ISSN 2385-2755
DiSSE Working papers
[online]

SAPIENZA
UNIVERSITÀ DI ROMA

WORKING PAPERS SERIES
DIPARTIMENTO DI
SCIENZE SOCIALI ED ECONOMICHE

GIS Methods for the Impact Evaluation of a Project with Missing Data

Valerio Leone Sciabolazza Alban Lika

N. 3/2024

SAPIENZA - UNIVERSITY OF ROME
P.le Aldo Moro n.5 – 00185 Roma T(+39) 0649910563
CF80209930587 – P.IVA 02133771002

GIS Methods for the Impact Evaluation of a Project with Missing Data

Valerio Leone Sciabolazza (Sapienza Univ. Rome)

Alban Lika (Sapienza Univ. Rome)

Abstract

Aim of this paper is to show how GIS methods and secondary data can be used as a complementary set of tools to conduct the impact assessment of an intervention. This is done by using satellite data to observe changes in the economic activities of the areas where beneficiaries of the intervention reside, and in the surrounding areas. Then, using a synthetic control approach based on GIS methods, the performance of treated and control areas is compared to assess the impact of the intervention.

Keywords: impact assessment; missing data; GIS; synthetic control.

JEL: B41; C54; O47

Acknowledgement: this work received financial support by the Italian Ministry of Education, Universities and Research under award number 2017RJJNM5_001 (Information, firms' performance and workers' careers, PRIN 2017). The findings, interpretations, and conclusions expressed in this document are solely those of the authors. They do not represent the views of the International Fund for Agricultural Development, its staff, its affiliated organizations, the IOE, its member countries, its Executive Board, or the governments they represent.

1. Introduction

The impact assessment of a project is crucial for international organizations to justify the funding entrusted to them by its Member States, as they help ensure accountability and transparency and contribute to the global public good by generating knowledge, data, and evidence to improve transformation policies and programs globally.

Implementing a rigorous impact assessment is not always feasible. First, this exercise entails considerable financial resources to collect data, and depending on funding and local capacity for data collection and management, the quality of data for an impact assessment may largely vary. Second, sometimes data are available, but only to conduct ex-post quasi-experimental analysis, which reduces the credibility of the evaluation. Third, even when good and rich primary data at the household or farmer level is available, these are often not consistent across projects, making it impossible to compare results across different areas. As a result, despite the growing number of impact evaluations, international projects still largely rely on a limited number of outcome indicators from the theory of change that are part of the Monitoring and Evaluation System or the Results and Measurement Framework of international organizations.

With the COVID-19 pandemic, international organizations began to explore alternative approaches to collecting data in person.[1] The Internation Fund for Agricultural Development (IFAD) for instance published a set of tools relying on Geographic Information System (GIS) methods and secondary data for enhancing the impact assessment operations (Mabiso et al. 2022), which were later refined and further extended to analyze the universe of the last wave of IFAD's project (IFAD 11).

In this paper, we present a more sophisticated and detailed use of these GIS methods to understand the extent to which GIS methods and secondary data can be used as a complementary set of tools to conduct project impact assessment, reduce the cost of data collection and overcome data quality challenges due to local capacity. This new

[1] https://blogs.worldbank.org/education/remote-data-collection-during-covid-19-thing-past-or-way-future

method consists in using only few GIS indicators on beneficiaries and avoid the additional costs and challenges to build a reliable counterfactual, a task that is not easy to conduct and limit the possibility to have a sound comparison group.

GIS data is obtained from multiple open-source databases containing e.g., satellite data, remotely sensed data, digital terrain models. It is often employed to map economic data with a spatial component and complement official statistics to generate additional spatial data as inputs to statistical analyses. It can thus be used to obtain socio-economic indicators, living standards measures, land resources and environmental data, and vegetation indexes, among other indicators, in order to estimate economic growth, the spread of economic activities, the quality of political institutions, the access to specific areas, the geographical distribution of agricultural practices, the development of infrastructure networks, environmental policies, and conflicts (Chen et al., 2013; Henderson et al., 2012; Hodler and Raschky, 2014; Donaldson and Hornbeck, 2016; Michalopoulos and Papayoannu, 2016; Rogall, 2021; Prem et al, 2023).

In addition to this, a large array of research used GIS data in combination with primary data to design credible counterfactual data to rigorously assess the impact of interventions, see among others: Banerjee et al. (2020), Dinkelman (2011), Duflo and Pande (2007), Faber (2014), Michaels (2008), Michalopoulos and Papayoannu (2013), Nunn (2008), and Qian (2008). This paper builds on this latter strand of research.

We propose to evaluate the effect of an intervention by comparing the performance of a treated area with that of a neighboring area, which is used as a control area. By being located close to one another, the two areas are likely to be comparable, in that they will have similar observed and unobserved characteristics and will be exposed to the same shocks.

A problem with comparing treated and control areas in our setting is that there might be unobserved differences between the treated area and its surrounding which may hinder our analysis. For this reason, we adopt a synthetic control approach (see e.g., Abadie and Gardeazabal, 2003, Abadie 2010, Abadie, 2021) and recombine the socio-economic characteristics of the surrounding of the treated place to obtain a plausible synthetic control group, considering both time-varying and time-invariant differences between treatment and control. In essence, this approach is therefore similar to the difference-in-differences analysis by Druckenmiller and Hsian (2018), where the unobserved component is captured by the difference in space.

Our method involves three steps. First, the control area is split into four quadrants of the same geographical extension, and satellite data is used to register relevant characteristics of these quadrants and the treated area. Second, we recombine the metrics associated with each quadrant to construct a so-called synthetic control: i.e., a control area with characteristics most similar to the treated area before the intervention took place. Intuitively, this is done by taking a weighted average of the metrics associated with the four quadrants, where weights are chosen so that the characteristics of the synthetic area best resemble the characteristics of the treated area before the intervention: i.e., differences between the metrics registered in the treatment and the control areas are minimized. Differences in the outcome of interest between the treatment area and the synthetic control after treatment took place are then used to assess the impact of the intervention. Third, we compare the outcome of interest for treatment areas and their relative synthetic control area that are immediately adjacent, thus assuming that adjacent observational units are comparable to one another but not comparable to distant units. By restricting comparisons to adjacent neighbors in our procedure, the influence of all omitted variables that are common to neighboring units are differenced out.

In order to showcase our method, we present a mock exercise. We consider a project to promote employability and entrepreneurship conditions and support the creation of SME’s and start-ups in the rural areas of Colombia, we identify the centroid of the provinces in which the project took place, and we assume the area surrounding these points to be the treated area. Then, we construct a synthetic control area to evaluate the effect of the intervention. This is done by collecting satellite data to control for differences in treatment and control areas which may explain differences in their outcome of interest after the intervention took place. Results from this exercise will be hardly reliable: since we do not have the exact location of treated households, a number of spurious correlations can emerge from our analysis. Nevertheless, the exercise has the merit to walk the reader throughout all the necessary steps to follow our procedure.

While our methodology can be generalized to all cases in which GIS data on treated individuals is available for conducting an impact assessment exercise, it is important to keep in mind that the construction of a synthetic control area requires deep knowledge of the intervention to be evaluated. In fact, it is crucial to understand which differences between treated and control areas need to be accounted for in order to difference out from the estimates the role of omitted variables common to the neighboring units which can hinder the correct identification of the treatment effect.

The remainder of this paper is organized as follows: Section 2 lays out the main features of our methodology. Section 3 uses a case study to show what information is important to collect in order to construct a synthetic control. Section 4 shows how to link theory to data in order to conduct a study using our methodology. Section 5 shortly illustrate the empirical results obtainable with our methodology. Section 6 draws conclusions.

2. Methodology

The core insight of our approach is that differences in outcomes across areas is explained by heterogeneity in their characteristics, but this heterogeneity can be eliminated when the spatial position of the treated population can be located. In fact, when the location of the treated population is known, it is possible to compare it with the population of a neighboring area, used as a control group. By being located close to one another, the two population are likely to be comparable in that they will have similar observed and unobserved characteristics and will be exposed to the same shocks.

The primary information needed to apply our methodology is therefore the geo-localization of the area where treatment takes place. The identification of the treated area requires special attention because of the potential presence of spatial spillovers, that is when the treatment effect for the treated units could affect the trajectories in nearby areas. For this reason, the researcher needs to determine what could cause treatment spillover effects. For instance, spillover effects can depend on whether there are complementarities between the treated areas and the nearby areas: e.g., an increase in economic activities in the treated regions could bring more economic activity to the entire region thus benefiting nearby areas. In this case, estimates would be underestimated. However, also the contrary could also occur: e.g., there is a reallocation from nearby areas to the treated ones which could lead to an overestimate of the parameter of interest.

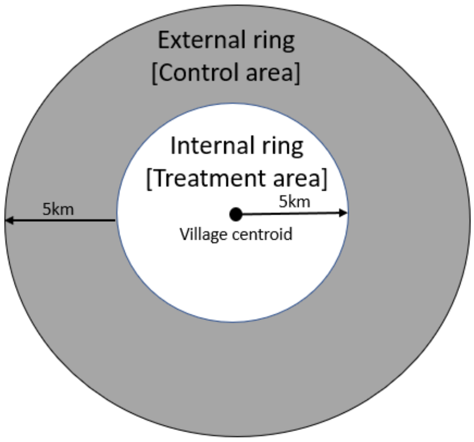
Appropriate methodologies to deal with the problem of policy impacts that do not stop at the border of the discontinuity (i.e., at the border of the treated area) are still being elaborated in the literature (see, among others, Jardim et al. 2022; and Butts, 2023) and future research may benefit from the elaboration of these methods to produce more accurate results. In what follows, for the sake of simplicity, we assume that this area is large 5km. This area is the main unit of observation and we refer to it alternatively as *treated area* or *internal ring*.

Once the internal ring has been identified, it is possible to obtain a control area by drawing a circle (5km radius) around each internal ring, to which we refer as to the *external* ring. The result of this exercise for a generic area is reported in Figure 1.

A problem with comparing the internal and the external ring is that there might be differences between them even if these are located close to one another and they are likely very similar. For this reason, caution is required when comparing them. For this reason, we assume that comparability between treated and control areas is possible only conditional on some observable characteristics, so that differences between the outcome of treated and control areas are attributable only to the treatment. From an econometric standpoint, this implies that for our methodology to work, a conditional independence assumption must credibly hold in the data.

Figure 1

Construction of treatment and control areas

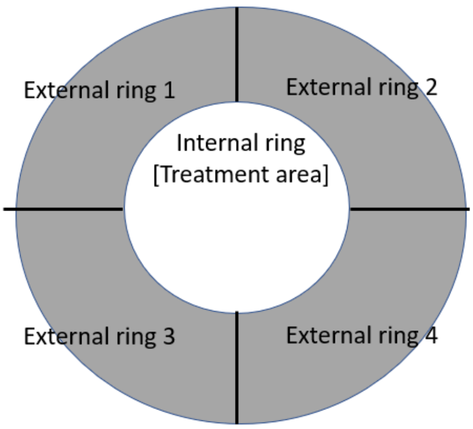


In order to satisfy the conditional independence assumption, we adopt the following multi-step procedure. In the first step, we split the external ring associated with a treated area into four quadrants with the same geographical extension (the result of this exercise for a generic area is reported in Figure 2). Then, we collect satellite time-series data on characteristics of each quadrant and the internal ring which can explain differences in their outcome (e.g. a different economic endowment), so to meet the requirement of the conditional independence assumption required by our method.

In the second step, we recombine the metrics associated with each quadrant to construct a so-called synthetic area: i.e., a control area with characteristics most similar to the treated municipality before the intervention took place. Intuitively, this is done by taking a weighted average of the metrics associated with the four quadrants, where weights are chosen so that the characteristics of the synthetic area (in terms of economic activities, natural endowment, and conflicts) best resemble the characteristics of the treated area before the intervention: i.e., differences between the metrics registered in the treatment and the control areas are minimized.

Figure 2

Quadrants used in the construction of the synthetic municipality



Finally, we compute a weighted average of the outcome of interest of quadrants after the intervention, using the weights obtained in the second step, to obtain the performance of the synthetic area after the treatment of the internal ring. We use this measure to obtain a counterfactual scenario to observe what would have happened to the treated area if had not received the treatment.

3. Constructing a synthetic control

In the previous section, we anticipated that for our methodology to work, a conditional independence assumption must credibly hold in the data. In other words, this implies that we need to identify which characteristics of the quadrants and the treated area can explain differences in their outcome. In order to do so, the researcher is required to conduct an intensive background research on which building a theory on what are the factors to be considered in order to meet the conditional independence assumption in her analysis. In what follows, we present an example of background research related to our case study, that is an intervention which took place from 2013 to 2023 to promote employability and entrepreneurship conditions and support the creation of SME's and start-ups in the rural areas of Colombia.

An example of background research

The agricultural sector is crucial to economic growth, accounting for 4% of global gross domestic product (GDP) and in some least developing countries, for more than 25% of GDP (WB, 2023). Colombia, despite being an upper middle-income country (the per capita GDP is US$14,164 in purchasing power parity),[2] still largely relies on this sector. In fact, agriculture generates 6.1% of the national GDP, and provides employment to 16.3% of the population (IFAD 2016). Nonetheless, its performance has been disappointing over the past 25 years, with a growth rate barely half that of the national economy. For this reason, the Colombian government adopted a new strategy in the last decade, aiming at providing better opportunities for rural farmers and the agriculture sector as a whole, as well as to support rural farmers through subsidies, farm inputs, agriculture machinery, and technical assistance, and to enable rural

[2] International Monetary Fund, World Economic Outlook Database.

farmers' market accession and enhance productivity and ensure resilient farming (WFP, 2017; OECD, 2021). This was the national response to the rapid increase of rural and structural transformation of Latin America and the Caribbean (LAC) economies. This is in addition of food system transformation in those countries and, i.e. in Colombia where modern food systems impact LAC supply and demand through midstream and downstream processing and wholesale, retail and transportation methods but also and the subsequent food system. Providing support to SME of the midstream, by linking them to upstream (e.g. agricultural inputs) and downstream (e.g. final consumption) IFAD's projects strengthen market linkages between small-scale producers' production and final consumption, thereby taking advantage of the increased connectivity and thus promoting improved food and nutrition security of IFAD's beneficiaries.

Providing the correct incentives and investments to enhance the economic welfare of rural areas and improve the living standards of its inhabitants is not an easy task, especially considering the many years of conflicts endured by these areas, and the unstable peace under which they currently live. Indeed, policies aiming to increase the welfare of rural people may have the effect to lower conflict by reducing the opportunity cost of appropriate resources violently. At the same time, a rise in contestable welfare may increase violence by raising gains from violent appropriation: i.e., the rapacity effect (Dube and Vargas 2013). But while the relation between welfare and conflict can be ambiguous (Caselli et al. 2015), in the case of Colombia we already have strong evidence of the fact that a rapacity effect is at play in rural areas (Angris and Cugler, 2008). Another key challenge in designing policies in support of Colombian rural areas is represented by the fact that people often engage in illegal farming (i.e., coca leaf) in order to increase their income and fill the gap of food shortages to which they are subject (UNODC, 2023). As a result, policies aiming at diffusing new and legal entrepreneurial practices also find the resistance of those controlling the market of illegal crops, increasing the level of conflict in rural areas.

In order to deal with this situation, the government of Colombia with funding from IFAD, the Agricultural Ministry of Colombia, the Spanish Fund, and the project beneficiaries, promoted "El Campo Emprende," or The Building Rural Entrepreneurial Capacities Programme: Trust and Opportunity (TOP) project from 2012 to 2023. Aim of the project is to promote employability and entrepreneurship conditions and support the creation of SME’s and start-ups in the rural areas of Colombia. Overarching goal of the project is to develop local productive organizations in regions characterized by insecurity, inequality, and post-conflict environments to improve employment, income, and living conditions.

Specifically, the objectives of the program include increasing food security, facilitating rural services, consolidating strategies for financial and physical assets, planning for rural youth, and promoting rural investments. To achieve these goals, TOP aims at forming groups (or making the informal ones formal) and to provide financial support to cover start-up costs and or working capital. Given the importance of Gender for the SDG agenda, the program prioritizes gender inclusion strategies by seeking gender equity in the staff composition, information sharing concerning the projects’ obstacles and activities among leader women, coordination with other programs for gender equity, and prioritizing women household-head groups for saving incentives, among others.

A crucial feature of TOP with respect to previous forms of intervention in Colombia is that, rather than simply strengthening regulations and punishment to motivate farmers to abandon illegal crop cultivation, the TOP project, alike many IFAD’s project, provide technical assistance and training to farmers in order to strength their productive capacities and, by heling them to create formal group, provides farmers with access to credit, training for establishing or improving an entrepreneurial activity, and programs for raising awareness on the social and economic costs of illegal activities and ultimately contribute to promote social cohesion.

Building on our background research, and consistent with the actual literature, we hypothesize that a number of factors can influence the success of the TOP project: the economic endowment of the area, the density of the present population, economic inequality, the presence of agricultural activities, and the number of conflicts.

4. An example of GIS data collection

Consistent with our background research, we collect the following data for each month from January 2014 (i.e., the first year in which all GIS data sources are available) to December 2021, for the internal and external rings constructed around the centroid of the provinces treated by TOP. The location of these provinces is reported in the appendix, Figure A1.

Gross dry matter productivity (GDMP) GDMP registers the overall growth rate of the vegetation of an area and is directly related to the productive capacity of land and changes in aggregate yields. Data is registered every 10 days at a spatial grid resolution of around 300x300m, with units customized for agro-statistical purposes.[3] Data is recorded by the Copernicus Global Land Service (European Union's Earth observation program) from January 2014, and it is widely used in agricultural studies. The unit of measurement is the number of kilocalories per hectare per day. The distribution of the value of GDMP by year in treated areas is reported in Figure A2 of the Appendix. The average value of GDMP is used in our study as the main outcome variable to assess the impact of TOP, while its standard deviation is used to determine the concentration of agricultural activities in an area.

Nighlights (NL): We use the NL data as recorded by the United States Air Force satellites and distributed by the Visible Infrared Imaging Radiometer Suite (VIIRS).[4]

[3] For additional information, see the website of the Copernicus Global Land Service at https://land.copernicus.eu.

[4] VIIRS data are designed to consistently measure the radiance of light coming from earth in a wide range of lighting conditions. They feature a high spatial accuracy and comparability across time, and they are often used as a proxy for GDP in rural and urban areas of developing countries (Gibson et al., 2020).

VIIRS data is registered every month at a spatial grid-resolution of around 300x300m, and it is available from 2012. A stray-light procedure is applied to data to correct for cloud coverage and stray light.[5] The unit of measurement is nano Watts per square centimeter per steradian. The distribution of the average values of nightlights by year in treated areas is reported in Figure A3 of the Appendix. The average value of NL is used as a proxy of the economic endowment and the population present in the area, while the standard deviation is assumed to be a measure of the dispersion/concentration of the population and the economic welfare in the area. In the Colombian case, Prem et al. (2023) show that there is a high correlation between NL and economic activity indicators as well as development indicators at the municipality level.

Conflicts We measure the level of conflicts in one area by using data on conflict events drawn from the Armed Conflict and Location Event (ACLED) dataset, a widely used data source in this field (see for instance Harari et al. 2018). The ACLED dataset provides the latitude and the longitude of the centroid of the municipality in which the conflict event occurred, date, number of fatalities, and additional characteristics of a wide range of conflict-related events. Event data are derived from a variety of sources, mainly concentrating on reports from war zones, humanitarian agencies, and research publications. Information from local, regional, national, and continental media is reviewed daily (Raleigh et al., 2010). The distribution of the average number of conflicts values another descriptive statistic in Colombia and in treated areas is reported in Table A1 of the Appendix.

[5] For additional information, see https://eogdata.mines.edu/products/vnl/

5. From theory to data: an empirical application

Once data has been constructed, the average treatment effect of the intervention can be estimated using the following regression model:

$$\Delta y_{i,t} = \alpha\mathbf{1} + \zeta_t + \eta_i + \epsilon_{i,t} \text{ (1)}$$

Where $\Delta y_{i,t}$ is the difference in the outcome of interest between the internal and the external ring i in each considered period t after treatment, α is the estimated value associated to the intercept $\mathbf{1}$, ζ_t indicates the time fixed effect, and η_i corresponds to the fixed effect for the internal and the external ring i.

Importantly, η_i is estimated under the constraint that $\sum_{i=1}^{N}\sum_{t=1}^{T}\eta_i = 0$: i.e., panel fixed effects sum to 0 across all observations in the sample. Because of this constraint, the parameter α can be interpreted as the grand average of the average difference in the outcome of each internal and external ring during the time window considered. It follows that α represents the parameter of interest for our analysis, because it corresponds to the average treatment effect.

Using this model, we analyze the impact of TOP's intervention by looking at the difference in agricultural productivity between treatment and control municipality in terms of GDMP in the period subsequent to the intervention.

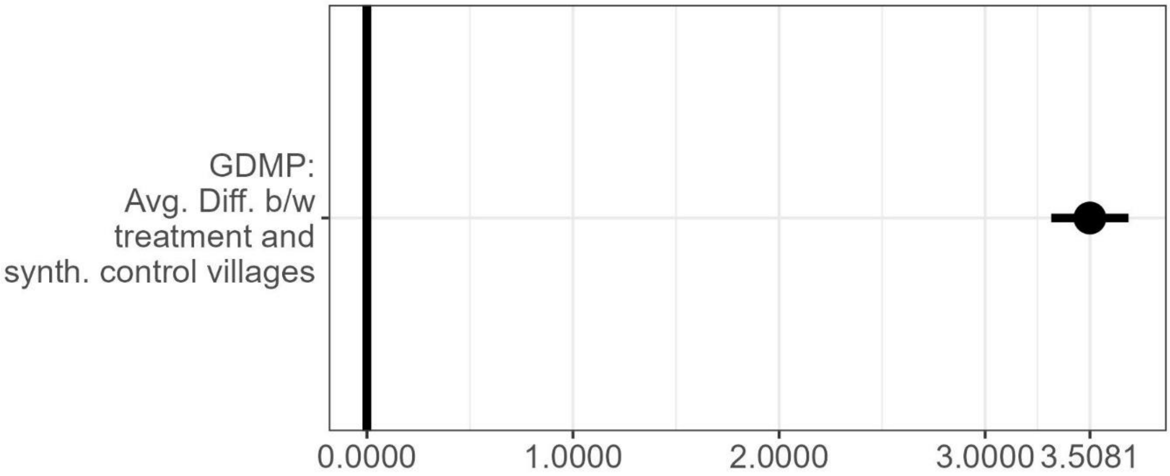
The result of this exercise is reported in Figure 3. The dot indicates the value of the average difference between internal and external rings in terms of agricultural productivity (as measured by GDMP), while the segment crossing the dot indicates the variation of this value within a 95% confidence interval.

The figure shows that it exists a positive and statistically significant difference between treatment and control municipalities after this intervention. Specifically, their agricultural productivity differs by 3.5081 points. Therefore, given that the average value of the GDMP in the treatment municipalities before treatment was 157.01, we can say that the effect of treatment was to increase agricultural productivity by roughly 3.5081/ 157.01=2.23%, which is thus the average effect size of TOP.

Again, we reiterate that given the data limitations, the interpretation of these results should be taken with extreme caution. The primary reason for presenting this result is to help the reader in understanding how to properly comment on the findings obtained with our methodology.

Figure 3

Average Treatment Effect



Possible extensions

Notably, our method can be adopted also to assess possible heterogeneities in the treatment effect.

For instance, it is possible to test whether the impact of treatment varies across treated areas. In order to do that, one could simply estimate (1) excluding the intercept and the estimation constraint that $\sum_{i=1}^{N}\sum_{t=1}^{T}\eta_i = 0$. In this way, the estimated coefficient associated to the generic parameter η_i can be interpreted as the average difference in the outcome of the internal and the external ring *i* during the time window considered. It follows that heterogeneities in the impact of treatment can be visually examined by plotting the distribution of η.

Another possibility is to test whether the impact of treatment varies across rings with similar characteristics. In this case, one should estimate (1): without the intercept, replacing fixed effects with group fixed effects (where the group correspond to the sample of rings with similar characteristics), and removing the estimation constraint that $\sum_{i=1}^{N}\sum_{t=1}^{T}\eta_i = 0$. Similar to the previous case, one can then visually inspect heterogeneities in the impact of treatment by plotting the distribution of η.

Finally, it is possible to investigate the impact of treatment across time. In this case, one can simply plot the estimated impact of each time fixed effect from (1), which indicate the average value of treatment for all rings observed in the same time period, and observe changes from one period to another.

6. Conclusions

This paper outlines a new methodology to conduct the impact assessment of a project, which is crucial for international organizations to justify the funding entrusted to them by its Member States. Specifically, this methodology is useful when standard rigorous impact assessment is not feasible, because of the lack of data and or financial resources.

This new methodology allows to conduct impact assessment when the only available data is the location of the treated population. Building upon the idea that population residing close to one another have similar characteristics and they are exposed to the same shocks, our methodology consists in comparing the outcomes of a treated area with those of a neighboring area.

Since space contiguity does not fully ensure that neighboring areas can be correctly compared, we use a synthetic control method to compare treated and control areas under a conditional independence assumption. The creation of a synthetic control requires to understand what differences between treatment and control may explain heterogeneities in their outcomes. We show how background research provides crucial insights in designing the correct synthetic control.

We show how the average treatment effect, obtained as the difference between treatment and synthetic control, can be estimated using a simple regression model. Small modifications of this model further allow the researcher to study heterogeneities in the impact of the treatment across treated areas.

The reader is walked through this methodology using a mock example constructed using the geo-localization of Colombia provinces recently involved in the TOP program of IFAD.

References

Abadie A. and Gardeazabal J. (2003). "The Economic Costs of Conflict: A Case Study of the Basque Country." American Economic Review. 93(1), 113-132.

Abadie A., Diamond A. and Hainmueller J. (2010) Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California’s Tobacco Control Program, Journal of the American Statistical Association. 105:490, 493-505.

Angris J. D and Kugler D. A. (2008). Rural Windfall or a New Resource Curse? Coca, Income, and Civil Conflict in Colombia. The Review of Economics and Statistics. 90(2), 191-21.

Arias M.A., Ibáñez A.M., Zambrano A. (2019). Agricultural production amid conflict: Separating the effects of conflict into shocks and uncertainty. World Development. 119, 165-184.

Banerjee A., Duflo E., and Qian N. (2020). On the road: Access to transportation infrastructure and economic growth in China. Journal of Development Economics: 102442.

Bernal, C., Prem, M., Vargas, J. F., & Ortiz, M. (2024). Peaceful entry: Entrepreneurship dynamics during Colombia’s peace agreement. Journal of Development Economics. 166, 103119.

Butts, K. (2023), Difference-in-differences estimation with spatial spillovers, mimeo, available at: arXiv:2105.03737.

Caselli F., Morelli M. and Rohner D. (2015). The geography of interstate resources wars. The Quarterly Journal of Economics. 130, 1, 267-315.

Cavallo E., Galiani S., Noy I. and Pantano J. (2013). Catastrophic nature disasters and economic growth. The Review of Economics and Statistics. 95(5), 1549-1561.

Chen Y., Jin G.Z., Kumar N. and Shi G. (2013). The promise of Beijing: Evaluating the impact of the 2008 Olympic Games on air quality, Journal of Environmental Economics and Management. 66, 3, 424-443.

Dinkelman T. (2011). The effects of rural electrification on employment: New evidence from South Africa, American Economic Review. 101, 7, 3078-3108.

Druckenmiller H. & Hsiang S. (2018). Accounting for unobservable heterogeneity in cross section using spatial first differences (No. w25177). National Bureau of Economic Research.

Dude O and Vargas J. F (2013). Commodity price shock and civil conflicts: Evidence from Colombia. The Review of Economic Studies. 80, 4(285),1384-1421.

Duflo E. and Pande R. (2007). Dams, The Quarterly Journal of Economics. 122, 2, 601-646.

Faber B. (2014). Trade integration, market size, and industrialization: evidence from China's National Trunk Highway System. Review of Economic Studies. 81(3), 1046-1070.

Gibson J., Olivia S., Boe-Gibson G. and Chao L. (2020). Which nightlights data should we use in economics, and where? Journal of Development Economics. 149, 102602.

Henderson J.V., Storeygard A. and Weil N. D. (2012). Measuring economic growth from outer space. American Economic Review. 102, 2, 994-1028.

Henderson V., Storeygard A. and Weil D. N. (2011). A Bright Idea for Measuring Economic Growth. American Economic Review: Paper & Proceedings. 101, 3, 194-199.

IFAD – International Fund for Agriculture and Development (2016). Republic of Colombia country strategic opportunities programme, Note to Executive Board representatives, EB 2016/119/R.14 - 11(a), available at: https://webapps.ifad.org/members/eb/119/docs/EB-2016-119-R-14.pdf?attach=1.

IFAD - International Fund for Agriculture and Development (2016). Republic of Colombia country strategic opportunities programme, available at https://webapps.ifad.org/ members/eb/119/docs/EB-2016-119-R-14.pdf?attach=1

Jardim E. S., Long M. C., Plotnick R., van Inwegen E., Vigdor J. L. and Wething H. (2022). Boundary Discontinuity Methods and Policy Spillovers. National Bureau of Economic Research, No. w30075.

Korovkin V. and Makarin A. (2023). Conflict and Intergroup Trade: Evidence from 2014 Russia-Ukraine Crisis. American Economic Review. 113(1), 34-70.

Mabiso A., Patacchini E., Savastano S., and Leone Sciabolazza V. (2022). Using GIS as a Potential Methodology to Assess the Spillover (Indirect) Effects of the International Fund for Agricultural Development's Interventions, IFAD technical report. Available at https://www.ifad.org/en/web/knowledge/-/using-gis-as-a-potential-methodology-to-assess-the indirect-effects-of-ifad-interventions.

Michaels G. (2008). The Effect of Trade on the Demand for Skill -Evidence from the Interstate Highway System. Review of Economics and Statistics. 90, 683-701.

Michalopoulos S., and Papaioannou E. (2016). The long-run effects of the scramble for Africa. American Economic Review. 106, 7, 1802-1848.

Montalvo J. G. and Renayal-Querol M. (2021). Ethnic density and growth: Revising the evidence. The review of Economics and Statistics. 103(3):521-532

Nunn N. (2008). The Long-term Effects of Africa's Slave Trades. The Quarterly Journal of Economics. 123(1), 139–176.

OECD (2021). Agricultural Policy Monitoring and Evaluation 2021 Addressing the Challenges Facing Food Systems. OECD Publishing.

Prem M., Purroy M. E., Vargas J. F. (2023). Landmines: The Local Effects of Demining. Working paper.

Raleigh C., Linke A., Hegre H. and Karlsen J. (2010). Introducing ACLED: An Armed Conflict Location and Event Dataset. Journal of Peace Research, 47(5): 651–660.

UNODC - United Nations Office on Drugs and Crime (2023). World Drug Report 2023, United Nations publication.

WB - World Bank (2023). Agriculture Overview: Development news, research, data, available at https://www.worldbank.org/en/topic/agriculture/overview

WFP - World Food Program. (2017). Colombia country strategic plan (2017-2021), available at: https://www.wfp.org/operations/co01-colombia-country-strategic-plan-2017-2021.

Appendix

Figure A1 displays the distribution of treated municipalities at the Colombia country level. The white color within the map surrounded by boundaries in black indicates districts of Colombia, and areas in green indicate treated municipalities in each district. The figure indicates that provinces are heterogeneously positioned across the country, and they have notable differences in terms of area.

Figure A1

Treated municipalities.

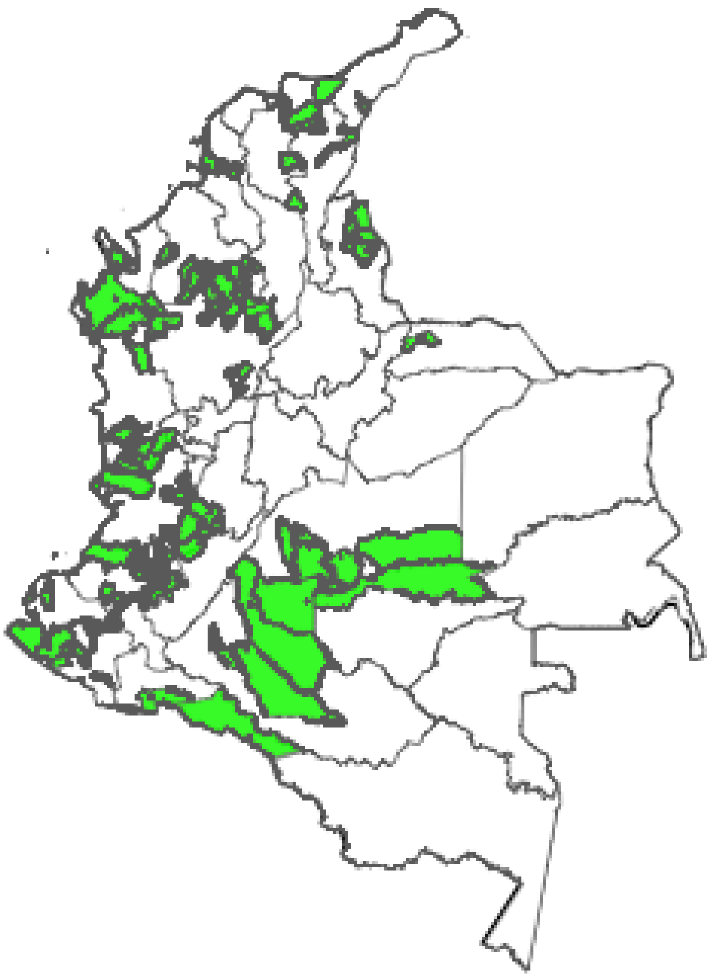
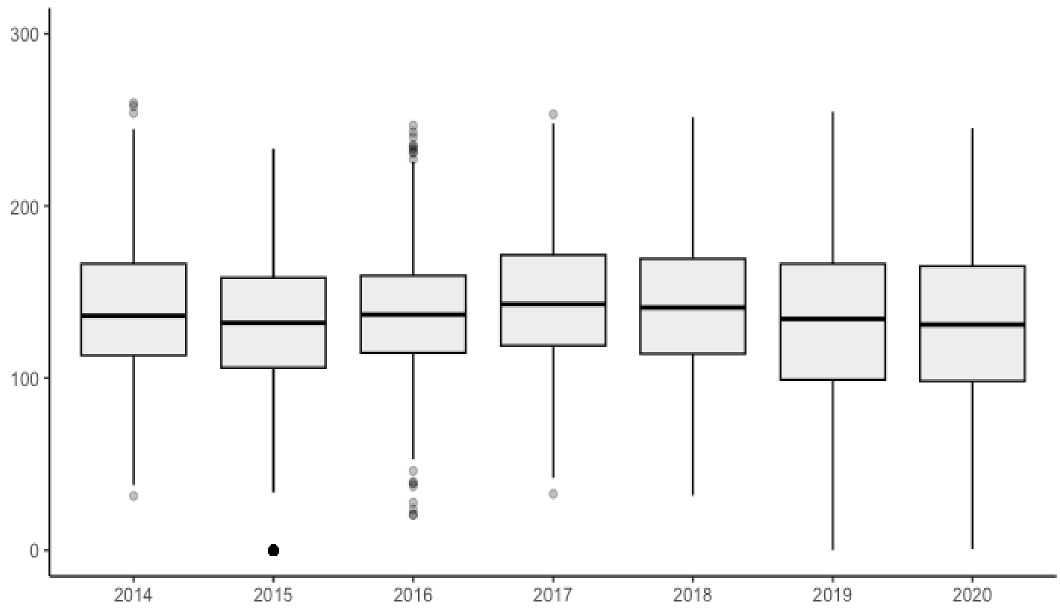


Figure A2[6] shows the average value of agriculture productivity in treated provinces. The figure shows that on average agricultural productivity remained stable in the considered time window.

Figure A2

Quantile Distribution of GDMP by year



The boxplot in Figure A3 illustrates the average distribution of nightlight intensity in treated provinces. The boxplot shows highly skewed distributions, indicating large differences in the intensity of nightlights across provinces. Moreover, we observe significant variations in this distribution in more recent years.

[6] Figure A3 illustrates the average variation of the agriculture productivity values before and after the treatment.

Figure A3

Quantile Distribution of Nightlights by year

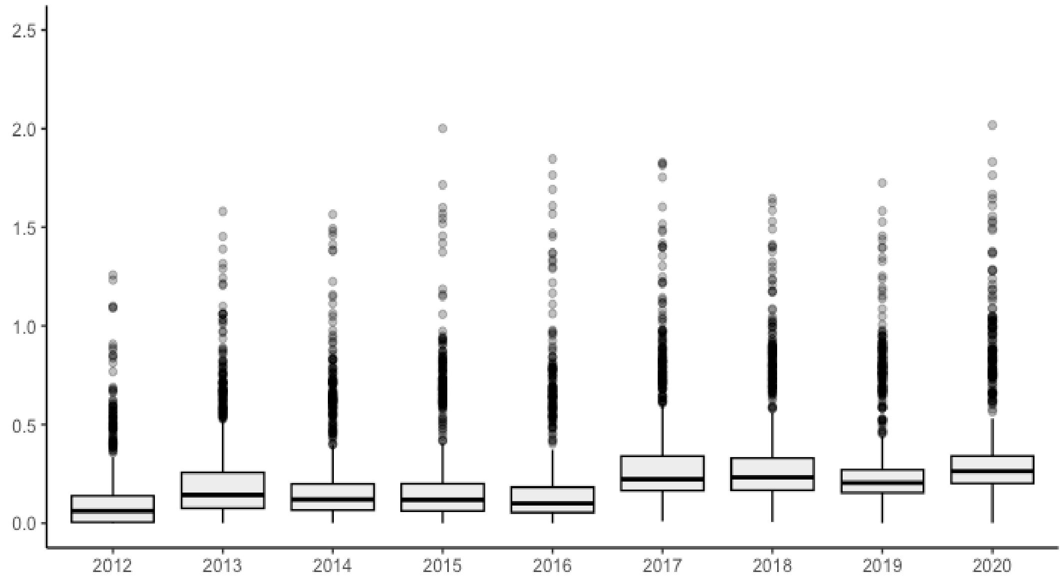


Table A1 report summary statistics on the number of conflicts experienced by all Colombian provinces, and treated provinces. Interestingly, we observe that treated provinces are significantly less affected by conflict with respect to the rest of the country.

Table A1 Number of Conflicts in Colombia and in Treated Municipalities

Item	**Minimum**	**Mean**	**Median**	**Maximum**
Colombia	3	31.56	26.0	117
Treated municipalities	0	3.24	2.5	31