How impact evaluation methods influence the outcomes of development projects? Evidence from a meta-analysis on decentralized solar nano projects

Fatoumata Nankoto Cissé

UNIVERSITÉ PARIS 1 PANTHÉON SORBONNE

Documents de Travail du Centre d'Economie de la Sorbonne

CES Working Papers

CES
Centre d'Économie de la Sorbonne
UMR 8174

How impact evaluation methods influence the outcomes of development projects? Evidence from a meta-analysis on decentralized solar nano projects

Fatoumata Nankoto CISSE

2022.08

Maison des Sciences Économiques, 106-112 boulevard de L'Hôpital, 75647 Paris Cedex 13
https://centredeconomiesorbonne.cnrs.fr/
ISSN : 1955-611X

How impact evaluation methods influence the outcomes of development projects? Evidence from a meta-analysis on decentralized solar nano projects

Fatoumata Nankoto CISSE*

2022

Abstract

This study analyzes the effect of impact evaluation methodologies on the positive and negative outcomes of decentralized solar nano projects in developing countries. Data originate from the Collaborative Smart Mapping of Mini-grid Actions (CoSMMA) developed by the Foundation for Studies and Research on International Development (FERDI). This study is based on a total of 727 tested effects from 10 decentralized solar nano projects which have been measured by experimental and quasi-experimental approaches. Using a multinomial-logit regression shown that randomized and non-randomized evaluation methods have a similar probability of generating a proven favorable outcome on the sustainable development of decentralized solar nano projects. By estimating a complementary log-log model, projects are most often evaluated as successful when effects on education are tested. In addition, a discrepancy of impacts is found between randomized control trials and difference-in-difference strategies in proven-unfavorable outcomes of projects. This analysis also highlights the convergence of impacts between randomization and matching techniques on projects implemented in Africa. Findings from this paper provide strong evidence for development practitioners to choose the appropriate impact assessment method.

Keywords: Impact evaluation; Meta-analysis; Experimental methods; Quasi-experimental methods; Randomized control trials; Matching; Difference-in-difference; Decentralized electrification; Sustainable development.

JEL Classification: C18, C90, F63, O12, O13, O22, Q01, Q42

---

* Centre d'Economie de la Sorbonne, University Paris 1 Panthéon-Sorbonne/Investisseurs et Partenaires (I&P). Email: Fatoumata.Cisse2@etu.univ-paris1.fr. The author is grateful to Pr. Mathilde Maurel, Pr. Jean-Claude Berthélémy, and Pr. Lisa Chauvet for the stimulating discussions and guidelines. But also, to the FERDI for providing the data through its "Electrification Access Program". The author expresses its gratitude to Samuel Monteiro and Jean-Michel Severino for their helpful comments and advices. The author also would like to thanks Mamadou Barry and Adrien Desroziers for their proofreading and comments. The author also appreciates all the comments from the informal seminar for PhD candidates in development economics affiliated with Paris 1 Panthéon-Sorbonne University. The usual disclaimers apply.

Content

Figures

Tables

1 Introduction

Over the past several decades, debates within the development economics field regarding impact evaluation methodologies have rapidly increased. Impact assessment provides guidance that informs policy-makers and practitioners in their decision-making strategies. Funders and operators increase their investment in impact assessment and strengthen their partnerships with researchers in order to accurately identify the results of their projects and programs. Specifically, the expansion of impact assessment in development economics highlights the lack of knowledge regarding the effects of impact evaluation tools on assessed outcomes.

The need for short and mid-term results increases the demand for diverse, flexible, and rigorous impact measurement tools. To that extent, Africa's Pulse report (2018), published by the World Bank, gathers numerous studies which use several impact assessment techniques. The classic impact methods, performed through quasi-experimental strategies, are based on observational models and standard theories. According to Banerjee (2020), the rise of natural experiments during the last two decades has created new frameworks and theories about poverty and development issues.

This paper contributes to the impact evaluation literature by analyzing the influence of impact evaluation methodologies on the results of sustainable development projects. The objective is to detect whether randomized and non-randomized methods have the same impact or not on the probability of finding success or failure of decentralized electricity programs. Or rather, if the influence of their estimates differs.

This study also contributes to the pioneering impact evaluation literature initiated in the 1970s, based on the need to assess the macroeconomic impacts of World Bank public policies (Guillaumont, 1985). Therefore, the assessment of development project outcomes is becoming an increasingly important issue for both academics or public and private operators. Cameron et al., (2016) observe a record of 4,600 published evaluations in June 2018, while only 132 experimental and quasi-experimental evaluations were published prior to 2000. According to development actors, project impact assessments are a key matter in development economics. Different tools and assessment methodologies are used, which mainly rely on experimental and quasi-experimental strategies.

However, the main constraints of impact evaluation arise from statistical and technical characteristics. For instance, there is a mismatch of time frames between academics, investors, and operators in the development field. This is supported by Bédécarrats et al., (2020), who argue that method constraints can force researchers to restrict themselves to short time frames and specific populations or geographic areas, and therefore produce unusable results. On the other hand, the cost of impact evaluation is a major challenge which can constitute 10 percent of the total cost of small-sized projects (FERDI, 2019) and can thus influence the choice of methodology employed. For example, the cost of specific methodologies may not be feasible for some private operators and investors for whom short term financial performance and gain are the main decision criteria.

A broad swath of the literature finds evidence of statistical and technical differences among impact evaluation methods (Duflo, 2005), while a minority concludes that, for specific development projects, the outcomes do not really differ according to the methodology that has been used (Rodrik, 2008). Currently though, there is limited empirical evidence on the influence of impact evaluation methodologies for a consistent set of development projects (Glazerman et al., 2003).

Thus, using an original meta-data set which gathers 151 scientific assessment studies of 483 decentralized electrification projects (DEP) in 62 countries for the period 1981-2019, I contribute to the broader literature of impact assessment by simultaneously comparing the effect of both randomized and non-randomized methodologies on project outcomes. Furthermore, this research highlights the conditions under which there is a discrepancy or convergence of the estimates generated by experimental and quasi-experimental methods. As observed by

Camfield and Duvendack (2014), there is a need to conduct meta-analyses, which are still too scarce in development economics.

As meta-analysis is a valuable technical input which allows us to address the matter by comparing the impact of several studies. The interest of the meta-analysis is to provide a consolidated result and to control for a certain type of bias from first-hand studies. The main objective of this paper is to identify which impact evaluation methods provide a positive and significant effect of decentralized electrification projects, in order to identify which methodologies generate significant and negative effects on project outcomes.

Obviously, the purpose of this paper is not to renew the debate on randomized control trials (RCTs) versus non-RCTs but to go beyond these discussions by identifying what impact assessment methods function better under the implementation of small-scale infrastructure projects, and to what extent. Therefore, this study is a plea for economic development actors to adopt an economic perspective rather than a technical one. To shed light on this issue, I simultaneously analyze the outcomes of impact evaluation methods on four possible evaluation outcomes based on the solar nano systems characteristics: *proven-favorable, proven-unfavorable, unproven-favorable, unproven-unfavorable*.

To the best of my knowledge, no other impact evaluation study has rigorously identified the influence of econometric assessment methods on the outcomes of small-scale infrastructure projects in developing countries. Indeed, using the results of similar studies to highlight the impact of statistical techniques is challenging. There is only a sole meta-study developed by Berthélemy and Millien (2018) which looks at the impact of DEP characteristics on sustainable development.

Empirical findings from my research reveal that there is no discrepancy between the likelihood of randomized and non-randomized methods to find a significant positive outcome from decentralized solar nano projects, mainly when I estimate the effects of projects designed for the educational sector. I also find that quasi-experimental approaches such as the Difference-in-Difference (DiD) method have a significantly lower probability (-3 percentage points) of finding a negative and significant project outcome, compared to RCTs. This means that, in terms of finding significant failure in decentralized solar nano projects, the DiD method is less pessimistic than are RCTs.

My baseline results hold when, as a robustness check, I restrict my sample to decentralized solar nano projects implemented in Africa. This is a valuable contribution to the discussions concerning the extension and generalization of impact assessment policies and the context settings. This evidence also provides stronger empirical conclusions on the uneven geographical distribution of impact evaluations, which are influenced by the connections between researchers and local entities such as non-governmental organizations (NGOs) (Ravallion, 2020).

This paper presents the related literature and research question in Section 2. In Section 3 the data and the key descriptive statistics are explored. Section 4 sets the methodological strategy. Section 5 presents the main empirical results and Section 6 challenges the robustness of the evidence. Finally, I conclude and discuss the findings in Section 7.

2 Related literature and research question

2.1 The contribution to the meta-analysis literature in development economics

Stanley and Jarrell (1989, p.161) define meta-analysis as "the regression analysis of regression analyses". Originally developed in medical research, meta-analysis in economics has been used to “calibrate structural models, examine patterns of publication bias, and explain the differences in the results of individuals studies” (Berthélemy and Millien, 2018, p.7). Doucouliagos and Paldam (2007), who perform a quantitative forensic analysis of the aid effectiveness literature.

Meta-studies identify the weakness of primary research and give more precision and representativeness to an estimate of treatment effects (Pang et al., 1999; Simes and Glasziou, 1992; Mugford et al., 1989). The methodology of meta-regressions goes beyond classification studies and qualitative review of structural models and estimates.

In the matter of energy efficiency, Labandeira et al., (2020) conduct a meta-analysis to measure the effects of energy efficiency policies regarding energy demand and the price of associated durable goods. They use 366 research papers on energy efficiency policies, which provide 1,375 estimates of the impact on energy demand and 108 estimates on the price of durable goods that consume energy. They find a significant reduction in energy demand, which is even greater for studies based on experimental designs, mostly because experimental studies are more likely to get published in top ranking journals.

On another issue, Corduneanu-Huci et al., (2021) published a seminal work in the field of political economics by assessing the impact of political environments on public policy design. They argue that RCTs occur mostly in Indian state jurisdictions where political competition is in a paroxysm. The authors identify this relationship through a "supply" (from researchers, institutions, NGOs, and donors) and a "demand" (from governmental decision-makers and channel drivers), which reveals the "political site selection bias" in the RCT field. Their study echoes the debate regarding what impact assessment method selection should be used to evaluate the effects of the Sustainable Development Goals (SDGs). It is shown in **Section 5** that this cognitive bias also affects the outcomes generated by off-grid solar nano projects.

2.2 Why is impact assessment fundamental for development economics?

Pamies-Sumner's study (2014) highlights the wide expansion of impact evaluations of development projects during the 1990s, when funders and researchers in the development field examined the effectiveness of development assistance. According to Gertler et al., (2011), the prospective impact studies aim to measure whether development projects achieved their expected results or not, and thus evaluate several policies to measure these results. Thus, impact measurements allow us to empirically assess the outcomes of projects by ensuring that development actors make the trade-off between the potential benefit and harm of their activities. Empirical investigations allow researchers, donors, policymakers, and operators in development economics to identify what projects work and in which context.

Thanks to the measurement of welfare changes for recipients, impact evaluations are seen today as global public goods, as they reflect reliable toolkits for public and private decision-makers, especially with the fundamental shift in development economics toward the issues of private goods access in rural development, health and educational matters that occurred in the 1970s (Morduch, 2020).

Hence, impact evaluations rely on a wide range of measurement techniques that effectively isolate project impacts (Baudet, 2019). These causal inference methodologies calculate a counterfactual which compares the changes for recipients to the prevailing situation if the project had not been implemented.

However, the greatest challenge of impact assessment is to find out what works and where, under different circumstances (Deaton, 2020). Ravallion (2020) highlights the importance of the choice of impact method according to the identification of the relevant counfounders, the type of project, the budget cost of the evaluation, the sample size covered by a project, and the impact parameters. This allows policymakers and practitioners to acknowledge the benefits and failures of their funded projects. The identification of a project's relevance requires knowledge on structural-econometric methods (Ravallion, 2020). On the other hand, impact evaluations also provide evidence on the efficacy of development projects and can help to improve some aspects of their implementation (Legovini and al., 2015).

2.3 Convergence or discrepancy between impact assessment methodologies?

For the past decades, several impact evaluation techniques have been applied in order to analyze the successes and failures of development projects. These methodologies rely mainly on quantitative and scientific parameters, under experimental and quasi-experimental modeling.

Experimental methodologies are mainly based on randomized control trial settings, which isolate the impact of projects by comparing, through a lottery draw, for instance, the behavior of recipients and non-recipients. According to promoters of randomization, such as Esther Duflo (2005) from the J-Pal[1] research laboratory, RCTs enable researchers to measure causal inference and to better correct selection bias, unlike non-randomized techniques. RCTs are considered as the gold standard research design, followed by observational studies (Pang et al., 1999). Therefore, development economists have been widely using RCTs for the past two decades (Deaton, 2020).

However, (Deaton, 2009; Rodrik, 2008) highlight the lack of statistical power and the operational limits of the randomization approach. Indeed, RCTs encounter the same issues as observational studies of internal validity caused by the circumstances of the project and the behavior of recipients and non-recipients throughout the intervention, as well as by the external validity of the evaluation design, based on the context and the wider application of development projects. These limits raise the question of representativity of the samples

[1]Abdul Latif Jameel Poverty Action Lab

compared to the targeted populations. As expressed by McKenzie (2012), this weakens the ability to draw conclusions and would require unrealistic sample sizes. Deaton (2020) adds that RCTs are affected by the same issues of inference and estimation of other methods.

Quasi-experimental methods try to correct these limits and reduce the selection bias by simulating the framework conditions of an experiment with an ex-post design of the counterfactual, such as the matching technique, which identifies and matches recipients to the most equivalent non-recipients based on their observed characteristics. Moreover, the Difference-in-Difference (DiD) approach can estimate the counterfactual for the impact variation of recipients by using the variation of non-recipients. Next, the Regression Discontinuity Design (RDD) measures the local average impact near an index of eligibility, which is a threshold that differentiates beneficiaries and non-beneficiaries. Therefore, the debate regarding the robustness of impact estimates has led to a dichotomy framework between experimental and quasi-experimental methodologies in the sustainable development field (Bédécarrats et al., 2017). To that extent, Lalonde (1986) conducts a RCT study of the US Job Training Partnership Act program and finds that several papers show significant disparities between experimental and quasi-experimental methods.

As one can observe in **Appendix 9.1**, the application gap among impact evaluation methods has increased during the past few decades, mainly with the predominance of RCT studies. The International Initiative for Impact Evaluation (3iE) reveals that more than 300 RCT studies were conducted in 2012 (Review 3ie, 2018). The prevalence of RCTs is also reflective in the decentralized electricity access field, with more than 200 effects measured and recorded in the Collaborative Smart Mapping of Mini-grids Actions (CoSMMA) database developed by the FERDI (**Appendix 9.2**). This evidence converges with the trend noted by Picciotto (2020), who observes that 62 percent of impact evaluations included in the 3iE repository used only RCTs, and 5 percent used a mix of RCTs and quasi-experimental techniques. This finding is supported by Banerjee et al., (2016) who argue that between 1990 and 2015, two-thirds of articles published by development economics journals were RCT evaluations. However, the trend presented in **Appendix 9.2** describes that between 2017 and 2018, the application of quasi-experimental methods was regaining the impact assessment field of decentralized electrification projects.

On the other hand, Heckman et al., (1997) assess that there is a quasi-neutrality of the selection bias when non-experimental techniques are applied. Afterwards, Glazerman et al., (2003) come to a similar finding with the alleviation of the bias between experimental and quasi-experimental estimators, mostly when the Matching measure has been applied and the sample is weighted. Vivalt (2020) fails to reject the null hypothesis that effect sizes estimated by RCTs and non-RCTs are the same for a sample of 635 studies in international development. It is worth noting that Buddelmeyer and Skoufias (2004) find an equivalence between the randomized and non-randomized outcomes of Mexico's *PROGRESA* conditional cash transfer program evaluation. Their findings are substantial, given that broader RCT studies have been conducted to assess this same conditional cash transfer program in Mexico (Gertler and Boyce, 2003).

2.4 Research question

The aim of this paper is to compare the effects of impact assessment methods on the outcomes of development projects, using data collected in the CoSMMA database and developed by the FERDI (2019). Ravallion (2020) highlights that there is a need for more research on the distribution of estimates from observational studies, and that they should be compared with estimates from RCTs in the same setting. With the introduction of the characteristics and effects of decentralized solar nano projects, I address the following research questions: i) Do impact evaluation methods have an influence on the outcomes of decentralized solar nano projects? ii) Is there an equivalence or discrepancy between the impacts of randomized and non-randomized methodologies on the results of decentralized solar nano projects? iii) Which evaluation methods are most likely to conclude success or failure of decentralized solar nano projects?

Obviously, the RCT versus non-RCT debate is well-known, which is why the additional value of this study is to provide insights and empirical facts to the impact evaluation literature, in line with the meta-analysis conducted by (Berthélemy and

Millien, 2018). While most of the studies in the literature perform distinguished analyses of impact methodologies, to my knowledge my empirical study is the first to carry out a simultaneous comparison of impact evaluation techniques for the same setting, which assembles decentralized solar nano projects which have been implemented in developing countries.

Furthermore, this research provides evidence that indicates whether experimental and quasi-experimental approaches find divergent or convergent outcomes from decentralized solar nano projects. These findings can shed light on best practices that public and private developers can rely on for the impact assessment and monitoring policies of their development projects.

To address the previous questions, this research focuses on small-scale infrastructure projects from the decentralized electrification sector, which enjoys the availability of data and also represents a growing research matter in development economics. As the decentralized electrification sector focuses on the last mile problems context, policymakers and practitioners need to evaluate their investments. In order to test my assumptions, the empirical model analyzes at the baseline data of a homogenous sample of decentralized solar nano projects, but one with disparities on governance schemes, geo-graphical locations, and produced effects.

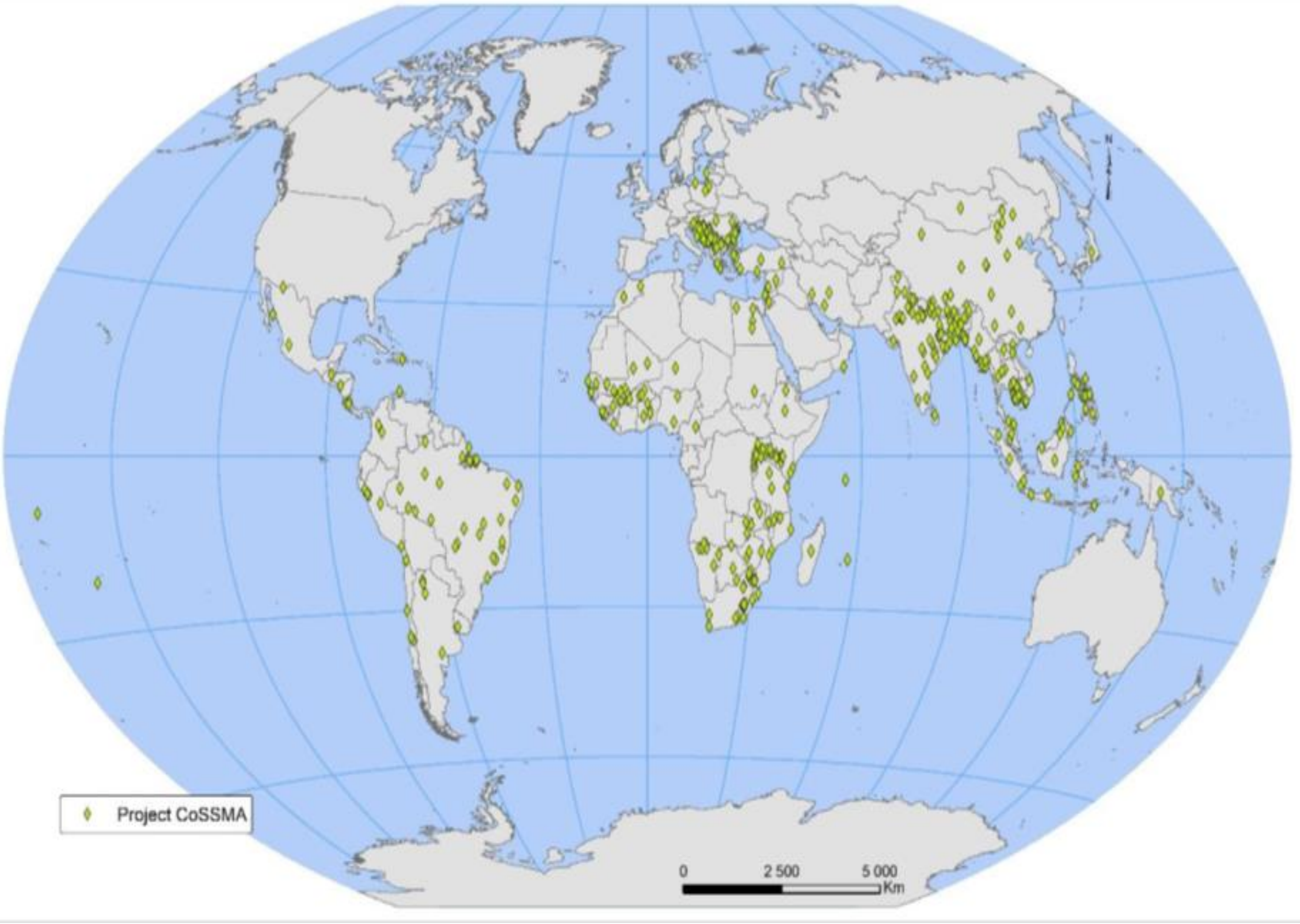
3 Data source and stylized facts

3.1 The CoSMMA database

Following the methodological steps of Berthélemy and Millien (2018), my study is the first meta-analysis in the development economics literature which attempts to relate impact evaluation techniques with the outcomes generated by off-grid solar nano projects. Such an approach is relevant for economists, given that the combination of data from studies using different designs is advantageous to external validity strategies and policy decision-making (Pang et al., 1999).

The CoSMMA database, developed by the FERDI and the University of Paris 1 Panthéon-Sorbonne, is an information instrument for the analysis and comparison of decentralized electrification projects. The CoSMMA illustrates the assessment of socio-economic benefits of decentralized electrification based on the specifications and on the observed, measured, and tested effects of projects, with the aim of highlighting the most efficient projects, namely those which generate the expected socio-economic effects. Therefore, this geographical and technical mapping tool allows the identification of implemented and evaluated projects, as well as for the identification of their impacts on targeted beneficiaries. **Figure 1** shows a broad and proper geographical distribution of DEP in developing countries, collected by the CoSMMA.

Figure 1: CoSMMA map



Source: CoSMMA, 2018

The source of information of the CoSMMA comes from two large document categories:

- Research papers reviewed by a referee committee and published in a scientific.
- Journal.
- Institutional reports.

The PRISMA scheme in **Figure 2** shows that to build the CoSMMA, developers relied on four scientific and economic database searches gathered by the online *EBSCO* library through keyword searches which produced six sets of articles called "packs" or "bunches"[2]: *Academic Search Premier*, *Business Source Complete*, *EconLit*, and *GreenFILE*. This sampling by keyword searches gives a random selection of DEP assessment studies. Almost 90 percent of the evaluation documents are published or working papers from scientific journals, and 10 percent are from institutional or corporate reports in the energy sector.

As a result of this selection, the CoSMMA has collected the following data:

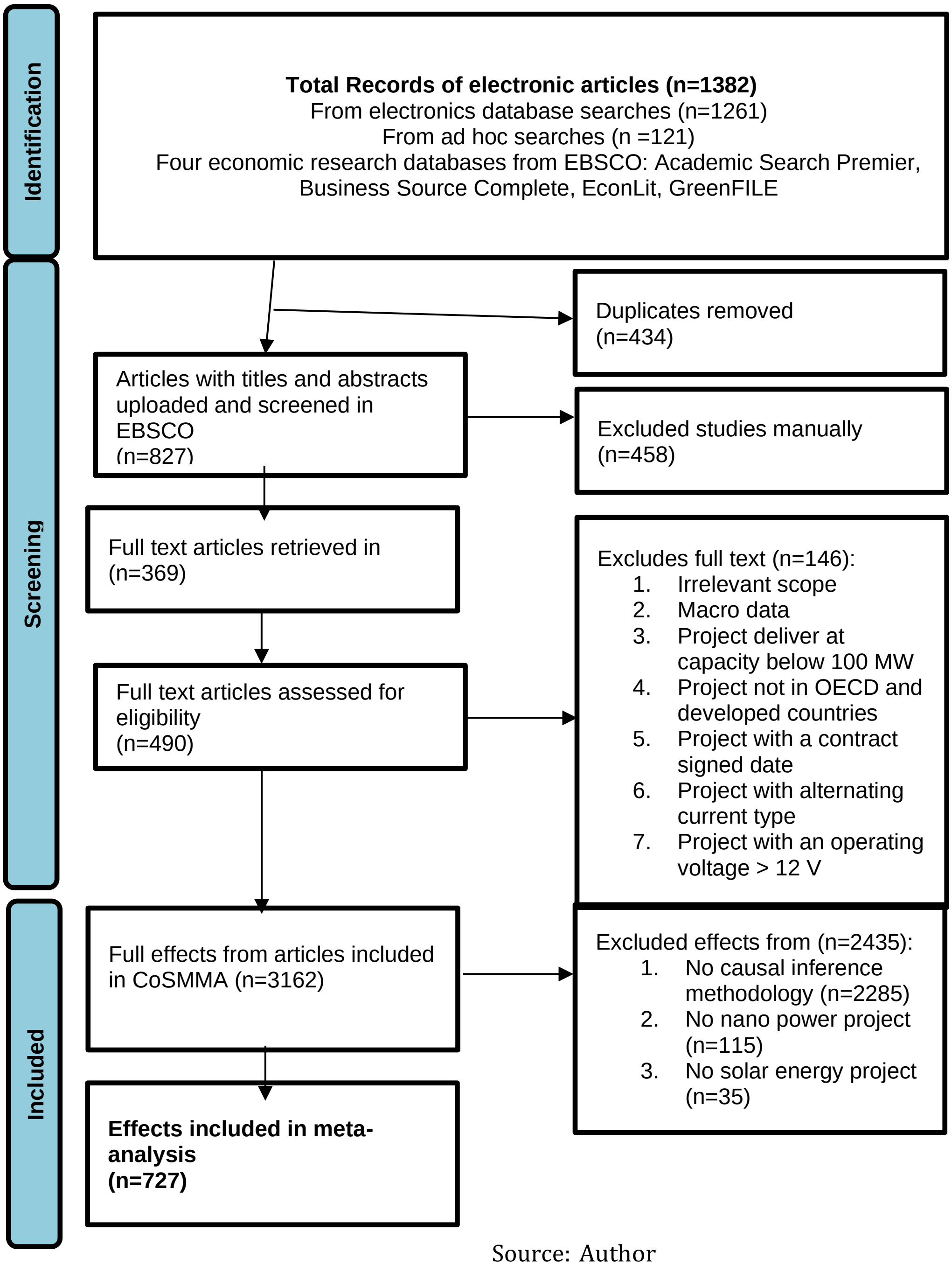
- **483** decentralized electrification **projects**;
- deployed in **561 production units**;
- located in **62 countries**;
- from **151 scientific assessment documents**;
- that describe **3,162 observed and tested effects**.

Also, one of the main rules of the CoSMMA data collection is based on the decentralization principle, which relies on the following principles:

- The production, transportation, and distribution of electricity without establishing contracts with the national network.
- The energy generated does not result in a price fixed by a clearing agency.

[2] See the extraction codes in Millien (2019).

Figure 2: PRISMA Scheme



Source: Author

3.2 Descriptive statistics

The assessment methodologies collected by the CoSMMA are classified as miscellaneous" (Berthélemy and Millien, 2018), and range from observed effects to measured and tested effects, through descriptive observations. The interest of my paper is that it focuses solely on measured and tested effects in order to analyze the impact of experimental and quasi-experimental approaches on the outcomes of decentralized solar nano projects. This leads to a limitation of the sample of the scientific data presented in **Table 1**. The subset consists of projects which have been evaluated by econometric measurement tools with standard deviations. Given that, the interest provides to issues in statistical identification is increasing in the economics field (Ravallion, 2020).

Table 1: Distribution of observed CoSMMA data

Denomination	Type	Nb of Obs (N)	Distribution
Scientific data	Quantified effect with variance	N>1	1775
Expert data	Quantified effect without variance	N=1	398
Expert data	Documented effect from research	N=0	868
Expert data	Unmeasured effect	N=0	121
Total			**3162**

3.2.1 Which evaluation methods are used to assess the impact of solar nano projects?

The subset of scientific data aggregates several evaluation methods which allow for the assessment of effects generated by decentralized solar nano projects. These methods are experimental (RCTs), quasi-experimental (DiD and Matching), and non-inferable econometrics.

Experimental techniques isolate the average impact of a project between random treated and non-treated units, while quasi-experimental methods are observational studies in which the assignment is purposive and not random, based on the prior characteristics of the recipient and non-recipient units (Ravallion, 2020). Units in the treatment and nontreatment groups are designed to ensure that they are similar with respect to the characteristics that may influence the outcome (Picciotto, 2020). Thus, the non-inferable econometrics approach examines the correlation between two variables without rectifying the endogeneity bias problem, hence without a causal inference strategy, although this latter methodology is not included in my empirical strategy because it represents only 4% of the measured effects generated by decentralized solar nano projects.

The distribution of the impact assessment approaches of my subset sample is presented in **Table 2**.

3.2.1.1 Experimental method: Randomized Control Trials

Randomized control trials measure the impact of a project by comparing the mean outcomes for beneficiary units (treated group), those who randomly have access to the project, with other random units who do not benefit (control group) from the project. RCTs are mostly suited to projects with clearly identified beneficiaries and non-beneficiaries in short-time horizon deployments (Ravallion, 2020). RCTs involve experiments in real settings and are more easily suited to private goods than to public goods interventions (Morduch, 2020).

Pritchett (2020) also assesses that, to be successful, RCTs need a clean assignment of units to the treatment and control groups and enough units for adequate statistical power. Hence, the RCT method is better applied to individualized interventions rather than to national development or sector-wide transformations. The analysis of Bédécarrats et al., (2019) provides a descriptive affirmation that less than five percent of RCT impact assessments conducted by the J-PAL have led to scaled-up policy changes.

The interpretation of RCTs requires assumptions and a specific survey, which is characterized by lower sample sizes and higher variances than observational studies for a given budget (Ravallion, 2020). Indeed, the cost of a classic RCT is rated between $500,000 and $1,500,000 (Bédécarrats et al., 2020). The World Bank specifies that RCTs evaluation costs are higher than those of observational studies. The timeline to implement an experimental approach can range from one year to a decade[3], considering the time needed to collect baseline data and project effects. RCTs can also require several data collection processes (Gertler et al., 2011).

Table 2 shows that the RCT approach has been applied to assess 59 percent of the effects generated by decentralized solar nano projects. These RCT evaluations randomly isolated the recipient and non-recipient units of decentralized solar nano features. Yet RCT studies have usually faced ethical issues, compared to observational studies (Teele, 2014). Moreover, Abramowicz and Szafarz (2020) argue that the principle of equipoise faced in the medical world also exists in the development field, through experimental strategies.

3.2.1.2 Quasi-experimental methods: Matching and Difference-in-Difference

Quasi-experimental studies estimate the causal impact of a project on a targeted population, without random assignment. Assessors use some statistical criteria other than random selection to design the treatment condition of the projects. Non-randomized techniques are particularly used when it is not practical or reasonable to carry out randomization assignments. There are several types of quasi-experimental designs which include Matching and Difference-in-Difference (DiD) models.

Rosenbaum and Rubin's well-known study (1983) revealed that the predicted values of the propensity score matching model are propensity scores used in selecting observational balanced treatment and comparison groups. The Matching approach identifies and matches specific recipients with the most comparable non-recipients, based on their characteristics. This impact evaluation technique has been applied for 25 percent of the tested effects generated by decentralized solar nano projects (**Table 2**).

The DiD methodology evaluates the impact of a project by considering the overtime invariant differences between recipients and non-recipients (Le Ster, 2011). Gertler et al., (2011) specify that the DiD estimate compares the average change over time in the treatment group outcome variable with the average change over time in the control group. **Table 2** shows that this statistical technique has been used to assess 11 percent

[3] J-Pal website

of the effects in my sub-sample.

Unlike randomized experiments, quasi-experimental evaluations are relatively inexpensive. There is no precise estimate of the cost of non-randomized studies, but Gertler et al., (2011) highlight that impact assessment costs do not exceed, on average, 4.5 percent of the total project cost. Usually, the cost of data collection represents the highest share of the impact assessment cost, an average of 60 percent. Observational studies can rapidly be conducted (between months and years), as they can rely on existing data from national statistics institutions (Gertler et al., 2011).

Table 2: Distribution of impact evaluation methods

Evaluation methods	b	pct
Identification - RCT	431	59.3
Identification - DiD	82	11.3
Identification - Matching	181	24.9
Non-Inferable Econometrics	33	4.5
Observations	727	

3.2.2 Impact assessment methods according to the geographic location of decentralized solar nano projects

In the context of decentralized solar nano projects, **Table 3** reveals that RCTs are mainly conducted in African (34%) and Asian (26%) regions, while only 7 percent of the outcomes generated by decentralized solar nano projects in Asia have been estimated by the Matching strategy. This descriptive evidence supports the potential *Asian cultural bias* due to the major presence of RCT advocacy structures in Asian countries, mostly in India, through the implementation of J-Pal experiments.

Also, the DiD (11%) approach has mostly been used on the African continent. Therefore, Africa is the only location where all causal inference methodologies have been applied to assess project outcomes. Hence, as a robustness check, I focus my empirical strategy on decentralized solar nano projects implemented in Africa. On the other hand, the matching identification mostly took place in Latin America (13%), and the Non-inferable Econometrics strategy only occurs in Asia (4%). Given the heterogeneity of evaluation methods across continents, my empirical model includes a probability weighted factor based on locations.

Table 3: Distribution of impact evaluation methods across locations

Continents	Freq	Pct
Africa		
Identification -RCT	245	33.7
Identification -DiD	82	11.2
Identification -Matching	37	5.0
Non-inferable Econometrics	0	0.0
Total	364	50
Asia		
Identification -RCT	186	25.5
Identification -DiD	0	0.0
Identification -Matching	51	7.0
Non-inferable Econometrics	33	4.5
Total	270	37.1
Latin America		
Identification -RCT	0	0.0
Identification -DiD	0	0.0
Identification -Matching	93	12.7
Non-inferable Econometrics	0	0.0
Total	93	12.7
Total		
Identification -RCT	431	59.2
Identification -DiD	82	11.2
Identification -Matching	181	24.8
Non-inferable Econometrics	33	4.5
Observations	727	

3.2.3 What effects are assessed by the CoSMMA literature?

The CoSMMA has collected a dozen domains of effects, mostly related to the Sustainable Development Goals (SDGs)[4]. In most cases, the measured effects mainly address problems in the educational (28%) and health (23%) sectors, and on energy market (13%) access. These effect typologies are related to the basic needs that decentralized solar nano projects address. Consequently, this paper also provides evidence on the impact of evaluation methods on significant positive outcomes on both education and health challenges.

3.2.4 The outcomes of decentralized solar nano projects

Table 4 describes the four possible outcomes of tested effects on sustainable development, based on the statistical significance test of their estimates:

- Proven-Favorable (P-F)
- Proven-Unfavorable (P-U)
- Unproven-Favorable (U-F)
- Unproven-Unfavorable (U-U)

One can note that only 28 percent of the measured effects have a significant and positive outcome on sustainable development. For instance, these effects are mostly related to: business creation, access to electronic appliances, decrease of respiratory disease prevalence, increase of school attendance. Whereas, 11 percent of effects lead to a significant project failure, which is caused by an "increase of adoption and maintenance costs" of decentralized solar nano projects, among other factors.

Table 4 indicates that most of the tested effects are not statistically significant, whether in terms of favorable (34%) or unfavorable (27%) impact on sustainable development.

[4] See **Appendix 9.3**

Table 4: Distribution of project outcomes

Outcomes of effects (4)		
	Fr	pct
Proven -Favorable	206	28.3
Proven -Unfavorable	77	10.5
Unproven -Favorable	248	34.1
Unproven -Unfavorable	196	26.9
Total	727	100

4 Specification strategy

4.1 Objective of the comparative analysis

The goal of this paper is to contribute to the methodology debate on the influence of experimental and quasi-experimental techniques on the outcomes generated by sustainable projects. This debate has been carried out by development actors under the assumptions driven by the *randomistas*[5], which rely on the scientific rigor for project assessment and monitoring processes, while observational studies can also alleviate the statistical and operational limits of randomization studies through the design of the control group.

Therefore, the analysis of the convergence and the discrepancy between randomized and non-randomized impact assessment approaches allows the development economics field to develop new practices at both the methodological and operational levels. This is a fundamental contribution to the dynamic development economics sectors, such as decentralized electrification.

[5] Designation of development economists who support randomized impact evaluations

Consequently, my statistical specification emphasizes whether the randomized and non-randomized methodologies lead to equivalent or different outcomes of decentralized solar nano projects and therefore identifies best practices for public and private development actors.

4.2 Main specification: the role of indicators and controls

This study relies on the direction of tested effects, which are specified by four possible sustainable development outcomes. To that end, a Multinomial Logit (M-Logit) regression is applied to my main specification. This empirical regression allows me to simultaneously analyze the results of each impact evaluation method on each sustainable development outcome. Moreover, according to Berthélemy and Millien (2018), the application of the M-Logit regression is relevant for studies using categorical dependent and explanatory variables.

As explained in **Section 3**, my specification is limited to a scope of effects tested by measurement tools with standard deviations, with the aim of identifying the influence of each impact assessment method on the probability of evaluating a decentralized solar nano project as a success or failure. Therefore, this work mainly focuses on both the significant positive and negative results, even if the global inclusion of outcomes indirectly provides information on factors that can limit the significant results of projects (Berthélemy, 2019). Therefore, the empirical strategy relies on the following equation:

$$P(outcome_{ip} = k) = \alpha_{ip} + \beta.Methods_{ip} + \rho.Effects_{ip} + \mu.Characteristics_{ip} + \varepsilon_p \qquad (1)$$

Where:

- ***Outcome****:* k corresponds to one of the four categorical outcomes of sustainable development projects
- ***Methods****:* represent the impact evaluation methods (RCT, Matching, and DiD)
- ***Effects****:* nature of measured tested effects
- ***Characteristics****:* vector of project characteristics based on their geographical location, program cost, and governance
- ***β, ρ, μ***: parameters to be estimated
- ***P****:* probability of finding one of the four possible outcomes

- ***p***: solar nano decentralized project index
- ***i***: tested effect index
- **α**: the constant parameter
- ***ε***: error terms clustered, by project

This specification estimates the impact of evaluation strategies on the probability of evaluating significant positive and negative sustainable development project outcomes. The empirical regression takes into account the effects and characteristics of decentralized solar nano projects. As noted by Chauvet and Ehrhart (2018), the standard errors are clustered at the level of aggregation of the variable of interest in macro-micro studies. In this meta-study, given that the impact evaluation techniques are aggregated at the project level, I cluster the standard errors at the same level.

To implement the above model, it is necessary to have a range of explanatory variables at the effect level. These key variables are presented in **Appendix 9.4**, which includes control variables based on the project characteristics and effects.

5 Empirical results

5.1 The impact of evaluation methods on the outcomes of decentralized solar nano projects: A Multinomial-logit strategy

This section discusses the influence of evaluation strategies on the probability of generating different outcomes, focusing on the four columns of interest presented by **Table 5**.

5.1.1 The impact of randomized and non-randomized methods

The findings in **Table 5** describe the influence of impact assessment methodologies on the outcomes of decentralized solar nano projects after controlling for their characteristics. The results represent the average marginal effects (AME) of the probability of generating each of four different outcomes. Estimated AME represent the difference between the probability that a given category will generate an outcome and the probability associated with the reference category, denoted as "ref. = ." (Millien, 2019, p.101). Columns (1) to (4) show the estimated coefficients for each outcome on sustainable development. One can observe that the specification finds 174 proven-favorable observations and 228 unproven-favorable results, as well as 61 significant and negative observations versus 180 non-significant and negative outcomes.

As shown in column (1), the estimates of the DiD (-0.6 pp[6]) and Matching (+21 pp) methods are not significant. Therefore, compared to the experimental method, these two quasi-experimental methodologies have no significant impact on the probability of finding success outcomes from decentralized solar nano projects. Hence, there is no significant difference between RCT and non-RCT strategies on the probability of project success outcomes. Given that the statistical significance of the DiD and Matching techniques converge with the RCT references, quasi-experimental methods find equivalent success outcomes on decentralized solar nano projects as experimental techniques.

These results go beyond the findings of Glazerman et al., (2013), who illustrate a decrease in bias between experimental and quasi-experimental estimators, by purely replicating the same labor program with different impact assessment strategies. Their research reveals a greater decrease in bias when the sample is weighted and when the Matching approach has been applied. Their conclusions are consistent with the main result of my paper. Indeed, one can note the positive sign of the Matching estimate in column (1). Even if this estimate is not significant, the result indicates that the Matching methodology is the closest approach to RCTs for the assessment of off-grid solar nano projects.

Consequently, the findings in column (1) point out that quasi-experimental methods provide approximately the same outcome as do experimental approaches. As detailed in **Section 2**, quasi-experimental methods are less expensive than randomized evaluations

[6] Percentage points

but are also considered by a part of the literature as less rigorous than RCTs. In terms of policy decisions, the results in **Table 5** have major implications for public and private operators in the development field, mostly in the area of small-scale infrastructure projects. As a consequence, the results of this paper find that non-randomized methodologies, which are often less constraining at the operational level than randomized ones, also provide reliable and rigorous project outcomes.

One can also observe that the results in column (1) do not apply for Latin American projects, as they are significant less likely (-13pp) to generate a proven-positive impact on sustainable development than nano solar projects deployed in Africa. The convergence of results between randomized and non-randomized evaluations also applies to projects which have effects on basic electricity access (+48 pp), information and communication access (+27 pp), and energy market features (+23 pp). As a matter of fact, effects on basic electricity access represent de facto one of the main goals of the decentralized electrification sector.

This conclusion also applies when assessors display an independence note (+15.7 pp) regarding funder and practitioner pressures. However, the result regarding the author and publication characteristics appears to be quite counter-intuitive, given that with a non-independent assessment the researcher has a higher chance to receive compensation from operators, and therefore may attest to a positive significant project outcome. Berthélemy and Millien (2018) explain that a large proportion of papers collected by the CoSMMA report favorable effects without any scientific evidence. Thus, it appears that my scientific sub-sample pattern has collected works provided by *genuine assessors*. One can also note that in the energy sector it is more difficult to hide the negative effects of off-grid solar projects, than in other development economics matters such as aid policies. Decentralized solar nano projects are mainly designed at a small-scale level where the importance of technical maintenance increases over the long run.

On the other hand, **Table 5** indicates a discrepancy of results between randomized and non-randomized strategies regarding the failure outcome of decentralized solar nano projects, shown in column (2). Indeed, solar nano projects which have been evaluated by a

DiD approach are a significant 5.6 percentage points less likely to generate a proven-negative outcome on sustainable development than projects assessed by RCTs. Compared to RCTs, the DiD design underestimates the probability of finding nano project failure. Therefore, the DiD technique is significantly less pessimistic than are RCTs. Nevertheless, RCTs have a higher prevalence of assessing project failures.

Moreover, failed projects are more commonly identified in the transformation of the financial sector (+28.5 pp) rather than in education, whereas there is significantly less chance to identify a negative solar nano project outcome on health (-9.3 pp), gender (-10.5 pp), basic access (-10.5 pp), community (-10.4 pp), security (-10.5 pp), and information and communication (-10.5 pp) issues than on education challenges. These findings echo the study of Cameron et al., (2016) which reveals that RCT evaluations tend to be concentrated in the education, health, and information and communication technology sectors. While quasi-experimental studies are more common in agriculture, energy, environment, and private sector development. Thus, there is a cognitive and academic bias regarding the assessment of the SDG effects. Such is the case of the "political site selection bias" of RCTs identified by Corduneanu-Huci et al., (2021).

The likelihood of RCTs to conclude a significant negative impact of decentralized solar nano projects is emphasized when there is transparency regarding the inclusion of the project stakeholders (+23 pp). This result is quite counter-intuitive given that the inclusion of stakeholders in the governance structure of projects usually contributes to their success. Yet my estimate shows that there is a significant risk that projects assessed by an RCT technique will generate a proven-unfavorable impact. This finding might be explained by the analysis of Cook and Campbell (1979), which describes how contextual factors affect experiments, mostly with experimental controls.

The third column in **Table 5** reveals the convergence of impact evaluation methodologies on the unproven-positive outcome of solar nano projects. I find that, compared to the RCT estimates, the DiD and Matching ones are not significant. Hence, as in column (1), there is no discrepancy between non-randomized and randomized methods on the probability of generating an unproven project success. Therefore, randomized and non-randomized methods have the same ability to detect proven and unproven solar nano project successes on sustainable development, mostly when the cost of the project is higher than $100,000 (+16 pp) but also, when the time between a project implementation and the evaluation is high (+3.4 pp), which means that unproven-positive solar nano project outcomes mainly occur in the long term. In terms of policy, sustainable development projects are more likely to generate positive outcomes when the evaluation occurs several years after their implementation. This enables projects to have more time to generate effects and take into account the adoption period of recipients.

Next, column (4) reveals that the Matching evaluation has a lower probability (-30.5 pp) than RCTs to generate an unproven-negative impact of solar nano projects. Thus, compared to RCTs the Matching method is optimistic, given that it tends to underestimate the unproven failure of solar nano projects, as emphasized by the Matching specification of Arraiz and Calero (2015), who find an increase of 4.3 minutes for women hours of unpaid work in the province of Cajamarca (Peru). However, this effect is not significant.

Finally, one can note that publication bias is not an issue in this model, as the estimate for the proxy number of observations is null and non-significant for all the four possible outcomes generated and that publication biases depend mainly on the selection of journal editors, who are less likely to publish negative or null results.

Table 5: **Impact of evaluation methods on the outcomes of decentralized solar nano projects - Average Marginal Effect (AME)**

Explanatory variables	**(1) (Proven-Favorable)**	**(2) (Proven-Unfavorable)**	**(3) (Unproven-Favorable)**	**(4) (Unproven-Unfavorable)**
Proxy of no. of obs.	-0.000+	0.000+	0.000	0.000+
Delay of evaluation (years)	0.008	-0.028	0.034***	-0.014
Eval methods (ref. = RCT)				
Identification - DiD	-0.006	-0.056**	0.074	-0.012
Identification -Matching	0.211	0.205+	-0.110	-0.305**
Domain effects (ref. = Education)				
Income & living conditions (O1)	-0.041	-0.032	-0.017	0.090
Health (O3)	-0.009	-0.093*	0.128	-0.026
Gender (O5)	0.005	-0.105*	-0.024	0.124
Basic access (O7)	0.484***	-0.105*	-0.227**	-0.152+
Economic transformation (O8)	-0.072	0.080	-0.189**	0.181
Community (O11)	-0.206***	-0.104*	0.591***	-0.281***
Security (O16)	0.110	-0.105*	0.017	-0.022
Financial transformation	-0.084+	0.285***	-0.057	-0.144*
Housework	-0.001	0.238	-0.123*	-0.114
Information & communication	0.274**	-0.105*	-0.107	-0.062
Usable time and leisure	0.134	0.043	-0.062	-0.115
Energy (type, costs & faults)	0.232*	0.186	-0.242***	-0.176+
Location (ref. = Africa)				
Asia	-0.018	0.145	0.026	-0.153**
Lat. America	-0.130***	-0.060*	0.068	0.123
Program cost (ref. =≤ $100,000)				
>$100,000	-0.103+	-0.038	0.163***	-0.022
Role of project stakeholders	-0.002	0.233***	-0.225***	0.006
Independence note	0.157***	-0.435***	0.178***	0.100**
Constant	0.305***	0.113***	0.329***	0.253***
Observations	643	643	643	643
Obs. nb outcomes	174	61	228	180
Clusters	Yes	Yes	Yes	Yes

Average Marginal Effect of Multinomial Logit regression. LHS: Proven-Favorable, Proven-Unfavorable, Unproven-Favorable, Unproven-Unfavorable. Subset of 727 observations based on nano solar projects. Ref =: Reference category. Estimates controlled by: Number of observations in evaluation samples (N), Delay of evaluation, Domain of effects, and Project charactesristics. Probability weight: by evaluation methods/continents. Coefficients tell the difference in percentage points from the prediction of referral category. Variance: cluster by id prog2: unique numeric identifier. The variance-covariance matrix is estimated all at once for all four equations. +p <0.10, *p <0.05, **p <0.01, ***p <0.001: significance level occurs only with bootstrap estimator.

5.1.2 Type II error test of the main results

As the estimates of my preferred results are not statistically significant **(Table 5 - column 1)**, I conduct a power test. Several error sources can falsely accept the null hypothesis of my model. This null hypothesis determines the convergence between the impacts of RCT and non-RCT evaluations on the probability for decentralized solar nano projects to generate proven-favorable outcomes. These scientific treatment errors usually come from selection, measurement, and observation biases. Along the same lines, Ravallion (2020) specifies that bias is removed when the treatment status is conditionally exogenous, namely uncorrelated with the error term conditional on the covariates.

As described in **Table 6**, firstly there is the type I error which occurs when there is a rejection of the null hypothesis when it is true. This error is represented by the significance level of alpha (α). On the other hand, the type II error is failing to reject the null hypothesis when it is false. Meta-regressions increase the chances of detect some type II errors in the published papers.

The beta estimate (β) of the type II error represents the likelihood to attest an impact discrepancy between randomized and non-randomized methods, despite the conclusion that the data failed to confirm this status. The standard threshold set by the literature is beta equal to 0.20. Hence, the power of the test can be defined as the probability that one will make the right decision when the null hypothesis is not true, calculated by one minus beta (Daniels and Minot, 2019). According to Pang et al., (1999), the combination of many small studies in a meta-analysis can detect important effects and reduce the possibility of a type II error.

Table 6: Type II error test

Decision of the model (sample)		
Reality of the population	**Accept H0**	**Accept H1**
H0 is true	Correct decision Probability = 1 - α (0.95)	Incorrect decision Probability = α (0.05) **Type I error (false positive)**
H1 is true	Incorrect decision Probability = β (0.2) **Type II error (false negative)**	Correct decision Probability = 1 - β (0.8)

Therefore, I conduct the type II error test for each category of my explanatory variable regarding the proven-favorable outcome in column 1 of **Table 5**. The test is based on the following power equation using the "power" command in STATA (Weinberg and Abramowitz, 2020):

$$poweronemean\ H0\ H1,\ \alpha\ (0.05)\ n\ sd \tag{2}$$

Where:

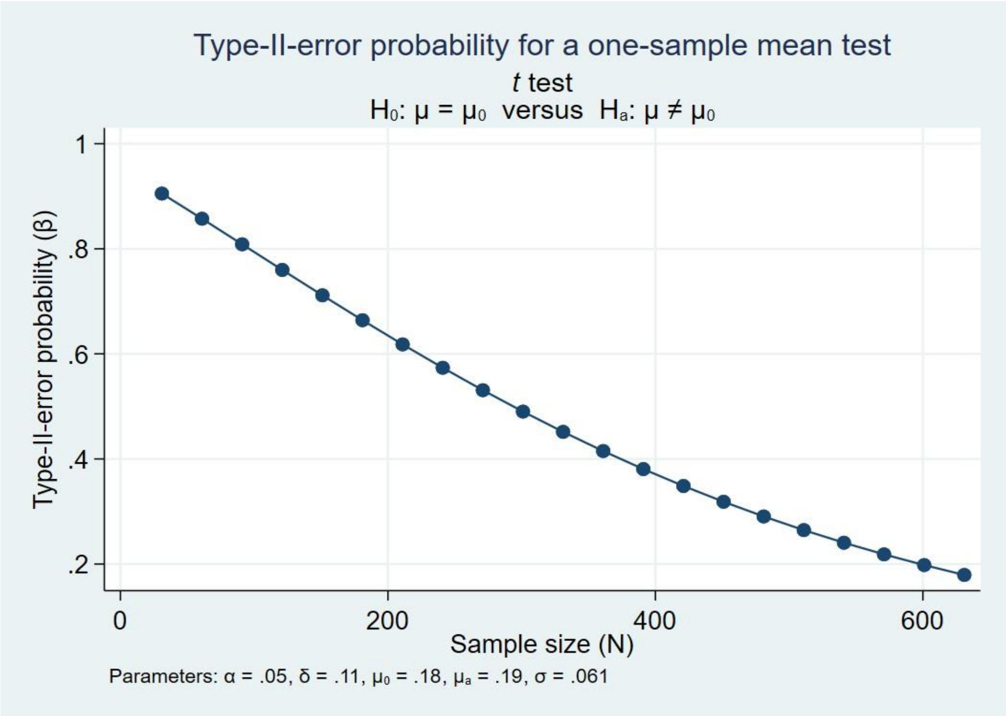
- $H0$ (null hypothesis) = DiD / Matching (convergence of outcomes by RCT and non-RCT methods)
- $H1$ (alternative hypothesis) = DiD / Matching (discrepancy of results by RCT and non-RCT methods)
- α = alpha corresponds to the significance level at 0.05 by default. Represents the probability that there is a convergence between groups but one concludes that there really is a difference
- n = sample of the estimation
- sd = standard deviation

The beta estimate of the DiD method in **Table 7** is equal to 0.172, which is lower than the threshold fixed at 0.20. Consequently, the decision to accept the null hypothesis is correct. This evidence confirms that there is a convergence of impact between RCTs and DiD methodologies on the likelihood to find a proven-favorable effect of decentralized solar nano projects. Moreover, the power graph in **Figure 3** highlights that the estimate reaches the beta threshold of 0.20 when my empirical model contains around 600 observations.

Table 7: Type II error test -DiD

Ho: m = m0 versus Ha: m! = m0								
alpha	**power**	**beta**	**N**	**delta**	**m0**	**ma**	**diff**	**sd**
0.05	0.828	0.172	643	0.115	0.18	0.187	0.007	0.061

Figure 3: Type II error - DiD

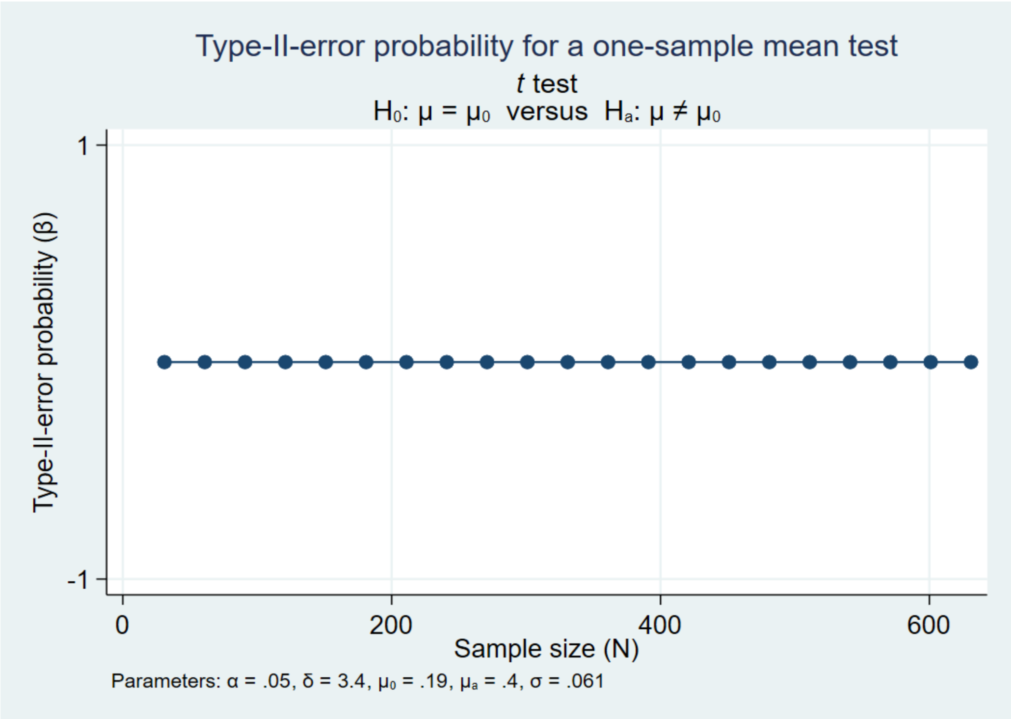


On the other hand, **Table 8** shows that the beta of the Matching estimate is equal to 0. This supports the finding that the results of RCTs and Matching techniques converge. This result is consistent with the graphical representation in **Figure 4**. To that extent, the sample of my main empirical specification is suitable.

Table 8: Type II error test - Matching

Ho: m = m0 versus Ha: m! = m0								
alpha	**power**	**beta**	**N**	**delta**	**m0**	**ma**	**diff**	**sd**
0.05	1	0	643	3.443	0.187	0.397	0.21	0.061

Figure 4: Type II error - Matching



5.2 The influence of impact assessment methods on generated effects of decentralized solar nano projects

The distribution of effects in the sample of decentralized solar nano projects is heterogeneous. As presented in **Table 9**, only education and health access represent more than 20 percent of assessed effects. Indeed, the electrification process is marginally supported by other development purposes such as access to education and health (Berthelemy, 2019). Health and education belong to the United Nations SDG framework and are key domains for the well-being of beneficiaries.

Table 9: Distribution of impact evaluation methods across effects

(**Nature of effects**)	Freq
Identification - RCT	
Health (O3)	22
Education (O4)	24
Total	46
Identification - DiD	
Health (O3)	0
Education (O4)	4
Total	4
Identification - Matching	
Health (O3)	7
Education (O4)	9
Total	16
Total	
Health (O3)	29
Education (O4)	37
N	66

Table 10 describes the impact of evaluation methods on the probability of finding that projects generate a positive and significant outcome on education and health. For each type of effect, a Complementary log-log (C log-log) regression is applied to this sub-sample in order to mitigate constraints based on the number of observations. The results show that there is no discrepancy between the impacts of experimental and quasi-

experimental (DiD and Matching) approaches on the positive outcomes of decentralized solar nano projects in the educational sector. Yet this finding is not statistically significant. Consequently, this result is in line with the findings in **Table 5**. The reason may be related to the fact that education is the only estimated effect which has been assessed by all the different methodologies.

The Matching's evaluations of Latin American projects (+45 pp) over a long-term period (-61 pp) have a higher probability (+27 pp) of generating a proven-positive effect on health challenges. In other words, compared to RCTs the Matching method tends to overestimate the chances of finding effects that address the third SDG: Good Health and Well Being. This result could be correlated with a "loss of memory effect" derived from the time scale between a project's commissioning and its evaluation (-61 pp).

Consequently, assessors should conduct their evaluations over a short-term period in order to find significant outcomes on health, considering that, the further we go forward in time, the less beneficiaries remember the occurrence of effects from decentralized solar nano projects. This finding is consistent with Bédécarrats et al., (2020), who claim that RCTs can only evaluate the short-term impact of causal chains. On the other hand, Ravallion (2009) notes that there is a "myopia bias" in development applications due to the scarcity of long-term impact evaluations, which allow participants to identify project externalities.

Table 10: Impact of evaluation methods on findings of proven-favorable effects

	(1) Education (O4)	(2) Health (O3)
Proxy of no. of obs	0.000	0.000
Delay of evaluation (years)	-0.021	-0.613***
Eval methods (ref. = RCT)		
Identification - DiD	0.307	
Identification - Matching	-0.209	0.272***
Location (ref. = Africa)		
Asia	0.169	0.614
Lat. America	0.417	0.452***
Independence note	-0.197	-1.621+
Observations	**194**	**179**
Clusters:	Yes	Yes

Complementary log-log regression. LHS: education (O4), health (O3). Subset of 206 proven-favorable observations. Ref = Reference category. Probability weight: by evaluation methods/continents. Estimates controlled by: number of observations in evaluation samples (N). Coefficients tell the difference in percentage. Variance: cluster by id prog2: unique numeric identifier of a project; +p<0.10, * p<0.05, ** p<0.01, *** p<0.001: significance level occurs only with bootstrap estimator.

6. Robustness checks

Given that meta-analyses can present some econometric issues, I have performed a set of robustness tests. As highlighted by Labandeira et al., (2020), meta-studies increase the risk of correlation among the paper-effects, mainly due to the use of multiple papers by the same author or by authors from the same institution. In the electricity off-grid sector, one should be aware that authors may be funded by the same source and that projects can be implemented by the same operator.

I include an additional specification strategy based on a logit estimation technique, as well as the restriction of my sample to the African continent.

6.1 The influence of impact methodologies on the success and failure of decentralized solar nano projects: Logit strategy

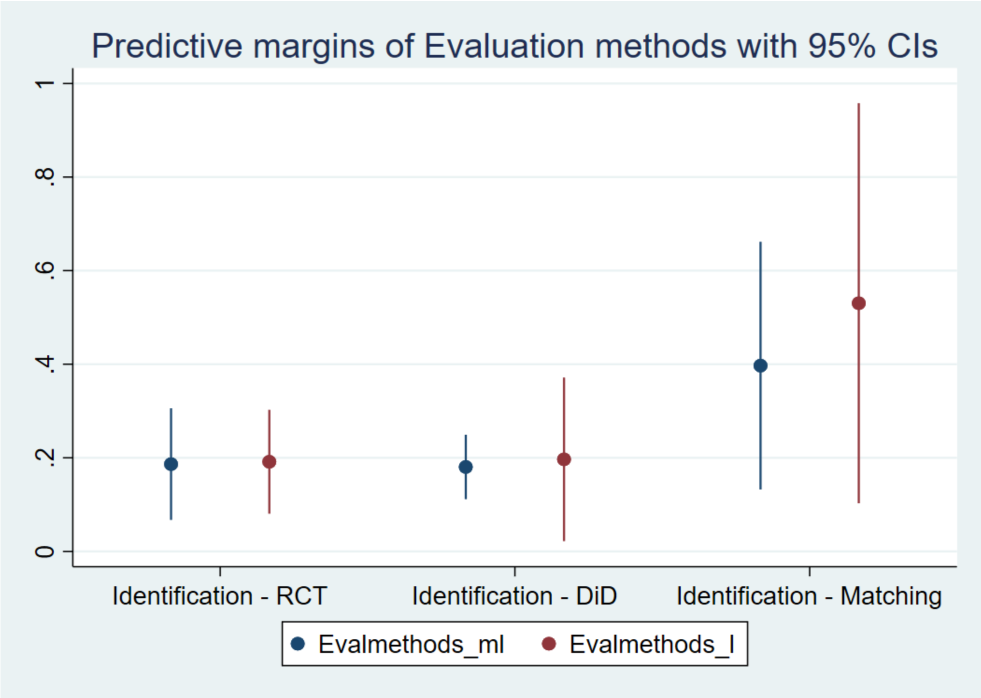
6.1.1 The success of decentralized solar nano projects

Table 11 describes the evolution of the findings in **Section 5** when applying a Logit specification, and removing the nature of effects, location, and program cost control variables. I include information about the type of assessor and the presence of a rural electrification agency for each of the decentralized solar nano projects. In accordance with the estimates of my preferred model, the empirical findings reveal that both experimental and quasi-experimental methods have no significant estimates regarding the likelihood for projects to generate a proven-favorable outcome (column 1). Evaluation techniques demonstrate an equal probability of assessing success of decentralized solar nano projects. This evidence is supported by the fact that assessors from academic entities are more likely (+36.5 pp) to generate proven-positive project outcomes. This result is important because it confirms the prominent role of academic research laboratories in adopting both randomized and non-randomized methodologies in their evaluation strategies.

Moreover, the local governance of projects, through the inclusion of all stakeholders (+41 pp), strengthens the convergence of impact methods on their ability to generate a proven-favorable outcome. For example, Furukawa (2012) uses an original database to conduct RCT and DiD evaluations for the Lighting Africa Program in Kyannamukaaka village (Uganda). In terms of evaluation policy, promoters should include academic assessors from the initial design of the project up to the post-evaluation stage in order to increase the chances of project success, as assessors possess several measurement tools which can provide the same findings.

As a consequence, the results of the two models, M-Logit and Logit, reach similar conclusions. As illustrated in **Figure 5**, the predictive margins of evaluation techniques for proven-favorable outcomes shows that evaluation methodologies behave the same way under the M-Logit (ml) and Logit (l) specifications. In addition, within the evaluation strategies, the Matching approach seems to be the more optimistic one, with higher M-Logit and Logit estimates. Therefore, the Matching methodology over-estimates the probability of project success, even if its estimates are not significant.

Figure 5: Margin effects of evaluation methods



6.1.2 Factors of project failure

To contribute to the knowledge of evaluation methodologies for the literature and for the operational scale of development, it is necessary to define the causes of project failure. Obviously, column (2) in **Table 11** reveals that when the model is limited to proven-unfavorable projects, the impacts of causal inference techniques are diverging.

Compared with RCTs, the DiD strategy is less likely (-3 pp) to generate a proven-negative outcome evaluation of solar nano projects. This finding is interesting as it claims that the DiD procedure is less pessimistic than RCTs in terms of the probability of project failures. Even with a lower coefficient, this result is consistent with the evidence of my preferred model in column 2 of **Table 5**. Disparities among the results of impact methodologies occur mostly in a context where the assessor is not independent (-7 pp) and may receive incentives from the project operator. As argued in **Section 5**, this result is counter-intuitive because the more independent the investigator, the more likely one can prove the failure of a project. In contrast, the Matching approach and RCTs give equivalent outcomes and have a higher chance of finding a proven-unfavorable project outcome than does the DiD approach. Thus, in terms of good practices, assessors and development actors can use quasi-experimental specifications such as DiD (ceteris paribus) to assess the projects they've financed.

Table 11: Impact of evaluation methods on the outcomes of decentralized solar nano projects - Logit model

Explanatory variables	**(1) (Proven-Favorable)**	**(2) (Proven-Unfavorable)**
Proxy of no. of obs	-0.000	0.000
Delay of evaluation (years)	0.198*	-0.021
Eval methods (ref. = RCT)		
Identification - DiD	-0.039	-0.031*
Identification - Matching	0.247	0.080
Type of assessors (ref. = minister or pub agency)		
Development bank	-0.282***	0.123
University/School - research lab	0.365***	0.186***
Rural electrification agency	0.330	-0.053
Role of project stakeholders	0.419***	0.230**
Independence note	0.537***	-0.072*
Observations	**565**	**565**
Clusters:	Yes	Yes

Logit regression. LHS: Proven-Favorable and Proven-Unfavorable. Subset of 727 observations based on nano solar projects. Ref = reference category. Estimates controlled by: number of observations in evaluation samples (N). Coefficients tell the difference in percentage. Variance: cluster by id prog2: unique numeric identifier of a project. Probability weight: by evaluation methods/continents. +p <0.10, * p<0.05, ** p<0.01, *** p<0.001: significance level occurs only with bootstrap estimator.

6.2 The impact of assessment methods on the outcomes of African decentralized solar nano projects

To attest the robustness of my main empirical model presented in **Section 5**, it is interesting to provide findings under different settings. Indeed, policymakers and practitioners want to learn from impact assessment operations and from the performance of measured outcomes in different contexts. Ravallion (2020) argues that one approach to learning from the external validity of projects is to repeat the evaluation in different contexts. Besides, meta-analyses facilitate subgroup analysis and can reveal variation of treatment impact across settings (Pang et al., 1999). Therefore, the aim of this section is to analyze if the estimates of experimental and quasi-experimental approaches hold, even in different geographical circumstances.

Table 3 shows that Africa is the only geographic area where all the evaluation methods have been applied. To that extent, the empirical evidence of this sub-sample analysis is to highlight that my main findings are still robust in a specific geographical context, mainly, if there is an "Asian cultural bias", given that most of the RCT evaluations are conducted in Asia, through the notable presence of the J-Pal research laboratory[7]. For instance, it is easier to perform an RCT design in India than in Burkina-Faso, thanks to financial and human resources. India has the largest production of experimental evaluations, mostly due to the presence of local NGOs. Corduneanu-Huci et al., (2021) attest that the location of J-Pal experiments reflects idiosyncratic organizational selection priorities. Moreover, India is the leading host of RCT experiments in the international development field. Hence, one can expect that my results in **Table 10** correct for this "J-Pal effect".

It turns out that the results in column (1) are robust to the African sub-sample. The estimates of RCT and non-RCT techniques are not significant; consequently, they have the same influence on the success of the African decentralized solar nano projects. This finding is important as it contributes to the debate regarding the limits of RCT studies. Indeed,

[7] **See references**

Section 2 of this paper highlights that one of the major limitations of RCT evaluations is the external validity of results. Often, the natural experiment does not generate the same outcomes according to the scale and the place where projects have been implemented. According to Deaton and Cartwright (2018), RCTs remain valid for a specific evaluation of a particular project. Results in **Table 12** check the external validity of the different outcomes of decentralized solar nano projects gathered in the CoSMMA database. When controlling for the effects, the governance, and the assessment conditions of projects. Thus, this specification is a major contribution to the impact evaluation debate.

This strategy also applies for the probability of evaluating failure of the African decentralized solar nano projects (column 2). One can observe that the DiD methodology still underestimates the probability of generating negative project outcomes compared to the RCT method, with a significant and negative (-14 pp) result. This finding holds when the domain of effects is related to educational matters, whereas the DiD method has a greater chance (+25 pp) of generating an unproven-positive impact of decentralized solar nano projects in Africa (column 3). This may be due to the fact that these evaluations mainly measure the effects on gender (+15 pp), community (+75 pp), and security (+17 pp) issues, while these matters are the least studied by the decentralized electrification literature.

However, in contrast with the main Multinomial-logit model, the African sub-sample indicates the presence of publication bias with small but significant estimates in columns (1), (2), and (4). This may reflect the lack of observations in the African continent compared to the main specification in **Section 5**[8]. Indeed, small sample size increases the risk of identifying publication bias (Pang et al., 1999). In addition, Doucouliagos and Paldam (2007) reveal that researchers are influenced by priors[9] and incentives.

[8] See **Appendix 9.5**

[9] The authors list the five most common priors: polishing, ideology, goodness, author history, and institutional interests.

On the other hand, the African sub-sample significantly lowered the Matching coefficient (-29 pp) regarding the unproven failure of projects. Even if, compared with RCTs, the Matching technique still underestimates the probability of unproven-negative outcome of projects. Therefore, the evidence assessed in this section supports the conclusions from **Section 5** with the illustration that the convergence of successful outcomes assessed through experimental and quasi-experimental methodologies holds even in a subgroup setting, as well as in the use of the DiD technique, which leads to a lower probability of failure of African decentralized solar nano projects than when assessed with RCTs.

Table 12: Impact of evaluation methods on the outcomes of African solar nano projects - AME

Explanatory variables	(1) (Proven-Favorable)	(2) (Proven-Unfavorable)	(3) (Unproven-Favorable)	(4) (Unproven-Unfavorable)
Proxy of no. of obs.	-0.001*	0.000***	-0.000	0.000***
Delay of evaluation (years)	0.031**	-0.016*	-0.003	-0.012
Eval methods (ref. = RCT)				
Identification - DiD	-0.026	-0.141***	0.246**	-0.079
Identification -Matching	0.213	0.149	-0.073	-0.289***
Domain effects (ref. = Education)				
Health (O3)	0.052	-0.197**	0.169	-0.024
Gender (O5)	-0.077	-0.223***	0.148***	0.151***
Basic access (O7)	0.368***	-0.222***	-0.017	-0.129
Economic transformation (O8)	-0.205***	-0.223***	-0.131**	0.559***
Community (O11)	-0.205***	-0.223***	0.746***	-0.319***
Security (O16)	0.034	-0.223***	0.173**	0.016
Housework	0.246	-0.222***	0.172	-0.195
Information & communication	0.262**	-0.223***	0.084	-0.124
Usable time and leisure	-0.084	-0.146	0.243	-0.013
Energy (type, costs & faults)	0.201***	-0.072	-0.066	-0.063
Role of project stakeholders	0.072	0.123***	-0.259***	0.064
Independence note	0.144*	-0.428***	0.100***	0.184***
Observations	364	364	364	364
Obs. nb outcomes	108	40	115	101
Clusters	Yes	Yes	Yes	Yes

Average marginal effect of multinomial logit regression. LHS: Proven-Favorable, Proven-Unfavorable, Unproven-Favorable, Unproven-Unfavorable. Subset of 364 observations based on African nano solar projects. Ref =: Reference category. Estimates controlled by: number of observations in evaluation samples (N), delay of evaluation, domain of effects, and project characteristics. Coefficients tell the difference in percentage points from the prediction of referral category. Variance: cluster by id prog2: unique numeric identifier. The variance-covariance matrix is estimated all at once for all four equations. +p <0.10, * p<0.05, ** p<0.01, *** p<0.001: significance level occurs only with bootstrap estimator.

7 Conclusion and discussion

This paper aims to measure the impacts of evaluation methodologies on the outcomes of development projects. Specifically, the study examines the results of experimental and quasi-experimental methods on the outcomes of decentralized solar nano projects. It describes how different impact evaluation methods influence the probability of finding the success and failure of decentralized solar nano projects in developing countries. Following Corduneanu-Huci et al., (2021), this work goes beyond methodological debates and sectoral approaches in the development literature. It contributes to gaining insight on impact evaluation policies for sustainable development stakeholders.

In the literature, the comparative studies of impact evaluation methods rely mainly on separated comparison models (*between comparison*). For example, Buddelmeyer and Skoufias (2004) find that the RCT and DiD approaches generate similar outcomes as part of the *PROGRESA* program in Mexico.

Several studies have conducted a comparative analysis of methodologies through a replication of an RCT evaluation by re-estimating the control group with a non-RCT strategy (*within comparison*). This strategy has been used by Glazerman et al., (2003) with the comparison of estimators from RCT and non-RCT techniques as part of a welfare program assessment in the United States. My study conducts a simultaneous comparative analysis of the experimental and quasi-experimental methodologies for the evaluation of decentralized solar nano projects by taking into account the project characteristics and their tested effects. In summary, the main findings reveal that RCT and non-RCT evaluations generate a similar impact on the probability of determining success of decentralized solar nano projects. This success mostly concerns the educational sector in developing countries. In addition, experimental and non-experimental methods have the same ability to detect unproven-positive project outcomes. The convergence of results between RCT and Matching methods is also robust at the African continental scale. This evidence echoes the conclusion of Glazerman et al. (2003), who attest that the bias between experimental and quasi-experimental estimators is limited when the matching technique is applied.

This comparative analysis also finds that the influence of methodologies on the determination of project success does not prevail in regard to the probability of determining failure. Indeed, compared with RCTs, the DiD approach is less likely (-3 pp) to find a proven-negative outcome of decentralized solar nano projects. Therefore, in terms of the failure of development projects, quasi-experimental techniques (DiD) are less pessimistic than experimental (RCT) designs. The findings of this paper highlight that project characteristics and effects are key in determining the influence of impact evaluation methods on outcomes. The inclusion of characteristics and *ex-ante* priors can provide protection to researchers in cases where they obtain null project results (Vivalt, 2020).

On the other hand, Callon (2006) indicates that the success of some methods and techniques rely on the common interests of development field stakeholders. Methodological tools should also be approached on a case-by-case basis, according to the prior knowledge available, the intervention design, and the particularities of the settings (Bédécarrats et al., 2020) in liaison with researchers, field operators, donors, and local beneficiaries.

To better choose the appropriate impact evaluation method, it is also important to identify the legal status of the project funders. For instance, RCTs are better suited for private goods such as solar home systems (Hammer, 2017), which are easy to assign across individual households or entities, rather than public goods, for which the benefits are shared across many beneficiaries (i.e., off-grid solar systems in a village or a school). Consequently, impact evaluation policies can stimulate development stakeholders to adopt best practices by alternating between experimental and quasi-experimental methods, based on the project context and characteristics. Ravallion (2020) specifies that methods which yield the most convincing and relevant answers in the context at hand are always the best ones. The planning of a rigorous impact assessment and quality follow-up are also key factors in development project success.

In the years to come, development economics can rely on a range of impact assessment methodologies based on interdisciplinarity and both quantitative and qualitative evidence. As explained by Bédécarrats et al., (2020), the qualitative methods (interviews, focus groups, case studies, beneficiaries' observations, etc.) can serve to contextualize project interventions and study the interactions between different entities. Picciotto (2020) adds that qualitative methods are better suited to determine the reasons for success or failure of effects, as they disentangle design issues and implementation problems.

Moreover, these evaluation tools are designed by the questions of interest for stakeholders and by the assumptions of project interventions (Bamberger et al., 2010). There is also an emerging strand of new impact evaluation strategies based on the combination of observational studies and satellite data (CLUB-ER and FERDI, 2019). As a consequence, the development economics field must be strengthened by including more description, more qualitative data, more big data, and more studies (Morduch, 2020). According to Ravallion (2020), the knowledge gains from an evaluation also bring benefits to future projects, which draw on the lessons learned from prior evaluations. Therefore, meta-analyses on the effects of evaluation tools on sustainable development projects should give decision-makers, developers, and academics confidence that they are implementing the best development policy designs.

Finally, the results of this study should be considered with some caveats since the stability of the model is at stake. For instance, my results are very specific to the energy sector, so their portability might be tricky. Future research should support the innovative aspect of the CoSMMA strategy by replicating the tool for different sustainable development sectors.

8 References

Abramowicz, M., Szafarz, A., 2020. Ethics of RCTs: Should Economists Care about Equipoise? in: Randomized Control Trials in the Field of Development: A Critical Perspective. Oxford University Press, p. DOI: 10.1093/oso/9780198865360.003.0012.

Arraiz, I., Calero, C., 2015. "From Candles to Light: The Impact of Rural Electrification". IDB Working Paper Series.

Bamberger, M., Rao, V., Woolcock, M., 2010. Using Mixed Methods in Monitoring and Evaluation: Experiences from International Development (World Bank Policy Research Working Paper No. No. 5245). World Bank, Washington, DC.

Banerjee, A.V., 2020. "Field Experiments and the Practice of Economics". American Eco-nomic Review 110, 1937-1951.

Banerjee, A., Duflo, E., Kremer, M., 2016. The Influence of Randomized Controlled Trials on Development Economics Research and on Development Policy. Paper prepared for The State of Economics, The State of the World Conference. Proceedings Volume. September 11:42.

Baudet, A., 2019. "Social Impact Assessment Tools for growing, managing and financing social enterprises".

Bédécarrats, F., Guérin, I., Roubaud, F., 2017. "L'étalon-or des évaluations randomisées : économie politique des expérimentations aléatoires dans le domaine du développement". ResearchGate.

Bédécarrats, F., Guérin, I., Roubaud, F., 2020. Editors' Introduction: Controversies around RCT in Development: Epistemology, Ethics, and Politics, in: Randomized Control Trials in the Field of Development: A Critical Perspective. Oxford University Press, p. DOI: 10.1093/oso/9780198865360.003.0001.

Berthélemy, J.-C., 2019. "Challenges of decentralized electrification for economic development: lessons from experience". FERDI.

Berthélemy, J.-C., Millien, A., 2018. "Impact of Decentralized Electrification Projects on Sustainable Development: A Meta-Analysis". FERDI Working paper.

Buddelmeyer, H., Skoufias, E., 2004. "An evaluation of the performance of regression discontinuity design on PROGRESA". (No. WPS3386). The World Bank.

Callon, M., 2006. Pour une sociologie des controverses technologiques, in: Fundamenta Scientiae. https://doi.org/10.4000/books.pressesmines.1196.

Camfield, L., Duvendack, M., 2014. Impact Evaluation -Are We "Off the Gold Standard?". The European Journal of Development Research 26, 1-11.

Cameron, D.B., Mishra, A., Brown, A.N., 2016. The growth of impact evaluation for international development: how much have we learned? Journal of Development Effectiveness 8, 1-21. https://doi.org/10.1080/19439342.2015.1034156.

Chauvet, L., Ehrhart, H., 2018. Aid and growth: evidence from firm-level data. Journal of Development Economics 135, 461–477.

CLUB-ER et FERDI, 2019. Atelier sur l'impact socio-économique de l'électrification. https://ferdi.fr/evenements/atelier-sur-l-impact-socio-economique-de-l-electrification-decentralisee

Cook, T.D., Campbell, D.T., 1979. Quasi-experimentation: design analysis issues for field settings. Houghton Mifflin, Boston.

Corduneanu-Huci, C., Dorsch, M.T., Maarek, P., 2021. The politics of experimentation: Political competition and randomized controlled trials. Journal of Comparative Eco-nomics 49, 1–21. https://doi.org/10.1016/j.jce.2020.09.002.

Daniels, L., Minot, N., 2019. An Introduction to Statistics and Data Analysis Using Stata®: From Research Design to Final Report. SAGE Publications.

Deaton, A., 2009. Instruments of development: Randomization in the tropics, and the search for the elusive keys to economic development. Proceedings of the British Academy, 2008 Lectures 162, 123–160.

Deaton, A., 2020. Introduction: Randomization in the Tropics Revisited, a Theme and Eleven Variations, in: Randomized Control Trials in the Field of Development: A Critical Perspective, Edited by: Florent Bédécarrats, Isabelle Guérin, and François Roubaud. Oxford University Press, p. DOI: 10.1093/oso/9780198865360.003.0002.

Deaton, A., Cartwright, N., 2018. Understanding and misunderstanding randomized controlled trials. Social Science Medicine, Randomized Controlled Trials and Evidence-based Policy: A Multidisciplinary Dialogue 210, 2-21. https://doi.org/10.1016/j.socscimed.2017.12.

Doucouliagos, H., Paldam, M., 2007. The aid effectiveness literature: The sad results of 40 years of research. University of Cambridge, Cambridge Working Papers in Eco-nomics. https://doi.org/10.17863/CAM.5143.

Duflo, E., 2005. "Évaluer l´impact des programmes d´aide au développement : le rôle des évaluations par assignation aléatoire". Revue déconomie du développement Vol. 13, 185-226.

Gertler, P.J., Boyce, S., 2003. An Experiment in Incentive-Based Welfare: The Impact of PROGRESA on Health in Mexico (No. 85), Royal Economic Society Annual Conference 2003, Royal Economic Society Annual Conference 2003. Royal Economic Society.

Gertler, P.J., Martinez, S., Premand, P., Rawlings, L.B., Vermeersch, C.M.J., 2011. "Impact evaluation in practice, World Bank Training Series". The World Bank.

Glazerman, S., Levy, D.M., Myers, D., 2003. "Non-experimental versus Experimental Estimates of Earnings Impacts". The Annals of the American Academy of Political and Social Science 589, 63-93.

Hammer, J., 2017. Randomized control trials for development? Three problems. Brookings. URL https://www.brookings.edu/blog/future-development/2017/05/11/randomized-control-trials-for-development-three-problems/.

Heckman, J.J., Ichimura, H., Todd, P.E., 1997. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme". The Review of Eco-nomic Studies 64, 605-654.

Labandeira, X., Labeaga, J.M., Linares, P., López-Otero, X., 2020. The impacts of energy efficiency policies: Meta-analysis. Energy Policy 147, 111790. https://doi.org/10.1016/j.enpol.2020.

LaLonde, R.J., 1986. "Evaluating the Econometric Evaluations of Training Programs with Experimental Data". The American Economic Review 76, 604-620.

Legovini, A., Di Maro, V., Piza, C., 2015. Impact Evaluation Helps Deliver Development Projects. Policy Research Working Paper No. 7157. World Bank, Washington, DC.

Le Ster, C., 2011. Revue de littérature sur les méthodologies d'évaluation. Entrepreneurs du Monde.

McKenzie, D., 2012. Beyond baseline and follow-up: The case for more T in experiments. Journal of Development Economics 99, 210-221. https://doi.org/10.1016/j.jdeveco.2012.01.002. Millien, A., 2019. Access to electricity and economic development: determinants of favorable impacts for households (These de doctorat). Paris 1. NNT: 2019PA01E038.

Morduch, J., 2020. The Disruptive Power of RCTs, in: Randomized Control Trials in the Field of Development: A Critical Perspective. Oxford University Press, p. DOI: 10.1093/oso/9780198865360.003.0005.

Mugford, M., Kingston, J., Chalmers, I., 1989. Reducing the incidence of infection after caesarean section: implications of prophylaxis with antibiotics for hospital resources. BMJ 299, 1003–1006.

Pamies-Sumner, S., 2014. Les évaluations d'impact dans le domaine du développement - Etat des lieux et nouveaux enjeux. Département de recherche AFD. https://www.afd.fr/fr/ressources/les-evaluations-dimpact-dans-le-domaine-du-developpement-etat-des-lieux-et-nouveaux-enjeux#:~:text=Les%20%C3%A9valuations%20d'impact%2C%20qui,%C3%A9valuation%20des%20projets%20des%20bailleurs.

Pang, F., Drummond, M., Song, F., 1999. The use of meta-analysis in economic evaluation. Working Papers, Working Papers.

Picciotto, R., 2020. Are the "Randomistas" Evaluators?, in: Randomized Control Trials in the Field of Development: A Critical Perspective. Oxford University Press, p. DOI: 10.1093/oso/9780198865360.003.0011.

Pritchett, L., 2020. Randomizing Development: Method or Madness?, in: Randomized Control Trials in the Field of Development: A Critical Perspective. Oxford University Press, p. DOI: 10.1093/oso/9780198865360.003.0004.

Ravallion, M., 2009. Evaluation in the practice of development. The World Bank research observer 24, 29–53. https://openknowledge.worldbank.org/handle/10986/6561.

Ravallion, M., 2020. Should the Randomistas (Continue to) Rule?, in: Randomized Control Trials in the Field of Development: A Critical Perspective. Oxford University Press, p. DOI: 10.1093/oso/9780198865360.003.0003.

Rodrik, D., 2008. The New Development Economics: We Shall Experiment, But How Shall We Learn? Harvard University.

Rosenbaum, P.R., Rubin, D.B., 1983. The Central Role of the Propensity Score in Observational Studies for Causal Effects. Biometrika 70, 41–55. https://doi.org/10.2307/2335942 Review 3ie, 2018. 2018 in review | 3ie.

Stanley, T.D., Jarrell, S.B., 1989. Meta-Regression Analysis: A Quantitative Method of Literature Surveys. Journal of Economic Surveys 3, 161–170. https://doi.org/10.1111/j.1467-6419.1989.tb00064.x.

Simes, R.J., Glasziou, P.P., 1992. Meta-Analysis and Quality of Evidence in the Economic Evaluation of Drug Trials. Pharmacoeconomics 1, 282–292. https://doi.org/10.2165/00019053-199201040-00005.

Teele, D., 2014. Field Experiments and Their Critics: Essays on the Uses and Abuses of Experimentation in the Social Sciences.

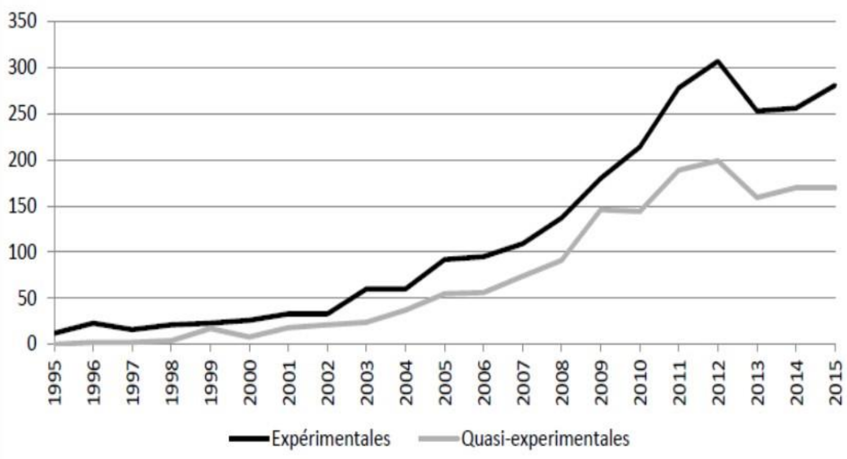
Vivalt, E., 2020. Using Priors in Experimental Design: How Much Are We Leaving on the Table?, in: Randomized Control Trials in the Field of Development: A Critical Perspective. Oxford University Press, p. DOI: 10.1093/oso/9780198865360.003.0013.

Weinberg, S.L., Abramowitz, S.K., 2020. Statistics Using Stata: An Integrative Approach. Cambridge University Press. https://www.povertyactionlab.org/fr/evaluations?sector=All&intervention_type=All&country=All&status=All&region=384&target_group=All&outcome_ of_interest=All

9 Appendices

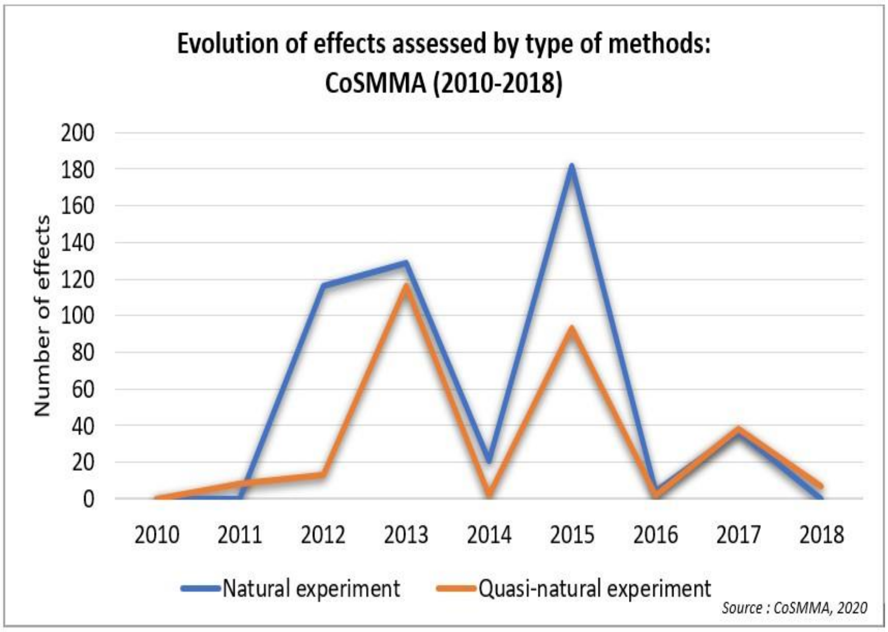
9.1 The evolution of impact evaluation methods in the development field

Figure 2: Evolution du nombre d'évaluations d'impact dans le champ du développement par type de méthode (1995-2015)



Source : 3ie Impact Evaluation Repository, mis à jour en Août 2016.

9.2 The evolution of effects assessed by method type: CoSMMA (2011-2018)



9.3 Domains of effects

Domains of effects	Freq	Pct
Income & living conditions (O1)	21	2.8
Health (O3)	166	22.8
Education (O4)	198	27.2
Gender (O5)	24	3.3
Basic Access (O7)	54	7.4
Economic transformation (O8)	32	4.4
Community (O11)	1	.1
Security (O16)	12	1.6
Financial transformation	5	.6
Housework	34	4.6
Information & communication	35	4.8
Usable time & leisure	51	7.0
Energy (type, costs & faults)	94	12.9
Total	727	100

9.4 Descriptive statistics

Descriptive statistics

Variable	Obs	Mean	Std.Dev.	Min	Max
Eval Methods	727	2.381	1.957	1	7
Proxy of No. of Obs	727	786.197	970.347	45	4000
Delay of evaluation (years)	694	3.304	2.547	1	10
Domain of effects	727	15.818	22.197	1	71
Location	727	3.627	.7	3	5
Program cost	643	1.302	.459	1	2
Role of stakeholders	727	.748	.434	0	1
Independence note	727	.502	.5	0	1
Type of assessors	598	2.246	.78	1	4
Rural electrification agency	727	.688	.464	0	1

9.5 Sample of the studies in the meta-analysis

Table 13 – **Sample size of the main specification and the African sub-sample**

Number of observations in the estimations of the studies					
Subsets	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
Large subset	727	786.1967	970.3469	45	4000
African subset	364	309.6621	260.3395	45	1200