defined. Furthermore, the coefficient matrices are obtained from those in (1) as follows: $\mathbb{A}_1 = (I_N \otimes A_1)$, $\mathbb{S} = (I_N \otimes S)$, $\mathbb{W} = (W \otimes I_n)$, and $\mathbb{B} = (I_N \otimes B)$. By defining $\mathbb{A}_0 = (I_N \otimes A_0)$, the stacked reduced-form residuals can also be expressed as $e_t = \mathbb{A}_0 v_t$, which clearly highlights how the district level relation between structural and reduced-form residuals is preserved with the stacked vectors by the block-diagonal structure of $\mathbb{A}_0$. In particular, the covariance matrix of the stacked reduced-form residuals is a block-diagonal matrix as well, and can be written as $\mathbb{E}(e_t e_t') = (I_N \otimes \Sigma)$, and the identification of A_0 would directly apply to each block of $\mathbb{A}_0$.

We next solve equation (5) for Y_t in order to be able to compute the time-space IRF

$$Y_t = \tilde{\mathbb{A}}_1 Y_{t-1} + \tilde{\mathbb{B}} X_t + \tilde{u} + \mathbb{M}_0 \mathbb{A}_0 v_t \tag{6}$$

where $\mathbb{M}_0 = (I_{nN} - \mathbb{S}\mathbb{W})^{-1}$, $\tilde{\mathbb{A}}_1 = \mathbb{M}_0 \mathbb{A}_1$, $\tilde{\mathbb{B}} = \mathbb{M}_0 \mathbb{B}$, $\tilde{u} = \mathbb{M}_0 u$, and the residuals are explicitly expressed in function of the structural innovations v_t. The matrix $\mathbb{M}_0$ embeds the effects of the spatial component on both the dynamics of the model, represented by $\tilde{\mathbb{A}}_1$, and the impact of the structural shocks on the system, since $\mathbb{M}_0$ pre-multiplies $\mathbb{A}_0$.

The structural identification of the time-space IRF relies on the contemporaneous restrictions of the Cholesky approach for $\mathbb{A}_0$, on one side, and the instrumental variable identification of S in the GMM estimation of the model, on the other.[12] By merging information from the within district ordering of the endogenous variables and from the spatial structure of the model, we follow the spirit of the identification strategy proposed by Di Giacinto (2010). In our case, however, the identification is streamlined by the separate estimation of S, which allows us to keep $\mathbb{M}_0$ and $\mathbb{A}_0$ independent. On the contrary, Di Giacinto (2010) estimates the reduced-form in (6) and imposes a set of restrictions on the joint term $\tilde{\mathbb{A}}_0 = \mathbb{M}_0 \mathbb{A}_0$.

We can now compute the time-space IRF, indicated by the $(nN \times 1)$ vector $\tilde{\Lambda}_{l,h}$

$$\tilde{\Lambda}_{l,h} = \tilde{\mathbb{A}}_1^h \mathbb{M}_0 \mathbb{A}_0 \tilde{\lambda}_l \tag{7}$$

where $\tilde{\lambda}_l$ is the selection vector of the stacked shocks, with $l = 1 \dots nN$ now. Any shock in any district can, in principle, have an impact effect on any variable in any district since $\mathbb{M}_0$ does not have a block-diagonal structure. Similarly, the transmission of the impulses across districts is then compounded by the spatial effects incorporated in $\tilde{\mathbb{A}}_1$. Furthermore, the responses of the same variable in two different districts to the same shock in the respective district will not unfold in the same manner due to the different weighting schemes and neighbors of each district – reflected by the differences in the rows of W.

It is clear that the number of impulse responses rapidly increases, especially for large cross sections, making it unfeasible to report all of them. For this reason, we compute and report the cross-section average of the time-space responses of the variables of a district to a shock in that district, which can be directly compared to the corresponding time-only responses given by (4) in order to assess the average effect of the spatial component of the model. These average IRF will be denoted by $\bar{\Lambda}_{j,h}$, where $j = 1 \dots n$ indicates the impulse variable as in the time-only IRF. Formally, $\bar{\Lambda}_{j,h}$ is computed as

$$\bar{\Lambda}_{j,h} = \frac{1}{N} \sum_{k=1}^{N} \tilde{L}_k \tilde{\Lambda}_{j+n(k-1),h} \tag{8}$$

[12]Since the weighting matrix W is given, knowing S is sufficient to fully determine $\mathbb{M}_0$.

where $\tilde{L}_k$ is an $(n \times nN)$ selection matrix with the n vectors $\tilde{\lambda}'_{1+n(k-1)} \dots \tilde{\lambda}'_{nk}$ as rows.

2.4 Solution to endogeneity of aid allocation

The local disaggregation of ODA likely mitigates many of the issues that contribute to the non-random allocation of ODA, such as the strategic interplay between donor and recipient at national level or the enormous heterogeneity across recipient countries. With respect to recipient heterogeneity, the concern is attenuated as estimation is across regions of more uniform climatic, institutional, and socio-economic structure. Similarly, the capability of Ugandan sub-national governments to implement effective strategic play with multiple OECD donor countries is also likely quite limited when compared to nation-states. Nevertheless, addressing the endogeneity of ODA disbursements is also a priority at the sub-national level.

In a VAR framework, a solution to the endogeneity problem is readily available through the identification of exogenous structural innovations to ODA that can be employed to correctly measure the effects of ODA on nightlights. Following the work of Sims (1980) in the empirical macroeconomics literature, the VAR is one of the most commonly used empirical tools and the choice of the identification restrictions is typically justified by an intuitive transmission mechanism, rather than an exact structural economic theory. With only two endogenous variables ($n = 2$) in the Sp P-VAR model, this narrative based justification is easily accomplished. In fact, there are only two possible cases to achieve identification of structural shocks: either we assume that ODA responds to a nightlight shock with one temporal lag or, vice-versa, that luminosity responds to an ODA shock with a lag.

We follow Lof, Mekasha, and Tarp (2015) and assume that nightlight shocks can impact ODA disbursements only with a temporal lag, allowing then nighlights to respond to ODA shocks within the period. The rationale for this identification strategy is that a local ODA disbursement could impact economic local activity via either the demand or supply side relatively quickly (within a year), while the response mechanism of ODA to an exogenous increase in nightlight is typically lengthy. This transmission mechanism is consistent with the complex process that culminates in aid allocation decisions and disbursements, as documented by Kilby (2013), for instance. In Section 4, and in Section S3 of the online Appendix, we discuss the importance of this assumption for the results since they would differ under the alternative identification scheme. However, we find it difficult to find a simple and plausible economic justification of this alternative transmission mechanism.

It should be noted this identification strategy only imposes a zero-restriction on the element of A_0 corresponding to the response of aid to nightlight, but it does not force any specific behavior on the response of lights to ODA shocks per se. This restriction, as any other assumption, needs to be justified and economically reasonable, whereas the response of nightlights to ODA is allowed to be freely estimated from the data. In principle, this response could be null or even negative and, in this respect, our identification strategy takes a fairly neutral stance on the effect of our main interest.

Once the structural shocks are identified, the effects of aid on luminosity can be estimated by computing the impulse response of nightlights to an initial shock to ODA. The relative magnitude of the shock and the response will represent the elasticity of luminosity to ODA that is at the heart of our analysis. The response functions also provide a dynamic representation of the transmission channel of transitory shocks in the short and long-run.

Other tools are also used to assess the transmission channel of ODA shocks; in particular, we analyze the long-run cumulative responses of lights to a one-time shock, the response to permanent shocks, and variance decomposition of forecast errors.

One important limitation of the VAR model is that, although the structural identification correctly reveals the dynamic interaction of the endogenous vector of variables in the system, the structural mechanism is still identified conditional on the model specification itself. Omitting other relevant endogenous variables or not controlling for relevant confounding factors might affect the dynamics of the impulse response functions and the magnitude of the identified responses. This issue has no easy solution in our context because of the difficulty of obtaining reliable sub-national data at annual frequency for a low-income country.

These concerns are mitigated in part through robustness checks and the use of fixed effects. Fixed-effects in the P-VAR model eliminates the effects of district level time invariant characteristics. These can include differences in governance, cultural, climatic, and socio-economic factors that vary across district, but not over time. Furthermore, we also include time-effects in the model by time-demeaning the variables, which allows us to control for common time-varying factors such as business cycles or political trends at the national level. Similarly, we use as a control variable rainfall at the district level, which is a meaningful factor explaining growth differentials in emerging economies, such as Uganda, where subsistence farming is still very important. Finally, we re-estimate the P-VAR with ODA and nightlights standardized by population, in order to control for an important factor that can influence both light emission and ODA allocation.

2.5 Linking economic activity to luminosity

After the impulse response of lights to ODA is estimated with the P-VAR, it remains to connect this impact to a more standard measure of economic activity. For this we rely on the predictive stage of Henderson, Storeygard, and Weil (2012)'s approach, in which a statistical measure of economic activity (official GDP, typically) is regressed on lights.[13] We adapt this methodology in a straightforward manner to explore the link between the growth in household real expenditure at the district level and the change in luminosity for the same set of cross-sectional units used in the P-VAR analysis. The predictive equation simply reads

$$x_{i,t} = \psi light_{i,t} + \epsilon_{i,t}, \tag{9}$$

where $(\epsilon_{i,t})$ are i.i.d. random variables, $\mathbb{E}(\epsilon_{i,t}) = 0$ and $\mathbb{E}(\epsilon'_{i,t}\epsilon_{i,t}) = \sigma^2 I$. Equation (9) is written in log-linear version as a two-period panel between 1999 and 2009. As in the P-VAR vector (2), $light_{i,t}$ is the log of the ratio of nightlight to the district surface area for district i in period t. We measure $x_{i,t}$ with either the log of the average household weekly consumption expenditure or, for robustness, the average household monthly expenditure in non-durable goods. Henderson, Storeygard, and Weil (2012) measure $x_{i,t}$ at the country level

[13] Henderson, Storeygard, and Weil (2012)'s methodology uses the estimates from this stage to ultimately infer the unobservable true growth of the underlying economic activity from an optimal combination of multiple signals correlated with it, such as GDP and lights. Since the purpose of our paper is to primarily document a transmission channel from ODA to household expenditure, we only focus on the first stage of their methodology.

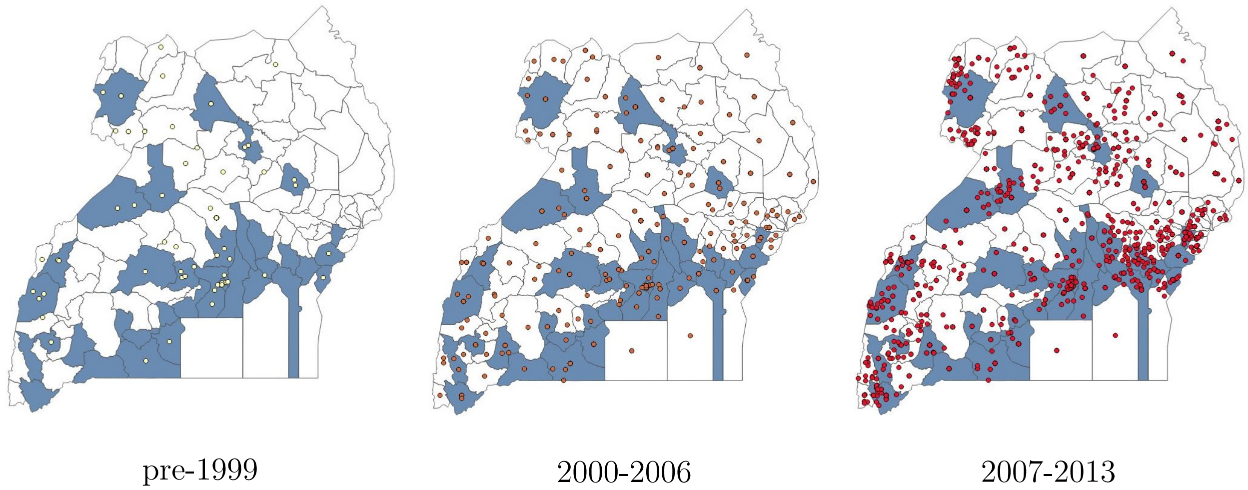


Figure 1: Distribution of the ODA project in Uganda since 1991. Data provided by AidData. Solid gray indicates the 35 districts included in the sample used in this study. White corresponds to the districts in the synthetic region.

using the log of national GDP instead; our estimates of $\hat{\psi}$ are very much consistent with theirs. The model is estimated using panel fixed-effects and robust standard errors. In line with the use of spatial components in the Sp P-VAR model, we control for potential cross-section dependence due to spatial factors by applying the Eigenvector spatial filter to the data before estimating the model (see, for instance, Griffith, 2003; Griffith and Peres-Neto, 2006, for a discussion of this filtering technique). Comparing the estimates of model 9 with and without pre-filtering the data we find a quite small effect of the spatial components in this predictive equation.

3 Data Description

Panel Structure – The panel used for the estimation of the P-VAR model covers the sample $1996-2012$, $T = 17$, and includes a cross-section of $N = 36$ sub-national regions. Details regarding the synthetic district can be found in Appendix A. Figure 1 shows the geographic location of the cross-section units of the analysis on the Ugandan map. The districts that are part of the synthetic district are white colored in the figure; the other 35 districts are illustrated by the solid gray areas instead.

ODA Data – Data for the aid projects is drawn from AidData's Uganda Geo-coded Dataset (Release IV). This dataset maps 420 projects over the sample $1981-2013$ distributed across the 112 Ugandan districts. The geographic precision of the disbursements for each project is reported on a scale from 1 to 8, where 1 indicates the knowledge of the exact geographic coordinates of the disbursement and 8 corresponds to projects at the central government level.[14] We focus on the spatial aid-growth nexus by filtering the projects to precision codes $1-3$, which correspond to ADM2 or lower administrative levels. This reduces the distance between where aid is spent and the region where it may produce positive

[14]Note that the precision scale skips the classification code 7 for unspecified reasons.

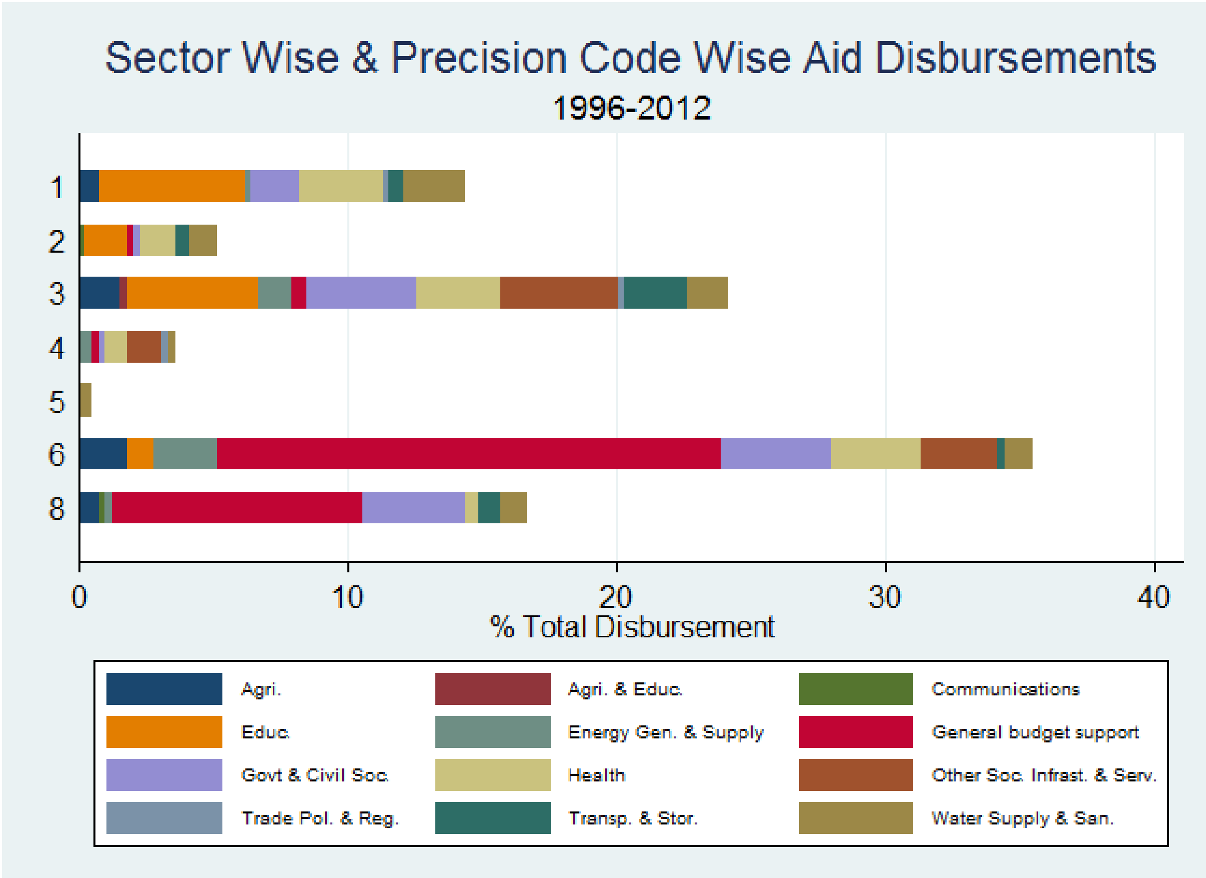


Figure 2: Disbursements classification by geo-location precision and by sector of activity.

economic effects. Figure 1 illustrates the spatial distribution of projects with precision $1-3$ over three time intervals. The number of precision 1 to 3 projects and the territory covered by them increase over time, especially in the last years of the sample.

In considering the sample $1996-2012$ for P-VAR estimation, note that before 1996 only two districts received aid for precision $1-3$ projects: the capital metropolitan area Kampala and, for a couple of years, the adjacent district of Wakiso, which includes some capital suburbs. We do not include these observations in the panel for two reasons. First, having only two cross-section units for so many periods is not desirable for our empirical model.[15] Second, even though recorded with high precision codes, these are projects close to the seat of government at a time when government institutions were often the primary beneficiaries of aid. With respect to our goal of local identification of the effect of aid, these projects likely match the criteria only weakly.[16] We also exclude 2013 from the sample due to some anomalous disbursements in two districts of the Northern region of the country that suddenly increase a hundred fold. These districts are part of the synthetic unit; the total disbursements of the synthetic district became five times bigger than those of the lit districts

[15]Furthermore, luminosity data is available only from 1992.

[16]For instance, the description of one of the projects mentions general government and civil society as main purpose. Two other projects are for the construction of the stadium and the international airport in Wakiso.

in 2013, whereas they are usually ten times smaller. This change makes this last observation behave like an outlier.

In addition to the geo-location of individual projects, the AidData dataset includes the annual disbursement flows for each project. Total district ODA is computed as the log of the ratio between the sum of aid disbursements for precision $1-3$ projects in a district (measured in real 2005 US dollar) and the district area measured in squared kilometers. Figure 2 decomposes the ODA disbursements by precision code and sector of activity for the sample $1996-2012$. This is a useful illustration of the implications of our geographic disaggregation strategy for the ODA disbursements. Two observations are worth noting. First, precision codes $1-3$ account for roughly half of the ODA disbursements. The purpose of our empirical approach is to separate the possible effects of locally circumscribed projects on local economies from the impact of a large number of projects that disburse funds at higher administrative levels. As noted, aid to the national government and ministries likely have a less direct connection to local growth than a project like a bridge or road. Second, the decomposition by sector also reveals a quite different structure between regional and national disbursements. Local disbursements exhibit a higher share of projects in education and agriculture, water sanitation, infrastructure and transportation, and health. On the contrary, general budget support - the main category for precision codes 6 and 8, are typically less focused geographically.

Figure 2 also shows that the share of projects related to energy generation and energy supply, which can inflate lights emission without an increase in economic activity, represents only about 5% of the precision $1-3$ disbursements. In one of our robustness checks, we repeat the analysis after removing these projects and obtain similar results. In principle, the information about project sector can be used to more accurately distinguish between projects which might affect economic growth more in the short-run from projects with a longer run impact (as proposed by Clemens, Radelet, Bhavnani, and Bazzi, 2012). This classification can be exploited in our study to shed some more light on the structural identification of the short-run effects of aid on luminosity. Unfortunately, the AidData dataset does not allow us to exactly replicate the Clemens, Radelet, Bhavnani, and Bazzi (2012)'s methodology because only 3-digit purpose codes are available. Moreover, at the precision levels used in our analysis, only a third of the projects are classified as early impact, leaving us with a quite uneven time and space distribution over districts. However, we can assemble a smaller set of 22 districts for which this distinction is feasible and conduct some analysis of the different implications of these two types of aid on the dynamics of the system. The results of this exercise, broadly consistent with the underlying scope of this classification of ODA, are in support of the structural identification scheme adopted in this paper.[17]

Finally, we also run a robustness check using aid project data from a second data base reporting disbursements exclusively for the World Bank as a single donor. This data base records aid projects financed by the International Bank for Reconstruction and Development (IBRD) and the International Development Agency (IDA), and it provides some advantages in terms of geographic precision and coverage over time (see the online Appendix, Section 3.2.)

[17] The details of the construction of this dataset are given in Appendix A. The results of the analysis are discussed in the empirical section of the paper below and Section 3.2 of the online Appendix.

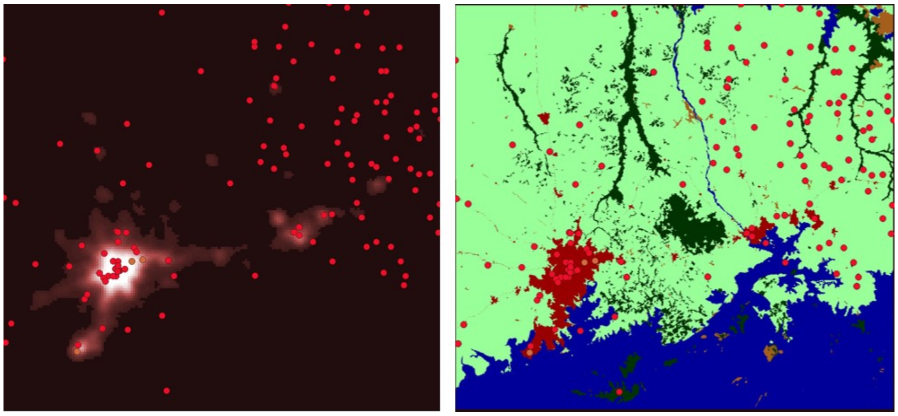


Figure 3: Comparison between land coverage and night luminosity. On the left side, the night lights are reported for on an area of about 30,000 Km^2 around the capital Kampala, for 2010. The red dots indicate ODA projects locations. The intensity of lights is represented in a black to white scale. On the right side, land coverage for the same area and the same year is shown. The color legend of the different uses is: red for urban coverage, dark green for forests, light green for agriculture and pasture.

Nightlight Data – Nightlight data is obtained from the Defense Meteorological Satellite Program (DMSP). These images are annual composites processed prior to release. Since data for distinct years can be provided by different satellites, an inter-calibration is applied to harmonize nightlight values across across years-satellites (see Appendix). As noted, persistent nightlight from this source is not detectable in all of Uganda's 112 districts over the DMSP sample, $1992 - 2012$. Figure 3 illustrates AidData project mapping overlaid with the 2010 nightlight image (on the left side panel) and land coverage (on the right hand panel) for a small region around Kampala, the capital of Uganda. In the land-cover image, red is used to indicate urbanized areas, while light green corresponds to rural and agricultural land.[18] Not surprisingly, nightlight signals are usually associated with greater urbanization. On the contrary, some rural regions do not produce any detectable luminosity signal for time intervals of several years. In any case, note that ODA disbursements, the red dots, are found in urban areas as well as rural areas with little detectable nightlight. For this reason, the construction of the synthetic region is important to avoid the complete loss of the information coming from the darker regions. Overall, 7 of the 35 individual districts have continuous detectable nightlight signals and received aid every year since 1996; the large majority of them receive ODA every year after 2000. Also the synthetic district exhibits positive ODA and nightlight signal for the entire sample.

Household Surveys – We construct the measures of economic activity at the regional level necessary to estimate the predictive stage of Henderson, Storeygard, and Weil (2012) using the Uganda National Household Surveys (UNHS) administrated in 1999 and 2009.[19] We present estimates based on both the district average household weekly consumption expenditure and the average household monthly expenditure in non-durable goods. Household

[18]The full legend of the land-cover colors is explained in the caption of Figure 3. This image is an authors' elaboration of (NASA) Landsat 7 multi-spectral satellite data.

[19]These surveys are similar in design to the World Bank Living Standard Measurement Surveys (LSMSs).

income data is also available from the UNHS, but given the well-known problems with income data in such contexts (e.g., high levels of informal activity, resistance to disclosing income, and recall error in the absence of written records), we believe that change in expenditure would be a better indicator of economic growth. Moreover, income and expenditure should be highly correlated in this poor environment with relatively little saving. Since aid disbursements may impact household durable expenditures as well, our impact estimates could be interpreted as conservative with respect to total expenditure growth, and by extension, to income growth.

Following Henderson, Storeygard, and Weil (2012), we then use the logs of the average expenditure in real 2005 US dollar terms by district as the dependent variable in the fixed-effects long-run model (9), and the log of lights per squared kilometer described above for the independent variable. Paralleling the panel of the main P-VAR estimation, we use the 35 districts and the synthetic district as cross-sectional units in the estimation of (9).

Rainfall – We start from the monthly series of total terrestrial precipitation (or rainfall) provided by the Climate Hazards Group (CHG) (see Funk, Peterson, Landsfeld, Pedreros, Verdin, Shukla, Husak, Rowland, Harrison, Hoell, and Michaelsen, 2015), in collaboration with the U.S. Agency for International Development (USAID) and the National Oceanic and Atmospheric Administration (NOAA). These series report data on a resolution grid of $.05 \times .05$ degree latitude/longitude. We compute the total annual rainfall by pixel and then aggregate it by district. The variable in the model is then computed as the log of the total annual height of rainfall per district.

Population – In a robustness check, we also utilize data for population series at district level to construct alternative normalizations of ODA and nightlights instead of normalizing by districts areas. We precisely describe the population data and its manipulation in Appendix A.

4 Empirical Results

As discussed above, our methodology consists of two linked estimation stages. In the first stage we estimate the responses of nightlights to an ODA shock in the Sp P-VAR. In the second stage we map the responses of luminosity to changes in local expenditure in the spirit of Henderson, Storeygard, and Weil (2012).

4.1 P-VAR Specification Assessment

It is worth noting again that the baseline specification of model (1) used in this Section is a Sp P-VAR(1,1) in the logs of ODA and nightlights normalized by districts' areas, which includes first-order time and spatial autoregressive lags. The choice of the lag specification is necessarily dictated by the relatively small sample size $T = 17$. However, the standard lag-selection tests do not suggest that higher order temporal lags are required.

In a panel with units geographically close to each other, cross-district spatial economic links can be significant. As a first step, then, we check for the presence of spatial autocorrelation in our two endogenous variables, *light* and *oda*. The Moran's I index, one of the most common indicators of spatial autocorrelation, for the two variables is reported in Table

S1 and discussed in the online Appendix. For conciseness, the fully detailed results and discussion are presented in the online Appendix. Luminosity exhibits a significantly positive autocorrelation in each year of the sample, while aid disbursements do not seem to have any significant autocorrelation pattern. Since the two variables are endogenously intertwined in the model, the use of spatial components in the empirical analysis is desirable.

As discussed in Section 2.2, we estimate a model with a spatial autoregressive term, equivalent to what is known as SAR in the literature. This is a feasible choice in the context of the GMM estimation of a dynamic model, since the new term can be handled by extending the same instrumentation approach of the GMM. Although some spatial autocorrelation is still found in some instances in the residuals of the ODA equation of the preferred baseline specification (see Table S2), alternative options would entail explicitly modeling the errors of the Sp P-VAR as spatially correlated terms. However, to the best of our knowledge, established estimation techniques that embed spatial errors model (SEM) into the P-VAR framework are not available yet. At least in part, the measurement of the spatial autocorrelation of the residuals might also be affected by cross-sectional dependence due to common determinants across districts. We can control for these common factors by using time fixed effects though, and we move to this aspect of the analysis next.

A key assumption of the P-VAR model is no cross-sectional dependence in the residuals. A Pesaran's CD test (Pesaran, 2004) for cross section dependence in panel data conducted on the residuals of each equation of the baseline model largely rejects the null hypothesis of no-dependence.[20] A full characterization of cross section dependence would rely on rich parameterizations of the covariance structure of the errors or some type of factor augmented model with unit-specific coefficients. In any case, given the limited size of the panel, especially in the time dimension, there would not be sufficient degrees of freedom to justify these approaches in this case.

In a an alternative and more parsimonious approach, Sarafidis, Yamagata, and Robertson (2009) devised a simple procedure to test whether including time dummies, or equivalently transforming the data in deviations from time averages as we do, is sufficient to eliminate any cross-section dependence in a dynamic panel model. We exploit this result to provide some evidence which supports the use of time fixed effects in the Sp P-VAR as an effective way of mitigating the cross-sectional dependence problem. We show in the online Appendix (Table S3 and related discussion) that replicating each equation of the Sp P-VAR as a separate dynamic panel estimated by GMM, the Sarafidis, Yamagata, and Robertson (2009)'s test does not reject the null of no residual dependence for either equation of the Sp P-VAR model.

In order to use the estimated model to compute the impulse response functions in Section 4.2, the Sp P-VAR must satisfy the invertibility conditions. The standard invertibility condition for the time-only IRF is simply based on the eigenvalues of matrix A_1 in equation (1). The baseline model satisfies this condition having two stable roots largely inside the unit circle, with the larger root of .77. The invertibility condition on the time dimension is then strongly satisfied, and possible non-stationarity of the vector of endogenous variables does not seem to be of any concern with our dataset. Similarly, we also need to confirm the model is invertible in the time-space dimension, in order to be able to use the spatial lags in the computation of the time-space IRF. This more general invertibility condition is based on

[20]Results reported in Section S1.2 of the online Appendix.

the roots of matrix $\tilde{\mathbb{A}}_1$ in equation (6) instead. This condition entails 72 possibly imaginary eigenvalues, and we find the maximum modulus of these eigenvalues is .85. The time-space invertibility conditions is largely satisfied as well.

The invertibility condition is necessary for the computation of IRF, but it can be verified only conditional on the output of the GMM estimation. The validity of the Arellano-Bond estimator used to estimate the reduced-form Sp P-VAR model also requires the endogenous variables of the model to be stationary processes. As discussed by Blundell and Bond (1998), if the endogenous variables are close to a random walk, then the difference GMM performs poorly since the original variables are weak instruments for the transformed variables. We assess this point by testing for a panel unit root in the two endogenous variables of the system.[21]

We use first the IPS test proposed by Im, Pesaran, and Shin (2003), a very popular so-called first generation panel unit root test that does not account for cross-sectional dependence, but allows for heterogeneity across panels. In spite of the spatial correlation documented above, the Pesaran CD test (Pesaran, 2004) does not reject the null of (generic) cross-section independence for either ODA or nightlights (see Section S1.3 of the online Appendix). Baltagi, Bresson, and Pirotte (2005) have shown the reliability of this test is preserved for moderate levels of spatial autocorrelation as well, which fairly corresponds to our situation where the Moran's I index does not exceeds .275. The IPS test is also used, for instance, by Beenstock and Felsenstein (2007) in a context similar to ours. We can hence apply this test with some confidence to our variables as well. The IPS test strongly rejects the null hypothesis of an integrated process with drift for both variables.

We further explore the stationarity condition with the Bai and Ng (2004)'s PANIC and the related Reese and Westerlund (2016)'s PANICCA tests, second generation tests that explicitly allow for cross-sectional correlation by decomposing the individual series of the panel into a deterministic component, a common component, and an idiosyncratic error term, the last two of which are then separately tested for stationarity. Stationary processes for both common and idiosyncratic components imply overall stationarity of the original variable itself. These tests reject the null hypothesis of integrated processes with drift for both components, consistently across different specifications of the test for both ODA and nightlights. We provide a more exhaustive illustration of these results in Section S1.3 of the online Appendix (see Tables S5 and S6).

As a last point, we check for (time) autocorrelation in the residuals in order to correctly inform the selection of the lagged instruments. We assess autocorrelation in two ways. First, we show some simple evidence that a specification of the model with only one-lag instrument, the most basic just-identified model, likely violates the GMM assumptions since it produces autocorrelated residuals. We fit a P-VAR(1) with the residuals of this model and find significant autoregressive coefficients, especially for the ODA equation. These results are reported in Table S4 of the online Appendix. However, we think this basic procedure is only plausible to check for low-order autocorrelation in our context.

As a second step, then, we return to the dynamic panel regressions applied individually to each equation of the baseline specification of model (1) used for the cross-section dependence tests. These dynamic panels are estimated by an analogous GMM procedure (see Roodman,

[21]We thank an anonymous referee for pointing out the importance of these unit root tests.

2009) and allow us to test for higher-order autocorrelations using the Arellano-Bond test (Arellano and Bond, 1991b). In line with the estimates at the previous point, we find that serial autocorrelation of the first order is present in the ODA equation, but not for lights. Furthermore, we detect autocorrelation of the fourth-order for the ODA residuals (see the result for AR(5) in column c of Tables S3).

Based on these results, our selection of instruments starts with the second-lag and excludes the fourth and higher lags. To gain precision in the estimation, our final choice for the baseline specification of the model is to use the second and third lags of both Y and $\bar{Y}$ as instruments along with the spatial lag of rainfall. The estimates of the long-run effects of aid are robust to a second-lag-only specification though. Finally, the Hansen's J statistic for the baseline Sp P-VAR model does not reject the validity of the over-identifying GMM restrictions with large p-values (.376 in Table A1).[22]

4.2 Impulse Response Functions

We begin with a presentation of the results for our preferred baseline specifications of the P-VAR model; we then provide a set of robustness checks at the end of this Section and in the online Appendix.

Figure 4 illustrates the impulse response functions of the model over a 10 year horizon in response to a one standard deviation temporary shock to ODA disbursements for the time-only responses, along with their 95% confidence intervals (in gray) computed by Monte Carlo simulation. On the left-hand side of the Figure, we can see the disbursement shock is large and relatively persistent (the response is significantly positive for about five years). Similarly, reported on the right-hand side, the response of luminosity is on average positive and persistent for a long horizon, but significant only for the first two to three years after the shock. Even though the response of nightlights is smaller and rapidly decays, it generates a non-trivial impact on luminosity, at least in the short-run.

Figure 5 compares the time-only and time-space responses of nightlights to the same shock to ODA as in 4. The effects of the spatial spillovers on the transmission mechanism of the shocks are quite evident. The response is reinforced by the spatial feedback from neighbor districts; it gains an additional 15% effect in the first year, and the contribution of the spatial component compounds over time (for instance, the effect is doubled five years after the shock). As a consequence, the impact of an ODA shock is more persistent and significant for a longer period in the time-space response. The impact shock to log-aid is around .61 and translates into a contemporaneous increase of about .12 units of log-luminosity. The response of log-luminosity after three years is at about .05 and .08 units for the time-only

[22] The Hansen's J statistics include the over-identifying conditions from the spatial components as well as the spatial term of rainfall. The Sp P-VAR model with lag instruments between 2 and 4 passes the over-identification test, but the p-value of the statistic drops below 20%. Relying again on the equation-wise dynamic panel models, we find evidence that the drop is driven by a significantly lower quality of the $Y-$lag instruments once the fourth lag is added to the instrument set; the J statistic is then prevented from falling further down by the other sub-set of exogenous instruments (see column c of Table S3). Another indication of the endogeneity of the fourth lag is obtained from the impulse response analysis where the model with first-lag instrumentation, clearly endogenous, and models including the fourth lag have similar responses which substantially differ from those of the baseline specification. We further investigate these points in Figure S1 of the online Appendix.

and time-space responses respectively. The time-space responses correspond to an elasticity at impact of 20% and a 2−year elasticity of about 16.3%.

Figure 6 documents the time-only responses of the two variables of the model to a one standard deviation temporary shock to nightlights. On the right panel, we see the lights shock has a magnitude and a significance pattern comparable to the ODA shock, but with a slightly larger persistence. The response of ODA, reported on the left, is significantly negative and U-shaped. This response function is quite relevant for the interpretation and corroboration of the identification ordering of the structural shocks.

The response suggests that the effect of luminosity shocks on ODA disbursements takes a few periods to propagate and reaches is highest strength only three years after the shock. In this identification scheme we constrain the response to be zero on impact, but after that the response is free to follow a negative path which is very consistent with the common interpretation of the endogeneity given to the response of aid allocation to economic activity. That is, more aid flows to places with lower income. In the alternative identification ordering, which is explored in the online Appendix, the ODA response is positive on impact for the initial period, and then turns negative (see Figure S17). This response is strongly at odds with that common interpretation of the aid endogeneity issue.

We further explore the underlying mechanism of our identification assumptions in Section S3.2 of the online Appendix, where we conduct a simple exercise to analyze the different effects of early-impact and late-impact ODA shocks. An important caveat in interpreting the results of this exercise is that their reliability is limited by the characteristics of the data we can construct, and the reduced statistically significance of the IRF we find. Although rigorous, the exercise hence provides some suggestive, but not fully conclusive, evidence. The main goal of this analysis is to show that the within period and short-term positive response of nightlights to ODA mainly comes in response to shocks to early-impact aid, while late-impact aid is responsible for the long-term dynamics of the responses. This decomposition of the total effects of ODA shocks would also confirm the empirical validity of the structural identification strategy we adopted in the paper, besides its theoretical justification. The analysis in Figure S21 in the online Appendix definitely support this conclusion.

Figure 7 focuses on the long-run effect of the temporary ODA shock, illustrating the cumulative response of luminosity for both the time-only and time-space cases. The cumulative impact on nightlights is large and statistically quite significant for the first five years after the shock for the time-only response. The response remains (on average) consistently positive for a long horizon, but the large confidence interval makes it not significant after five years. As expected from the previous results, the reinforcing effect of the spatial spillovers increases the magnitude of the time-space cumulative response, but more importantly it enhances the persistence and significance of the path of the response, which remains strongly significant up to the ten years horizon. Numerically, this time-space cumulative response is .73 at ten years. We can compute the cumulative elasticity of nightlight with respect to ODA by comparing the cumulative response to the corresponding cumulative change in ODA, which is 1.57 ten years after the shock. We find a substantial ten-year *cumulative* elasticity of about 46.5% (with an equivalent elasticity for the time-only case of 28%).

Another possible way to assess the long-run effects of ODA on lights is to compute the response of lights to a permanent shift in ODA. This can be easily done, for instance, in the simple case of the time-only dimension using the estimates of the coefficients in A_0 and A_1.

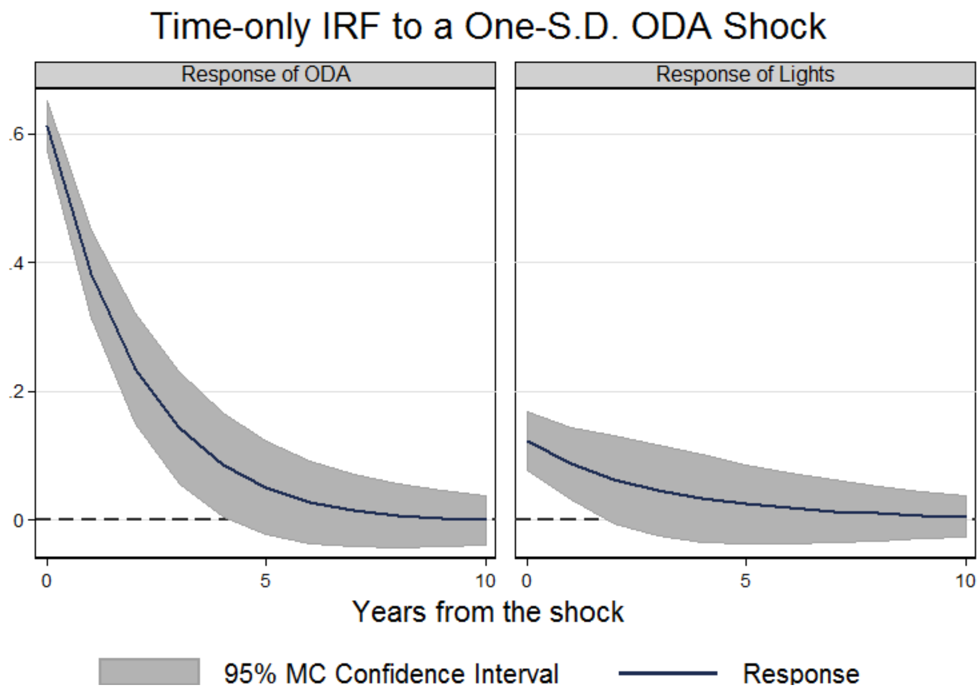
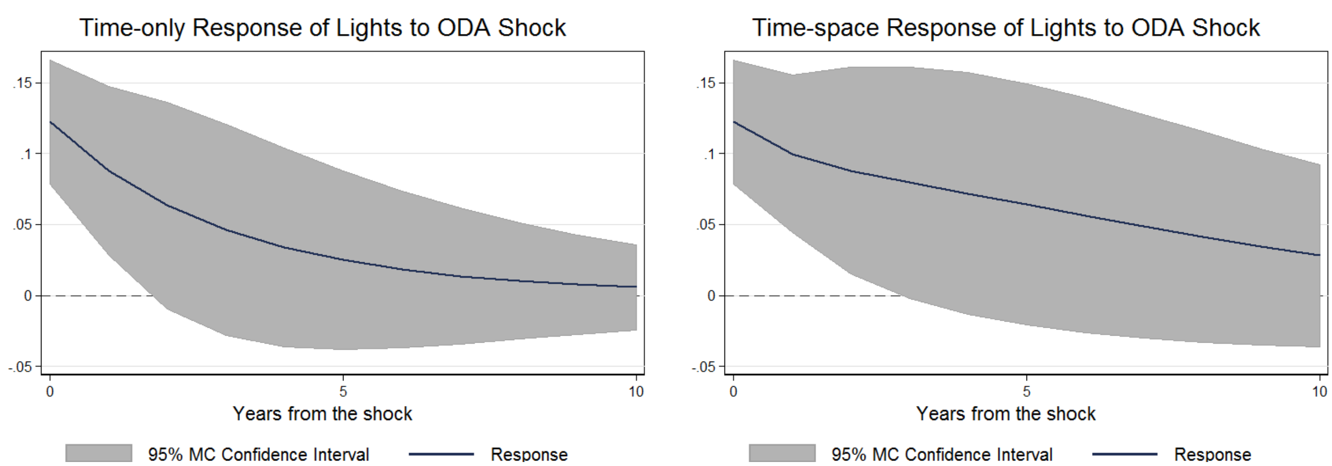


Figure 4: Time-only response functions to a one standard deviation shock to aid disbursements. Identification ordering: light, oda. Years from the shock on the x-axis.



(a) Correct answers.

(b) Time.

Figure 5: Comparison of the the time-only and time-space response functions of lights to a one standard deviation shock to aid disbursements. Identification ordering: light, oda. Years from the shock on the x-axis.

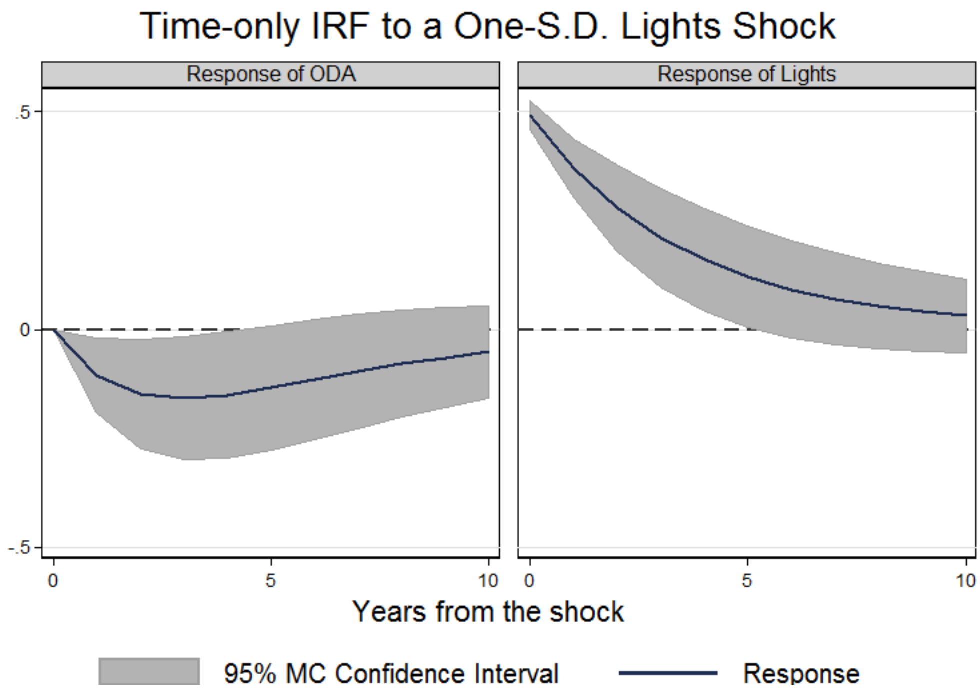
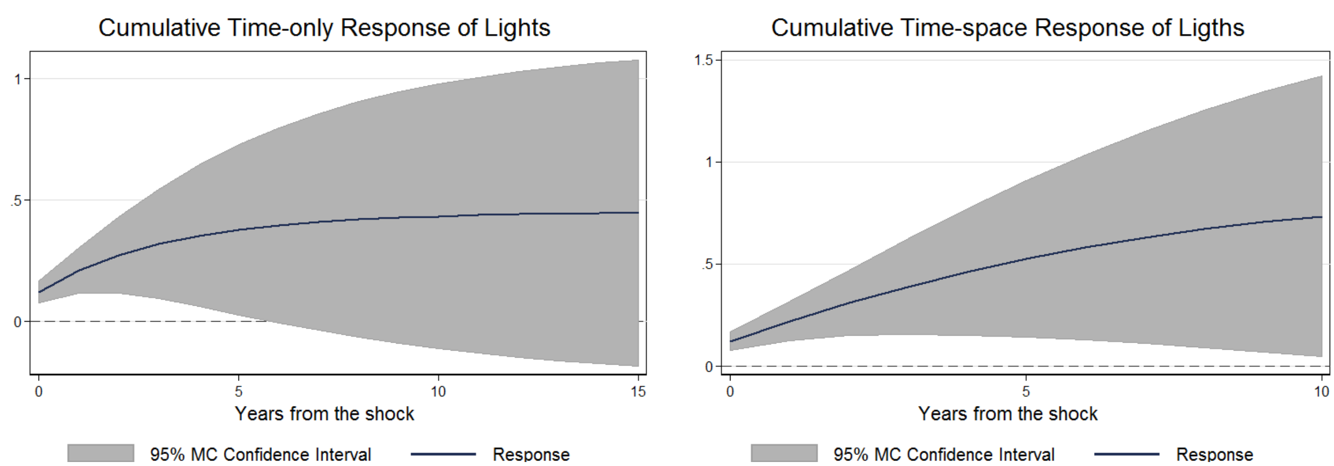


Figure 6: Time-only response functions to a one standard deviation shock to lights. Identification ordering: light, oda. Years from the shock on the x-axis.



(a) Correct answers.

(b) Time.

Figure 7: Comparison of the the time-only and time-space cumulative response of lights to a one standard deviation shock to disbursements. Identification ordering: light, oda. Years from the shock on the x-axis.

	Horizon: years									
	1	2	3	4	5	6	7	8	9	10
ODA on lights	5.8	5.6	5.5	5.4	5.4	5.3	5.3	5.3	5.3	5.3
Lights on ODA	0	2	5.4	8.8	11.7	13.9	15.5	16.5	17.2	17.7

Table 1: Share of the variance of lights and ODA explained by the shocks to the reciprocal variable. The decomposition is computed for the baseline identification ordering of the variables. Values expressed in percentage points.

The nightlights equation in structural form, under our baseline structural identification and omitting some of the terms for convenience of presentation, can be re-written as:

$$light_{i,t} = \alpha_1 light_{i,t-1} + \beta_0 oda_{i,t} + \beta_1 oda_{i,t-1} + ... + v_{2,i,t} \tag{10}$$

where α_1, β_0, and β_1 are combinations of the coefficients in A_0 and A_1, and $v_{2,i,t}$ is the structural shock for the light equation. The long-run elasticity to a permanent shift in ODA is computed as the ratio $\frac{\beta_0+\beta_1}{1-\alpha_1}$, and it is estimated to be 29.2%.[23]

As a last point, we also look at the variance decomposition in only the time dimension of the forecast errors of nighlights and aid disbursements at different time horizons to assess the relative contribution of the ODA shocks to the total variance of lights. Similarly, we can examine the contribution of lights shocks to the variance of ODA. This result is reported in Table 1 for the baseline orthogonalization of the shock structure. The ODA shock explains a relatively small, but stable share of the volatility of lights, ranging between 5 and 6%. Lights shocks, on the contrary, explain a more sizable and increasing share of the ODA volatility, ranging from just 2% at one year to 18% at ten years.

These results provide another interesting comparison between the two alternative structural identification orderings. In the baseline case, the majority of the variance of observed luminosity is determined by its own shocks, but the contribution of aid determines non-negligible fluctuations as well. The same variance decomposition under the alternative ordering shows that aid does not contribute at all to the fluctuations of nightlights, which can be then modeled as an independent autoregressive process.

We conclude this Section with a set of six robustness checks of the effects on the IRF of changing the specification of the model and the way data is treated. Figure 8 illustrates the time-space responses of nightlights to an ODA shock for these cases, with the exception of panel (e) in which the time-only response is considered. These main checks are a starting point to discuss a broader set of robustness exercises, which for sake of brevity are fully described in the online Appendix, Section S3.

The first two checks, in panels (a) and (b), modify the time sample of the estimation and document the stability of the estimation results over time. Panel (a) illustrates the time-space impulse response function for the latest part of the sample, after dropping the first four observations; panel (b) looks at the earlier years, dropping the the last four observations. In panel (a) the mean response is always positive, but the effects are smaller than in the

[23] Point estimates of the P-VAR coefficients and of the Cholesky decomposition matrix are reported in Appendix B.

baseline and not statistically significant. The trajectory of the response is hump-shaped, with a peak after about three years from the shock. In panel (b) the mean response is still positive, but smaller, and with much higher long-run persistence than in the baseline. The confidence intervals are also larger in both cases due to the smaller estimation precision.

These responses show that the results are relatively sensitive to the time sample, which is not surprising given that only 17 years are covered by the full sample. Interestingly, the final portion of the sample seems to contribute relatively more to the short-run dynamics of the response, while the earlier part of the sample contributes more to the medium and long-term persistence of the effects. This difference is consistent with the evidence about the effects of early-impact aid discussed above. Early-impact aid is more concentrated in the last few years of our sample and, as seen in Section S3.2 of the Appendix, better explains the short-run dynamics of the responses of lights. ODA in the late part of the sample contains a relatively larger share of aid in the early-impact category.

The results in these first two panels are closely related to the suggestion of an anonymous referee to use an alternative dataset of geocoded aid projects from the World Bank for a robustness check. Panel (c) illustrates this case. The World Bank dataset provides a more balanced coverage of aid flows over time on the one hand, but it is limited to a single donor on the other hand.[24] It allows us to check, then, whether the baseline results are dominated by the coverage of the AidData dataset. The AidData dataset has improved in precision in the most recent years of the sample, but it could potentially be skewed towards projects with specific characteristics. The panel shows this does not seem to be the case, since we find very similar time-space IRF also with the alternative data.

The fourth check, in panel (d), excludes aid disbursements from projects that are related to energy generation and power supply network enhancement. These projects could increase lights emission without a direct effect on the real economic activity. We notice some interesting differences. The time-space response is smaller on impact than in the baseline case. Since energy related projects are a typical component of early-impact ODA, this drop in the response would be consistent with the short-run effects of this type of aid. However, the confidence interval gets very wide, very rapidly and the statistical significance of the response is largely undermined.

In the fifth check, we explore the robustness of the results to changes in the composition of the cross-section units. Specifically, panel (e) drops the three neighbor districts of the metropolitan region of the capital of Uganda, Kampala. Dropping this set of districts, the time-only response is still positive on impact, but it is characterized by a more rapid decay and smaller long-run effects. A very similar result is obtained by excluding from the sample just the Kampala district. Kampala is the main city of the country, it is the center of a large urban area, and it has a more developed economic and business activity. It is not surprising that aid could be relatively more effective in this type of environment. Interestingly, the contribution of these districts is particularly relevant for the persistence of the light response.

This fifth robustness check is only an example of a very thorough exercise in which we systematically drop, one at the time, each district and small subsets of contiguous districts

[24]The World Bank dataset has some important differences, especially in relation to the application to the Sp P-VAR model, that should be noted. In particular, a much larger share of periods with no ODA disbursements than in the AidData data. We discuss the characteristics of the dataset in Section S3.2 of the online Appendix.

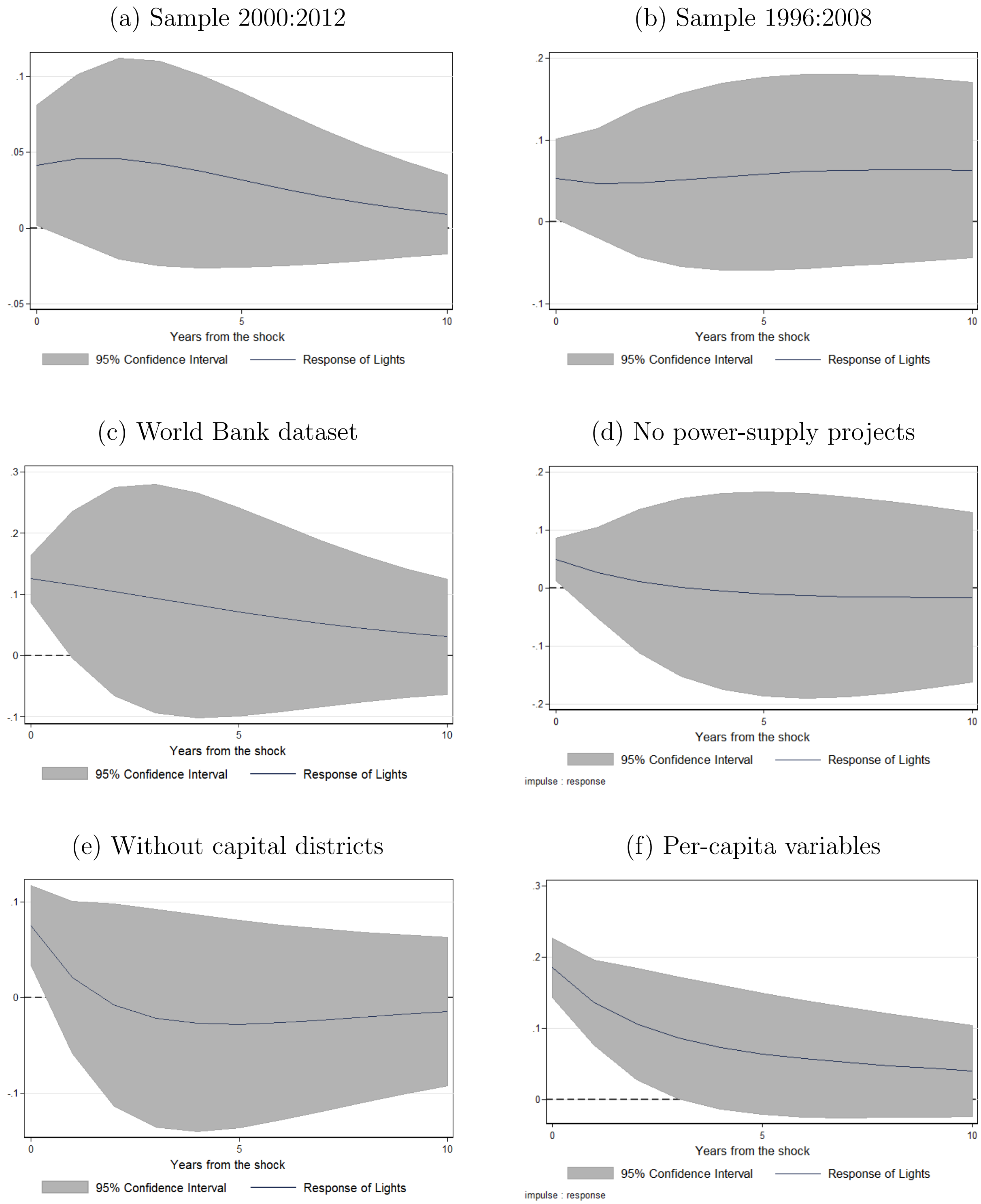


Figure 8: Robustness checks – Responses of nightlights to a one standard deviation shock to ODA. Years from the shock on the x-axis.

to check for the influence of possible outliers. The time-only responses of lights seem to remain very stable across the sets of districts. Since the time-space IRF are reported as the average of responses with spatial spillovers across units, they are not very sensitive to the exclusion of small subsets of districts. For this reason, this robustness check is simply conducted for the time-only IRF. The full set of impulse responses is documented in Section S3 of the Appendix.

Finally, panel (f) considers variables normalized by district population rather than district area. The time-space response remains significantly positive and persistent, and it follows a very similar trajectory. However, the key difference using this data treatment is that the ODA shocks are now larger, reducing the elasticity of the response to 14% on impact. Our preference for data normalized by district area is based on the evidence by Henderson, Storeygard, and Weil (2012), who show that nightlights per area is a strong predictor of economic activity. Lights normalized by district surface are also more tightly connected to household expenditure than lights per capita in our sub-national context, and are less distorted by the upper-bound light truncation. This truncation will limit the sensitivity of light to population growth in dense urban areas. Normalization by area, on the other hand, will capture light growth at the land extensive margin.[25]

4.3 Link to Economic Activity

We now turn to the second stage of the analysis that maps the ODA shocks to a traditional measure of economic activity. As discussed previously, given the type of data available from the household surveys, we believe household expenditure is the best measure of economic activity at the district level in this context. Many alternative measures of economic impact and time horizons are possible so these estimates might be considered baseline exercises. Table 2 below reports the ten-year elasticities of two average household expenditure measures to district luminosity (these are the $\hat{\psi}$, in equation 9).

Column (a) is the elasticity of household weekly consumption expenditure to lights and Column (b) is the elasticity of average household monthly expenditure in non-durable goods to lights. We find highly significant estimates in both cases, with a larger elasticity for the expenditure in non-durable goods (39.2% compared to 20.8% for the weekly consumption expenditure). The magnitude of both the effects is also strongly consistent with the .32 estimated by Henderson, Storeygard, and Weil (2012) for the analogous regression at country level.[26]

Using Table 2 estimates of $\hat{\psi}$ and the baseline Sp P-VAR specification, we can estimate the cumulative effects of an aid shock on household expenditure. The ten-year cumulative elasticity between ODA disbursements and nightlights found in Section 4.2 was 46.5%. This

[25] In Figure S18 of the online Appendix, we also check for the just-identified specification with the second-lag of Y used as an instrument. We find that, even though the short-run response is smaller than in the baseline, the responses are closely comparable in magnitude in the medium and long-run, preserving the long-run effects of aid on nightlight. See Section S3.1 of the online Appendix too.

[26] They utilize country-level GDP rather than district-level household expenditure. A diagnostic assessment of the residuals of these two models is left for the online Appendix, Section S2. Overall, the model assumptions are satisfied, with the exception of heteroskedasticity. We correct the standard errors accordingly in the regression.

	weekly expend.	monthly expend.
	(a)	(b)
log(lights/area)	.208	.392
	(.046)***	(.104)***
F.E.	Y	Y
Obs.	72	72
B/W R^2	.30	.50

Table 2: Estimates of the ten-year elasticity of household expenditure to district luminosity from model (9). Column (a): dependent variable is the log of the average real household weekly consumption expenditure; Column (b): dependent variable is the log of the real average household monthly expenditure in non-durable goods. Standard errors indicated in parenthesis, with significance levels of respectively 1%, 5%, and 10% indicated by *** , ** , and *. Survey years: 1999 and 2009.

indicates that a district-level cumulative increase in aid disbursements of 1% over ten years would generate an increase of about 18.2 basis points in the district average cumulative household expenditure in non-durable goods ten years after the shock. This increase is 9.6 basis points if we measure economic activity through average household consumption expenditure instead. Similarly, a temporary shock to ODA of 1% would cause an average response for the first two years after the shock in the average household non-durable expenditure of 6.4 basis points, and 3.4 points in the average household consumption expenditure.

As an alternative perspective on the economic magnitude of these effects, we can convert them into average dollar per-capita terms. For example, in 2009 for which we have both household expenditure and ODA disbursement data, the average 1% increase in per-capita real dollars ODA disbursements across the Ugandan districts corresponds to 4.5 cents per-capita. We find that a temporary positive shock to aid of this magnitude returns an increase of 7.75 cents in per-capita expenditure in non-durable goods and of 6.35 cents in per-capita consumption expenditure on average for the first two years.[27] This is equivalent to a short-run multiplier around $1.4 - 1.7$. Naturally, the effects of ODA strengthen in the long-run if we consider cumulative estimates. A cumulative increase of per-capita ODA equivalent to 7 cents over ten years produces a cumulative increase in non-durable expenditure of almost 35 cents and an increase in consumption expenditure of 28 cents. The cumulative multiplier is between 4 and 5 then. It is worth noting that these results are strongly in line with those

[27]The 1% increase in ODA in per-capita terms is computed from the corresponding increase in land-unit terms, which is the unit of measure used to compute the compounded elasticities from the empirical exercise. Since districts' areas are constant, the increase in ODA would fully come from a change in disbursements. The increase in ODA is then divided by the 2009 population size of each district to obtain the per-capita value. For the dollar responses of the expenditure measures, we start from the district average households expenditures in the same year. The real average household expenditure in non-durable goods is \$605; since the average household size is around 5 people, the per-capita expenditure in non-durable goods is equal to \$121. The corresponding values for the average annual consumption expenditure are \$935 and \$187 respectively. We finally apply the compounded elasticities to these averages.

reported by Lof, Mekasha, and Tarp (2015), who find an average multiplier effects around 4.5 for a panel of 59 countries.

5 Conclusion

The low-income country recipients of US and other OECD donor of foreign aid contain over 4 billion people, the majority of the global population. Yet the large literature attempting to measure aid impact at the country level using traditional data sources and estimation techniques has produced no consensus on the effects of aid on growth. This research shifts the analysis of aid impact from the country to the sub-national level, combines traditional data sources with remote sensing (satellite) data, and employs an estimation technique that accounts for the endogenous allocation of aid across the sub-national (district) units.

Our regional estimation strategy entails two-stages. In stage one we use a spatial panel vector-autoregressive model to generate the impulse response of luminosity to aid shocks. We find a positive, statistically significant, and persistent response to the shock. The second stage uses a traditional regression approach to generate coefficients estimates to map the impulse response to traditional local economic variables. Connecting the two stages indicates that the shock impact on expenditure is small, but non-negligible. We find the multiplier effect of ODA on household expenditure ranges between 4 and 5 in the long-run horizon.

Our approach is highly scalable across location, sector, and outcomes, and it holds promise as a flexible tool for policy analysis. The most immediate opportunity is application of our methodology to the other countries with geo-coded AidData (over fifteen countries at this writing). Examples of scalability beyond location include using this approach to measure local effects of alternative "treatments" (to official foreign aid) that can be tracked over time and for which the autoregressive methodology would be suitable. For example, a straightforward application would be to measure the local impacts of aid disbursements from large private foundations and NGO. Examples of scalability beyond economic growth could include the impact of a treatment upon investment decisions, health conditions, governance, and the environment. Additionally, the impact of specific categories of aid that we would not expect to have strong light-generating consequences could also be captured by satellite signals other than nightlights. For instance, agricultural land-use-change associated with irrigation or farmers education projects could be measured using infrared and near-infrared satellite data. This approach is the subject of ongoing research.

References

Addison, T., and F. Tarp (2015): "Aid Policy and the Macroeconomic Management of Aid," *World Development*, 69, 1–5.

Alesina, A., and D. Dollar (2000): "Who gives foreign aid to whom and why?," *Journal of Economic Growth*, 5(1), 33–63.

Arellano, M., and S. Bond (1991a): "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," *Review of Economic Studies*, 58(2), 277–297.

——— (1991b): "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies*, 58(2), 277–297.

Badinger, H., W. Muller, and G. Tondl (2004): "Regional Convergence in the European Union, 1985- 1999: A Spatial Dynamic Panel Analysis," *Regional Studies*, 38(3), 241–253.

Bai, J., and S. Ng (2004): "A PANIC Attack on Unit Roots and Cointegration," *Econometrica*, 72(4), 1127–1177.

Baltagi, B. H., G. Bresson, and A. Pirotte (2005): "Panel unit root tests and spatial dependence," *Journal of Applied Econometrics*, 22(2), 339–360.

Beenstock, M., and D. Felsenstein (2007): "Spatial Vector Autoregressions," *Spatial Economic Analysis*, 2(2), 167196.

Berthélemy, J.-C. (2006): "Bilateral Donors Interest vs. Recipients Development Motives in Aid Allocation: Do All Donors Behave the Same?," *Review of Development Economics*, 10(02), 179–194.

Blundell, R., and S. Bond (1998): "Initial conditions and moment restrictions in dynamic panel data models," *Journal of Econometrics*, 87(1), 115–143.

Bouayad-Agha, S., N. Tutpin, and L. Védrine (2013): "Fostering the Development of European Regions: A Spatial Dynamic Panel Data Analysis of the Impact of Cohesion Policy," *Regional Studies*, 47(9), 15731593.

Brady, R. R. (2011): "Measuring the diffusion of housing prices across space and over time," *Journal of Applied Econometrics*, 26(2), 213–231.

Briggs, R. C. (2012): "Electrifying the base? Aid and incumbent advantage in Ghana," *The Journal of Modern African Studies*, 50(4), 603–624.

Center for International Earth Science Information Network, Columbia University (CIESIN-CIAT) (2005): "Gridded Population of the World Version 3: Population Grids (GPWv3)," Palisades, NY: Socioeconomic Data and Applications Center (SEDAC), Columbia University, Downloaded on 4-11-2015.

Center for International Earth Science Information Network, Columbia University; United Nations Food and Agriculture Programme; and Centro Internacional de Agricultura Tropical (CIESIN-FAO-CIAT) (2005): "Gridded Population of the World: Future Estimates (GPWFE)," Palisades, NY: Socioeconomic Data and Applications Center (SEDAC), Columbia University, Downloaded on 4-11-2015.

Chen, X., and W. D. Nordhaus (2011): "Luminosity Data as a Proxy for Economic Statistics," *Procedings of the National Academy of Sciences*, 108(21), 8589–94.

Civelli, A., A. W. Horowitz, and A. Teixeira (2016): "Is Foreign Aid Motivated by Altruism or Self Interest? A Theoretical Model and Empirical Test," *BE Journal of Economic Policy and Analaysis - Forthcoming*.

Clemens, M. A., S. Radelet, R. R. Bhavnani, and S. Bazzi (2012): "Counting Chickens when they Hatch: Timing and the Effects of Aid on Growth," *The Economic Journal*, 122(June), 590–617.

Crost, B., J. Felter, and P. Johnston (2014): "Aid under Fire: Development Projects and Civil Conflict," *American Economic Review*, 104(6), 1833–56.

De, R., and C. Becker (2015): "The Foreign Aid Effectiveness Debate: Evidence from Malawi," AidData Working Paper 15/6.

Deaton, A. (2010): "Instruments, randomization, and learning about development," *Journal of Economic Literature*, 48(June), 424–455.

Di Giacinto, V. (2010): "On vector autoregressive modeling in space and time," *Journal of Geographical Systems*, 12(2), 125–154.

Dionne, K. Y., E. Kramon, and T. Roberts (2013): "Aid Effectiveness and Allocation: Evidence from Malawi," Paper Presentation, Conference 'Research Frontiers in Foreign Aid' at Princeton University.

Dreher, A., V. Z. Eichenauer, and K. Gehring (2014): "Geopolitics, Aid and Growth," Cepr discussion paper no. dp9904.

Dreher, A., and S. Lohmann (2015): "Aid and Growth at the Regional Level," IMF Working Papers 15/196, International Monetary Fund.

Duflo, E., R. Glennerster, and M. Kremer (2008): "Using randomization in development economics research: A toolkit," in *Handbook of development economics 4*, ed. by T. P. Schultz, and J. Strauss, pp. 3895–3962. North-Holland, Amsterdan.

Elvidge, C. D., F.-C. Hsu, K. E. Baugh, and T. Ghosh (2014): "National Trends in Satellite-Observed Lighting: 1992-2012," in *Global Urban Monitoring and Assessment Through Earth Observation*, ed. by Q. Weng. CRC Press, London.

Findley, M. G., J. Powell, D. Strandow, and J. Tanner (2011): "The Localized Geography of Foreign Aid: A New Dataset and Application to Violent Armed Conflict," *World Development*, 39(11), 1995–2009.

Funk, C., P. Peterson, M. Landsfeld, D. Pedreros, J. Verdin, S. Shukla, G. Husak, J. Rowland, L. Harrison, A. Hoell, and J. Michaelsen (2015): "The climate hazards infrared precipitation with stationsa new environmental record for monitoring extremes," *Scientific Data*, 2(66), 1–21.

Galiani, S., S. Knack, L. C. Xu, and B. Zou (2017): "The effect of aid on growth: evidence from a Quasi-experiment," *Journal of Economic Growth*, 22(1), 1–33.

Gillanders, R. (2016): "The Effects of Foreign Aid in Sub-Saharan Africa," *The Economic and Social Review*, 47(3), 339–360.

Griffith, D. A. (2003): *Spatial autocorrelation and spatial filtering: gaining understanding through theory and scientific visualization*. Springer-Verlag,, Berlin, Germany.

Griffith, D. A., and P. R. Peres-Neto (2006): "spatial modeling in ecology: the flexibility of eigenfunction spatial analyses," *Ecology*, 87(10), 2603–2613.

Hamilton, S. E., and C. Stankwitz (2012): "Examining the relationship between international aid and mangrove deforestation in coastal Ecuador from 1970 to 2006," *Journal of Land Use Science*, 7(2), 177–202.

Henderson, J., A. Storeygard, and D. Weil (2011): "A bright idea for measuring economic growth," *American Economic Review*, 101(3), 194–199.

Henderson, J. V., A. Storeygard, and D. N. Weil (2012): "Measuring Economic Growth from Outer Space," *American Economic Review*, 102(2), 994–1028.

Im, K. S., M. H. Pesaran, and Y. Shin (2003): "Testing for unit roots in heterogeneous panels," *Journal of Econometrics*, 115(1), 53–74.

Jablonski, R. (2014): "How Aid Targets Votes: The Effect of Electoral Strategies on the Distribution of Foreign Aid," *World Politics*, 66(2), 293–330.

Juselius, K., N. F. Moller, and F. Tarp (2014): "The Long-Run Impact of Foreign Aid in 36 African Countries: Insights from Multivariate Time Series Analysis," *Oxford Bulletin of Economics and Statistics*, 76(2), 153–184.

Kilby, C. (2013): "The political economy of project preparation: An empirical analysis of World Bank projects," *Journal of Development Economics*, 105(C), 211–225.

Kilby, C., and A. Dreher (2010): "The impact of aid on growth revisited: Do donor motives matter?," *Economics Letters*, 107(03), 338–340.

Lof, M., T. J. Mekasha, and F. Tarp (2015): "Aid and Income: Another Time-series Perspective," *World Development*, 69(C), 19–30.

Love, I., and L. Zicchino (2006): "Financial development and dynamic investment behavior: Evidence from panel VAR," *The Quarterly Review of Economics and Finance*, 46(2), 190–210.

Márquez, M. A., J. Ramajo, and G. J. Hewings (2015): "Regional growth and spatial spillovers: Evidence from an SpVAR for the Spanish regions," *Papers in Regional Science*, 94(1), 1–18.

Mosley, P. (1987): *Foreign Aid, its Defense and Reform*. University Press of Kentucky, Lexington, KY.

Pesaran, M. H. (2004): "General Diagnostic Tests for Cross Section Dependence in Panels," CESifo Working Paper Series 1229, CESifo Group Munich.

Pesaran, M. H., and R. Smith (1995): "Estimating long-run relationships from dynamic heterogeneous panels," *Journal of Econometrics*, 68(1), 79–113.

Pesaresi, M., G. Huadong, X. Blaes, D. Ehrlich, S. Ferri, L. Gueguen, M. Halkia, M. K. T. Kemper, L. Lu, M. A. Marin-Herrera, G. K. Ouzounis, M. Scavazzon, and P. S. V. S. L. Zanchetta (2013): "A Global Human Settlement Layer From Optical HR/VHR RS Data: Concept and First Results," *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, 6(5), 2101–2131.

Radelet, S. (2006): "A primer on foreign aid," Working paper, Center for Global Development.

Rajan, R., and A. Subramanian (2008): "Aid and Growth: What does the Cross-Country Evidence Really Show?," *Review of Economics and Statistics*, 90(4), 643–665.

Reese, S., and J. Westerlund (2016): "Panicca: Panic on Cross-Section Averages," *Journal of Applied Econometrics*, 31(6), 961–981.

Roodman, D. (2009): "How to do xtabond2: An introduction to difference and system GMM in Stata," *Stata Journal*, 9(1), 86–136.

Sarafidis, V., T. Yamagata, and D. Robertson (2009): "A test of cross section dependence for a linear dynamic panel model with regressors," *Journal of Econometrics*, 148(2), 149–161.

Sims, C. A. (1980): "Macroeconomics and Realty," *Econometrica*, 48(1), 1–48.

Temple, J. R. (2010): "Aid and conditionality," in *Handbook of development economics*, ed. by D. Rodrik, and M. Rosenzweig, chap. 67, pp. 4415–4523. Elsevier.

Tierney, M. J. (2011): "More dollars than sense: refining our knowledge of development finance using AidData," *World Development*, 39(11), 1891–1906.

Appendix

A Dataset Construction Details

This Appendix provides more details about the data collection and treatments following the outline of Section 3.

Panel Structure – The cross-section of the panel in the study includes 36 regions. Of these regions, 35 administrative districts (ADM2 level units), while the last one is a synthetic region obtained from the aggregation of all the other Ugandan districts. The geographic boundaries of the districts are obtained from the world administrative divisions layer provided by ESRI-ArcGIS. We adopt the most recent definition of districts established in 2010, also shared by the AidData dataset, which consists of 112 districts. The synthetic district aggregates 77 districts with low (or no) night luminescence corresponding to about 70% of the Ugandan territory, and 48% of the population on average over this sample.

Although the introduction of the synthetic district may seem an asymmetric treatment of the data at first sight, our choice is justified by two goals. On one hand, it satisfies the need of a sufficiently balanced structure of the panel to support the P-VAR estimation. Both nightlights and ODA disbursements are regularly observed for the main 35 districts, but not for the others. Taken individually, these districts have extremely sporadic luminosity and ODA series that make them unsuitable for the VAR model. On the other hand, the aggregation corrects for these problems and allows us to preserve the information coming from this set of districts without arbitrarily dropping any observation.

Overall, the synthetic district receives a less than proportional share of total aid, on average only around 11%, and produces a very small fraction of the country nightlights, never bigger than 4%. The reason for this weaker luminosity emission can be found in the smaller urbanization share in the synthetic region. As illustrated by Figure 3, there is a close correspondence between urban areas and luminosity. The districts in the synthetic region are primarily rural, with an urban share about 50 times smaller than that in the lit districts.[28] We maintain the comparability across units by standardizing the observations by the surface area of the respective district (or population in some robustness checks). Under the assumption of homogeneity of the effects of ODA across geographic locations implicitly imposed by the P-VAR approach, the use of the synthetic district weakens the benefit of the ODA geographic disaggregation strategy for this unit.

ODA Data – The ODA data frequently shows multiple disbursements made under the same project id in multiple locations, not necessarily in the same district or in the same year. In such cases, there is not sufficient information to assign all disbursements of a project to a single district and the records show equal aid disbursement for each of the multiple locations of a project. In these cases, we proportionally re-distribute the disbursements over the recipient districts of a project based on population size. Also, if a district does not receive any aid in one period, we substitute the observation with a small value, .0001, before taking the log. This occurs in about 20% of the observations.

[28]The share of urbanized land in 2014 was .8% in the lit districts and .014% in the no-light districts. These shares are computed using the dataset provided by Pesaresi, Huadong, Blaes, Ehrlich, Ferri, Gueguen, Halkia, Kemper, Lu, Marin-Herrera, Ouzounis, Scavazzon, and Zanchetta (2013) - Beta release.

Early and Late ODA – Clemens, Radelet, Bhavnani, and Bazzi (2012) classify aid projects in two categories, "early-impact" and "late-impact" aid, depending on the time horizon at which they might plausibly affect economic growth. In their words: *"Early-impact aid [...] might plausibly affect growth within a few years after it is disbursed, such as aid for road construction. [...] Late-impact aid is that which finances activities that are likely to take many years or even decades to affect growth, such as vaccination campaigns or basic education."* Their classification is based on the 5-digit purpose codes of the projects, which are not provided by AidData for its geo-coded disbursements; AidData provides only 3-digit codes, which correspond to a coarser classification. We then map the 5-digit classification into the 3-digit purpose codes by assigning codes to the short-term impact group whenever the majority of the 5-digit codes stemming from a 3-digit code are classified as "early-impact" by Clemens, Radelet, Bhavnani, and Bazzi (2012).

Their original classification is available from the technical note accompanying their paper. This translates into the following list of purpose codes for the projects found in the precision levels 1-3 of the AidData dataset: 160 - Other social infrastructures and services , 210 - Transport and storage, 220 - Communications, 230 - Energy generation and supply, 311 - Agriculture, 510 - General Budget Support. This additional filtering has a significant impact on the availability of of projects in two ways. First, the number of projects drops by two thirds with respect to the dataset in use in the main analysis. Second, the distribution of early-impact aid is not uniform across the 36 districts we consider anymore. Five districts do not receive any short-term impact aid, for example, while early-impact disbursements are concentrated in the very last years of the sample for an other handful of them. As a consequence, we can only define a smaller panel of 22 districts suitable for the analysis of the effects of early-impact aid. Even though these differences clearly limit the comparability with the main analysis, this exercise gives us the opportunity to obtain, at least, suggestive results to support the identification strategy of the structural shocks used in the main VAR model of the paper. Further details about this panel data are left for the online Appendix, along with a deeper discussion of these results.

World Bank Aid Data – We also run robustness checks using another aid project data base, reporting disbursements exclusively from the World Bank. This data base records all the aid projects from 1995 to 2014 financed by two World Bank lines: International Bank for Reconstruction and Development (IBRD) and International Development Agency (IDA). As in the aid data base used in this paper, each project can be identified by location with precision codes going from 1 to 8, as well as by sector and disbursement date. The main difference between the two datasets is that the World Bank database is smaller because it refers to one donor only; however, it has some advantages with respect to geographic precision, geographic coverage within the donor, homogeneity of classification over time. For example, the World Bank data bases has 73% of the projects classified with high precision codes 1, 2 or 3 (projects are located within 25 km around latitude/longitude given by the project or ADM2 centroid location) whereas other AidData aid datasets have, on average, just 60%. This is an interesting trade-off to explore, and we do that in Section 4.2 of the paper and in the online Appendix.

Nightlight Data – In order to harmonize luminosity data from different satellites, we apply the inter-calibration adjustment parameters provided by Elvidge, Hsu, Baugh, and Ghosh (2014). As is done for ODA, the luminosity variable in the Sp P-VAR is then constructed as the log of the sum of the luminosity index for all the pixels within a district boundaries standardized by the district area (measured in squared kilometers). In a very limited number of cases (7 periods in 4 districts) there is no luminosity in one year, even though these periods are preceded and followed by very clear nightlight signals. Since the DMSP satellites are meant to record the luminosity emission from stable human-based sources of light, we believe it is more plausible to interpret the lack of observations in these years as missing observations rather than actual zero values. Therefore, we substitute these observations with an *spline* piecewise polynomial interpolation. The choice of using these substitutions bears a small impact on the final results. We find nearly identical effects when we add a small value to the zero observations instead of applying the interpolation.

Household Surveys – The UNHS is designed to be representative at national and macro-region level; however, geographic coordinates are available for each surveyed household. We use this information to assign the households to the correct districts based on the 2010 administrative definition of districts. We then compute the district expenditure as the weighted average of the expenditures reported by the households within a district, using the survey multipliers to construct district rescaled weighting schemes. The surveys basically adopt a stratified two-stage sampling design, in which Enumeration Areas (EA) are first sampled with probability proportional to their population relative to the national aggregate, and then households are randomly sampled within each Enumeration Area. On average, we had at least 60-80 observations per district, although a few of the 35 districts are not sampled in either year of the surveys. Since the EA are sub-units of the districts, the sampling procedure does not necessarily cover every district, neither can it guarantee the full coverage of a district included in the survey. For each district in our panel, we identify all the EA belonging to that district and use the (national) multipliers of the EA to construct the relative within-district weights for the households in each EA. Since the multipliers designed to reflect the representativeness of the EA at national level, this approach is only an approximation, in particular when a small portion of the the EA of a district are sampled. However, the approach is quite satisfactory for the synthetic region and it is arguably more effective than simply treating the households sampled within a district as $i.i.d.$ observations, for example. As a robustness check, we estimate (9) also under this second scenario finding still very significant, but 30% smaller, estimates of ψ.

Rainfall – This variable was constructed using version 2.0 of the Climate Hazards Group InfraRed Precipitation (CHIRPS) dataset for global terrestrial precipitations (Funk, Peterson, Landsfeld, Pedreros, Verdin, Shukla, Husak, Rowland, Harrison, Hoell, and Michaelsen, 2015). The data is downloaded from the Climate Hazards Group website.[29] The dataset compiles information from different sources, in particular blending data directly obtained from in-locu stations and from interpolated gauge satellite datasets. The resolution of the dataset is relatively coarser than the rest of the dataset we utilize due to the specific difficulties of data collection for precipitation that rely on terrestrial weather stations. However, this is a relatively minor issue when considering geographic areas as large as districts and aggregation over annual periods, which are less sensitive to temporary variations of the geographic distribution of rainfall.

Population Series – The geographic distribution of population at relatively high resolu-

[29]At the web address http://chg.geog.ucsb.edu/index.html

tion (about 100 squared meters) is obtained from the Socioeconomic Data and Application Center (SEDAC) at $5-$year frequency for the period $1995 - 2005$. These are raster images of world population, harmonized with national and sub-national administrative population counts and United Nation country statistics, which can be aggregated at the desired geographic unit consistently over time. The raster images are part of the two Gridded Population of the World series provided by SEDAC, the GPWv3 and the GPWFE (CIESIN-CIAT, 2005; CIESIN-FAO-CIAT, 2005). The low frequency of observations requires a further manipulation to construct annual time series: we interpolate the 5-year data with an *spline* piecewise polynomial. Though the interpolation returns annual observations, this remains a quite noisy measure of population change, with only a mechanically predicted variability over time. However, since population dynamics are relatively slow and predictable, we can utilize it with some confidence to standardize the other variables of the model, but not directly as a control variable.

B Point Estimates from Baseline P-VAR

	$oda_{i,t}$ (a)	$light_{i,t}$ (b)
$oda_{i,t-1}$	.666	-.007
	(.051)***	(.035)
$light_{i,t-1}$	-.212	.753
	(.085)**	(.070)***
$rain_{i,t}$	1.33	.816
	(.653)**	(.498)
$\overline{light}_{i,t}$	-.155	-.302
	(.236)	(.117)***
$\overline{oda}_{i,t}$	-.243	-.239
	(.187)	(.109)**
F.E.	Y	
Obs.	468	
N. of Panels	36	
Instruments	$L2/3$, $\overline{rain}$	
Hansen J-test	$\chi^2(10) = 10.76$ [.376]	

Table A1: Estimates of the fixed-effects P-VAR(1) model in equation (1). Cross-section units are the Ugandan districts (including the synthetic district); t is expressed in years. The endogenous variables are the logs of the ratio of nightlight to the district surface area, $light_{i,t}$, and the ratio of aid disbursements to the district surface area, $oda_{i,t}$. Spatial terms $\overline{oda}_{i,t}$ and $\overline{light}_{i,t}$ are constructed based on the share-border contiguity matrix. Exogenous variable is $rain_{i,t}$, the log of total annual rainfall per district. Column (a): ODA equation; Column (b): nightlight equation. Instruments are the time-lagged endogenous variables and spatial terms, and the spatial-lag of the exogenous variable. Standard errors indicated in parenthesis, with significance levels of respectively 1%, 5%, and 10% indicated by *** , ** , and *. P-values in brackets for the Hansen's test. Sample years: 1996 and 2012.

	light	*oda*
light	.492	0
oda	.122	.611

Table A2: Estimates of the Cholesky factorization matrix. Variables order: light, oda. Column variable responds to row variable.