

Making impact with agricultural development projects: The use of innovative machine learning methodology to understand the development aid field

Citation for published version (APA):

Moore, L., van de Laar, M., Wong, P. H., & O'Donoghue, C. (2023). *Making impact with agricultural development projects: The use of innovative machine learning methodology to understand the development aid field*. UNU-MERIT. UNU-MERIT Working Papers No. 011
<https://www.merit.unu.edu/publications/wppdf/2023/wp2023-011.pdf>

Document status and date:

Published: 04/04/2023

Document Version:

Publisher's PDF, also known as Version of record

Document license:

CC BY-NC

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Download date: 23 May. 2024

Working Paper Series**#2023-011**

Making impact with agricultural development projects: The use of innovative machine learning methodology to understand the development aid field

Lindsey Moore, Mindel van de Laar, Pui Hang Wong and Cathal O'Donoghue

Published 3 April 2023

Maastricht Economic and social Research institute on Innovation and Technology (UNU-MERIT)

email: info@merit.unu.edu | website: <http://www.merit.unu.edu>

Boschstraat 24, 6211 AX Maastricht, The Netherlands

Tel: (31) (43) 388 44 00

UNU-MERIT Working Papers

ISSN 1871-9872

**Maastricht Economic and social Research Institute on Innovation and Technology
UNU-MERIT | Maastricht University**

UNU-MERIT Working Papers intend to disseminate preliminary results of research carried out at UNU-MERIT to stimulate discussion on the issues raised.



Making Impact with Agricultural Development Projects: The Use of Innovative Machine Learning Methodology to Understand the Development Aid Field

3 April 2023

Lindsey Moore, Mindel van de Laar, Pui Hang Wong, Cathal O'Donoghue

Abstract

This paper introduces a novel methodology aimed at addressing a critical knowledge gap related to the lack of a systematic understanding of agriculture projects across spatial and temporal dimensions. This gap has impeded efforts to enhance learning and accountability, thereby reducing the overall effectiveness of foreign assistance to the agriculture sector. To address this gap, deductive and inductive methodologies are applied to develop a standardized taxonomy for benchmarking United States Agency for International Development (USAID) agricultural projects. By applying this taxonomy to code all available final evaluations of USAID projects, a large qualitative dataset was generated. This dataset facilitates the analysis of the rich qualitative information available within public project evaluations and covers ninety countries over a span of six decades. The result of this research is a new dataset on the multi-layer composition of development projects, forming the foundation for a machine learning algorithm that expedites the process of synthesizing qualitative evidence and measuring the impact of development aid projects at a systems level. The overarching objective of this research is to contribute to the improvement of project and policy implementation in the field of agriculture development.

JEL codes: C40, F35, O13

Keywords: agricultural projects, development aid, interventions, machine learning, USAID

1. Introduction

The current body of research on individual agricultural development projects is extensive. However, the existing evidence largely consists of case studies from a single project or multiple projects in the same country, often studying only a short period of time. The absence of a systematic comprehension of projects has undermined efforts to enhance learning beyond micro-level projects and does not hold the institutions accountable for their work, thereby impacting the efficacy of foreign assistance to the agriculture sector overall (Fløgstad & Hagen, 2017; Oliver et al., 2014; Strydom et al., 2010; Takaaki et al., 2017; Tierney et al., 2011). The central challenge is the absence of a standardized taxonomy for benchmarking projects making it difficult to compare the evidence of project success across sectors, institutions, and countries (Belcher and Palenberg, 2018). Consequently, the development community lacks critical data and tools that make the application of data accessible and easily understandable.

To address this critical knowledge gap, this paper introduces a novel methodology that enables quantitative analyses of the rich qualitative information that is available within the public project evaluations. It is based on the creation of a comprehensive taxonomy that enables the comparison of project interventions and outcomes across spatial and temporal dimensions. Project evaluations serve as the primary data source as they are one of the most frequently used and most informative sources of development data for practitioners within the international development sector (Takaaki et al., 2017). The result of this research is a new dataset on the multi-layer composition of development projects. The research forms the foundation for a machine learning algorithm that expedites the process of synthesizing qualitative evidence and measuring impact of development aid projects at a systems level.

Efforts to advance aid effectiveness research agenda in recent years have focused primarily on expanding access to data and enhancing the quality of data related to the determinants and mechanisms of the development assistance (Addison et al., 2005; Alesina & Dollar, 2000; Boone, 1996; Headey, 2008). Although thought-provoking research inquiries and complex models can impact scholarship and policymaking, their conclusions may be misguided and prone to error if the foundational data on foreign aid flows are deficient or incomplete. Despite efforts to improve the quality and availability of data on the determinants and mechanisms of development assistance, few researchers have made contributions to enhance the breadth and depth of data related to the fundamental dimensions of aid allocation.

Tierney and colleagues (2011) have contributed to the overall comprehension of the basic aspects of aid allocation by utilizing the Creditor Reporting System to track aid flows reported by OECD member countries. Their dataset, AidData, specifies the financial flows and specific projects from 42 bilateral donors and 44 multilateral donors. Similarly, Honig et al. (2022) constructed a dataset based on success ratings made by donor staff and independent evaluators to evaluate the degree to which projects encourage the adoption of "Access to Information" and efficiently allocate resources. In another study, Denizer and colleagues (2003) examined 6,000 evaluations of World Bank projects over a period of 26 years to determine the project-level factors that contribute to success. Furthermore, Bulman, Kolkma, and Kraay (2017) evaluated 5,000 assessments of World Bank and

Asia Development Bank projects over a span of 30 years. However, none of these datasets include data at the intervention level, and all omit data from the largest bilateral donor, USAID. Consequently, an important literature gap exists in providing a fundamental understanding of the nature and composition of agricultural projects at the intervention level. Without this basic understanding of project characteristics, answering more complex questions surrounding the specific determinants and effects of foreign assistance is impossible.

In the realm of machine learning, limited research has been conducted to explore its potential in gaining insights into the complex landscape of development interventions. Toetzke et al. (2022) employed machine learning techniques to analyze textual descriptions of aid activities in developing countries, as reported by the Creditor Reporting System. Ricciardi (2020) conducted a literature review using machine learning to map existing research on on-farm interventions that improve the incomes or yields of small-scale farmers in water-scarce regions. Porciello (2020) conducted a machine learning-assisted review of text summaries from agriculture research. However, no attempts have been made to use machine learning to extract, label, and rate text from full evaluation reports.

In this research, we develop a methodology that contributes to the development communities' ability to leverage data from a vast number of reports to enhance project and policy implementation. Before doing so, it is imperative to establish a clear and concise understanding of the terminology and concepts that will be utilized throughout the analysis. The definitions of interventions and outcomes utilized in this study are adapted from the Office of the Director of U.S. Foreign Assistance of the State Department and USAID (2009: 6). Specifically, intervention refers to "an action or entity that is introduced into a system to achieve some result." An outcome is defined as "the results or effect that is caused by or attributable to the project, program or policy" (Office of the Director of U.S. Foreign Assistance of the State Department and USAID 2009: 8).

In Section 2, the sampling strategy for the inclusion of the projects in this study is explained. In Section 3, we present the taxonomic approach that will be used to categorize interventions and outcomes as well as the coding methodology of the qualitative evaluation reports. In Section 4 we present the data set and descriptive outcomes of the coding effort. Section 5 discusses the limitation and concludes.

2. Sampling Strategy

As the largest resource for USAID-funded technical and project materials, the Development Experience Clearinghouse (DEC) is an online database mandated to store all data collected by USAID and USAID's contractors (USAID 2015). USAID's Evaluation Policy (2016b) stipulates that all quantitative data collected by USAID or one of the Agency's contractors or for an evaluation must be uploaded and stored in the DEC (USAID, 2015). Such data could include many different types of materials such as text, images, video, audio, maps, charts, and raw data.

This research employs a sampling strategy that draws from a subset of the DEC database, encompassing all countries and years. Our sampling frame consists of all projects that have final

evaluation documents available and offer detailed information at both the intervention and outcome levels. For this study, we focus on the agricultural sector. This results in the selection of only those documents within the DEC database related to the agriculture sector as defined by the 2000 Famine Prevention and Freedom from Hunger Improvement Act and cited in USAID's Agriculture Strategy (2004: 1) as "the science and practice of activities related to production, processing, marketing, distribution, utilization, and trade of food, feed, and fiber." All evaluation documents within the agriculture selection were extracted through the DEC Application Programming Interface (API) using the search terms detailed in Table 1. Based on this sampling strategy, a total of 446 agriculture final evaluations was obtained

Table 1: Inclusion Criteria

Category	Search Strategy	Range and Keywords
Countries	All available	Afghanistan, Albania, Angola, Armenia, Thailand, Azerbaijan, Bangladesh, Belarus, Benin, Bosnia and Herzegovina, Botswana, Brazil, Burkina Faso, Burma, Burundi, Cambodia, Cameroon, Central African Republic, Chad, Columbia, Cote d'Ivoire, Cuba, Cyprus, Democratic Republic of Congo, Djibouti, Dominican Republic, Ecuador, Egypt, El Salvador, Eswatini, Ethiopia, Georgia, Ghana, Guatemala, Guinea, Guyana, Haiti, Honduras, India, Indonesia, Iraq, Jamaica, Kazakhstan, Kenya, Kosovo, Kyrgyz Republic, Laos, Lebanon, Lesotho, Liberia, Libya, Madagascar, Malawi, Maldives, Mali, Mauritania, Mexico, Moldova, Mongolia, Montenegro, Morocco, Mozambique, Namibia, Nepal, Nicaragua, Niger, Nigeria, North Macedonia, Pacific Islands, Pakistan, Panama, Paraguay, Peru, Philippines, Republic of the Congo, Rwanda, Senegal, Serbia, Sierra Leone, Somalia, South Sudan, Sri Lanka, Sudan, Syria, Tajikistan, Tanzania, Thailand, The Gambia, Timor-Leste, Tunisia, Turkmenistan, Uganda, Ukraine, Uzbekistan, Venezuela, Vietnam, West Bank Gaza, Zambia, Zimbabwe
Years	All available	1970-2020
Search Terms	Limited to Agriculture Sector	Agriculture (General), Agricultural policy, Agricultural markets, Agricultural management, Agricultural finance, Agricultural enterprises and companies, Agricultural education, Agricultural economics, Agricultural development, Animal husbandry, Animal nutrition and health, Aquacultures and fisheries, Agribusiness, Agricultural technology, Agricultural research, Fertilizers, Farming systems, Crop protection, Crop production, Crop pests and control, Crop diseases and control, Cash crops, Sustainable agriculture, Soil Sciences and Research, Livestock, Irrigated farming, Food supply, Food security, food crops, and Plant breeding, seeds and physiolog

To understand the representativeness of our sample, we compared the basic characteristics of our sample to USAID's total project base within agriculture for the period 2010-2020.¹ A total of 705 agriculture projects implemented during this period were identified, including the 183 evaluated projects that are sampled (see Figure 1). We, therefore, estimate that approximately 30% of all implemented agricultural projects are included in the sample for this research.

Table 2: Sample and Population Characteristics

Category	Sample (Evaluated Projects)	Population ² (All Projects)
Average Project Length	Six Years	Six Years
Average Project Budget	\$20,000,000	\$10,000,000
Number of Countries	90	90

As can be seen in the data in Table 2, the sample of projects is different from the population in terms of budget. While the average project length and geographical coverage are similar, the average budget size of the sample is approximately \$20 million while the average budget size of the population is approximately \$10 million. This difference is likely because larger projects are more often evaluated.

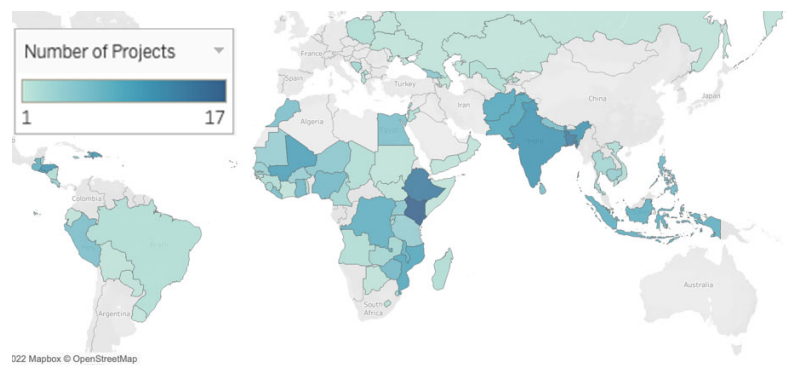


Figure 1: Sample Size by Country

¹ Determining the population size of agriculture projects conducted by USAID since its establishment in 1961 presented unexpected challenges. This is primarily due to the absence of a disaggregated record of USAID projects by sector before 2000 (Department of State, USAID, 2022). Consequently, the population size of agriculture projects is approximated by leveraging data on the obligation of funding managed by USAID in the agriculture sector from the U.S. Government's foreign assistance database from 2010 onward (U.S. Government, 2023). Each entry within this dataset underwent manual scrutiny to remove incorrectly tagged projects that did not fit the definition of agriculture. Given the possibility of incomplete data for the years 2000-2009, which could result in an underestimation of the population size, calculations were based on the period 2010-2020.

² Outliers, as defined by projects above \$400,000, were removed from the average calculation. In total, two projects from Afghanistan were removed.

While USAID's Evaluation Policy (2016b) stipulates that all quantitative data collected by USAID or one of the Agency's contractors must be uploaded and stored in the DEC, not all projects implemented have an evaluation in the DEC. This is either because the project was not evaluated, or because the project was evaluated but the evaluation was not uploaded.

It is however important to be mindful that small projects are being underrepresented in our sample

3. Taxonomical Approach based on Qualitative Data Coding

This study aims to generate a quantitative data set from the rich qualitative information available in evaluations to enable a comparison of different approaches taken in agriculture development across projects, countries, and years. Our methodological approach has two stages. In the first stage, we create a project taxonomy both inductively and deductively based on the text data extracted from the evaluation documents from the DEC database. In the second stage, we apply the taxonomy to projects to determine if the text of a project described an intervention, an outcome, or both.

3.1. Taxonomy

To achieve a comprehensive representation of all the various interventions and outcomes in agricultural projects, we employed a taxonomic categorization approach to disaggregate the data into quantitative data points. This approach provides a granular understanding of the projects, which has not been previously available.

The taxonomy was built using both inductive and deductive methodologies. Deductively, we adopted the variables of interest from a pre-existing list of interventions developed from a USAID evaluation synthesis based on 200 evaluations conducted between 2010 and 2015 across 64 countries (USAID, 2016a). During the coding process, it was discovered that the original list adopted from the USAID synthesis was not sufficiently comprehensive or nuanced. For example, many codes such as 'economic growth' were not specific enough to be classified as an intervention. Thus, inductively, through an iterative coding process described in Section 3.2, we identified and grouped interventions based on their similarities and added interventions that were not included on the initial USAID list. The final taxonomy catalogues the universe of all USAID agriculture interventions and outcomes including clear definitions for each intervention.

The resulting taxonomy consists of three hierarchical levels of codes, which span from the most general to the most specific. First-order codes, at the highest level, offer a general categorization of intervention types, as exemplified by 'value chain intervention' in Figure 2. Conversely, third-order codes, at the lowest level, provide the most detailed description of the intervention, such as 'vertical linkage' in the same example.



Figure 2: Example Code Orders

3.2. Coding Methodology

The application of the taxonomic categories to project documentation facilitates the identification and standardization of interventions and outcomes across projects. The research team was composed of a multidisciplinary group of technical experts in the field of agriculture, including both USAID technical experts and Ph.D. students. The coding process included 3 steps, namely 1) identification of excerpts, 2) labeling the excerpt, and 3) assigning a rating based on the extent to which the intervention achieved the described outcome.

The first step of the coding process involved identifying relevant text excerpts that described either an intervention, an outcome, or both. Each evaluation report is extensive, with an average length of 150 pages, including a seven-page Executive Summary. To identify the relevant excerpts, the researchers reviewed their assigned reports' Executive Summary and applied three-point criteria to identify relevant text. The criteria for text selection included: 1) text detailing the degree to which an intervention achieved a specific outcome; 2) text explaining the interventions implemented by the project; and 3) text about an outcome without explicitly attributing that success to a particular intervention. The researchers then thoroughly reviewed the report's body to ensure all interventions and outcomes were identified and labelled.

Secondly, the excerpts were labelled with the applicable codes from the taxonomy. To ensure the validity of the coding, each report was independently coded by two researchers and then reviewed by a third researcher for consistency. Any inconsistencies in the coding were discussed among the researchers to arrive at a consensus regarding the correct code. When an intervention or outcome was identified in an evaluation but not listed in the coding taxonomy, it was noted. If the same variable appeared in two additional reports, it was adopted as an official variable in the taxonomy. This process enabled the research team to continually improve the taxonomy to capture all interventions and outcomes implemented across projects historically.

Thirdly, the excerpts were assessed to rate the degree to which the intervention achieved the outcome. Consistent with the literature on the effectiveness of foreign aid, the researchers employed a Likert scale ranging from 1 to 4 to gauge the degree to which the intervention met its

intended outcome. Prior research has employed comparable approaches to evaluating project performance, such as overall success ratings that are assessed by both donor personnel and independent evaluators post hoc (Denizer et al., 2013; Dreher et al., 2021; Honig, 2014; Honig & Gulrajani, 2018). The ratings adhere to several widely accepted evaluation criteria of the OECD which capture two primary aspects of performance: 1) the attainment of project objectives and 2) the effectiveness, or efficiency, with which project resources are utilized (Honig et al., 2022). Specifically, the ratings were defined as follows:

1. A rating of 1 indicated that the intervention had little or no positive effect or mainly a detrimental effect on the stated objectives.
2. A rating of 2 indicated that the intervention contributed in a small way to achieving the objective with a minor negative effect on the stated objectives
3. A rating of 3 denoted a positive impact, indicating that the intervention had a largely positive impact on the project's objectives.
4. A rating of 4 signified that the intervention exceeded expectations by having a significant and measurable effect on the project's objectives, fully achieving or surpassing them.

The final data set was used to train a machine learning algorithm as a third verification method to reduce the potential for human error and bias in the coding process. By feeding the algorithm data that had been labeled with the correct taxonomical terms, it was trained to accurately classify and label different interventions and outcomes. To test the accuracy, a subset of the final data set was reserved for validation purposes. The algorithm was then applied to this subset of data, and the predicted classifications were compared to the true classifications to measure the accuracy.

4. Descriptive Outcomes

The final dataset comprises approximately 6,000 rows of data. Figure 3 provides a snapshot of the final data set. Each row includes details on the project name, extracted excerpt, intervention labels, implementing organization name, project start and end date, outcome codes (where relevant), and the numerical ranking of intervention success in achieving the outcome. For example, the first row of the dataset pertains to the Accelerating Sustainable Agriculture Program (ASAP) in Afghanistan implemented by Chemonics from 2006 to 2011 in Afghanistan. The extracted excerpt in the first row states "Efforts to upgrade farm-level production technologies involved numerous field demonstrations and trainings. Impact studies indicated greater revenues from these new inputs and technologies" (Wazir et. al, 2010: 2). The excerpt was labelled with the intervention labels 'extension services', 'inputs', and 'technology', and rated a '3' for the outcome 'increased revenue generation' as it states that impact studies found farmers to have achieved greater revenues owing to these interventions.

Project Name	Excerpt	Interventions	Country	Implementer	Start Date	End Date	Increased Exports	Increased Revenue or Income Generation
Accelerating Sustainable /	Efforts to upgrade farm-level	Extension Services	Afghanistan	Chemonics International	2006	2011		3
Accelerating Sustainable /	Six agribusinesses were supported	Value Chain Interventions	Afghanistan	Chemonics International	2006	2011		
Accelerating Sustainable /	The effort to upgrade and improve	Value Chain Interventions	Afghanistan	Chemonics International	2006	2011	4	4
Accelerating Sustainable /	The establishment of AgDe	Value Chain Interventions	Afghanistan	Chemonics International	2006	2011		4
Accelerating Sustainable /	The intent was to continue	Value Chain Interventions	Afghanistan	Chemonics International	2006	2011		
Access to Mechanization f	Achieve project objectives	Extension Services	Georgia	Citizen's Network Foundation	2009	2012		
Access to Mechanization f	AMP's contribution to rural	Extension Services	Georgia	Citizen's Network Foundation	2009	2012		4
Access to Mechanization f	Data gathered from FGDs in	Extension Services	Georgia	Citizen's Network Foundation	2009	2012		4
Adaptive Symbiotic Techn	The seed treatment innovation	Agriculture Technology	India	Adaptive Symbiotic Technologies	2020	2020		
Advanced Marketing and	Appropriate systems or products	Value Chain Interventions	Egypt	ACDI/VOCA	2015	2018	1	
Advanced Marketing and	Catalyze market-oriented	Agriculture Technology	Egypt	ACDI/VOCA	2015	2018		
Advanced Marketing and	Conclusion 1.1: The AMAL	Agriculture Technology	Egypt	ACDI/VOCA	2015	2018		3
Advanced Marketing and	Conclusion 1.2: Both the	Value Chain Interventions	Egypt	ACDI/VOCA	2015	2018	1	
Advanced Marketing and	Conclusion 2.1: The GDA r	Business Services	Egypt	ACDI/VOCA	2015	2018		
Advanced Marketing and	Conclusion 3.1: The lack o	Human and Institutional	Egypt	ACDI/VOCA	2015	2018		
Advanced Marketing and	Conclusion 4.3: Training a	Extension Services	Egypt	ACDI/VOCA	2015	2018		3
Advanced Marketing and	The GDA between USAID a	Value Chain Interventions	Egypt	ACDI/VOCA	2015	2018	3	3

Figure 3: Example of Final Dataset

Note: For enhanced visualization, this table presents only two of the 33 outcome variables, namely "increased exports" and "increased revenue or income generation."

The 6,000 rows of data, taken as a corpus, offer the first systematic view of USAID projects at the intervention level, which lends itself to various forms of analysis. For instance, as shown in Figure 4, the evidence can be used to visualize trends in the popularity of interventions over time. To calculate these trends, a panel data set was created by computing the number of projects that used each intervention per year between 1975 and 2016.³ The prevalence of each intervention per year was calculated as the ratio of projects that used the intervention to the total projects implemented that year, representing the intervention's popularity in the overall agriculture portfolio. For example, in 1991, 73 projects employed the 'strengthening government Institutions' intervention, while the total number of projects implemented that year was 95. Therefore, the prevalence of 'strengthening government institutions' intervention in the agriculture portfolio was calculated to be 76%, indicating that most agriculture projects that year included a component working with government institutions. The resulting panel dataset aggregates the prevalence of each intervention over time, cataloguing the multi-layered composition of agriculture approaches in 457 development projects across six decades (1970-2015) in 65 countries.

³ The start date of 1970 was selected as it is when the first USAID evaluation is available, while the end date of 2016 was chosen as the year with all evaluations likely published, given the average six-year time lag between project inception and evaluation publication.

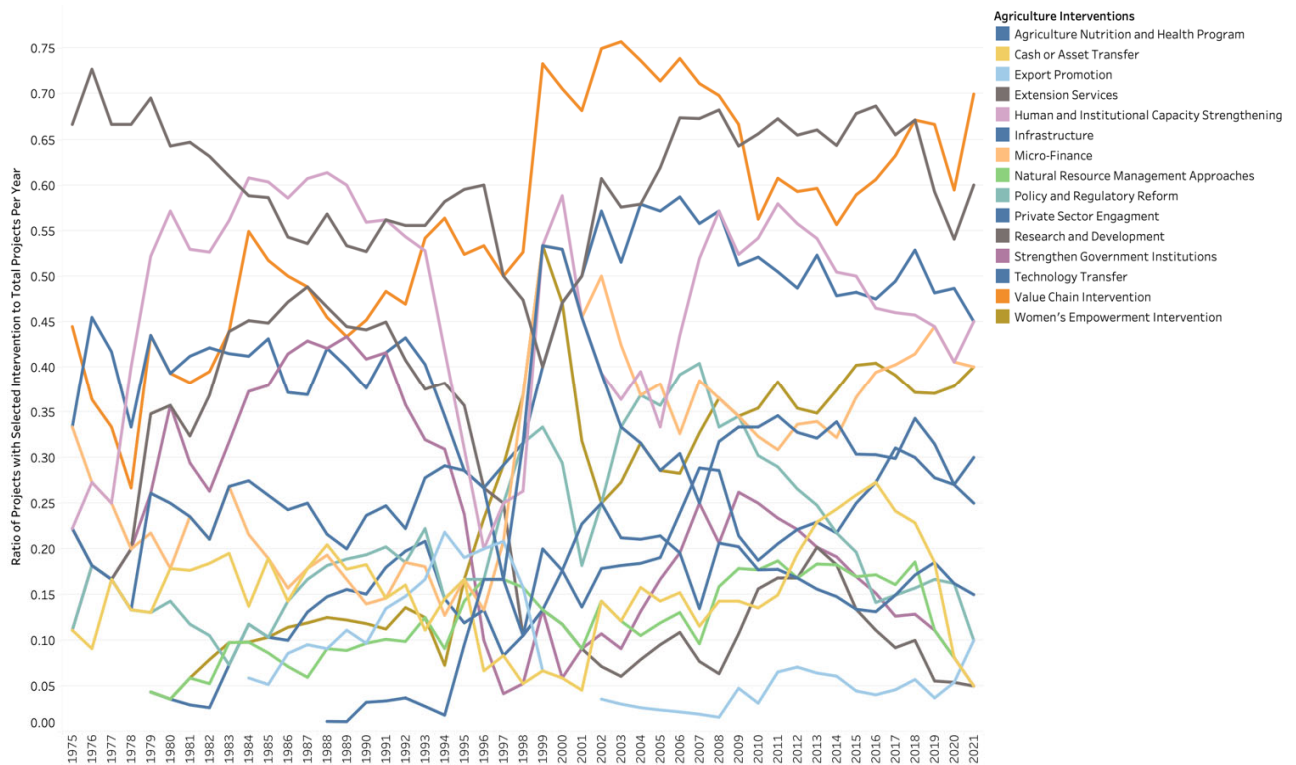


Figure 4: Popularity of Interventions Over Time

Overall, a comprehensive understanding of the trends in the popularity of interventions over time can provide valuable insights to inform foreign assistance strategies. For example, as depicted in Figure 5, the popularity of the 'Research and Development' intervention appears to have declined markedly starting around 1987. This could be related to the Bumpers Amendment in 1986 which, in the words of Bumper was to "prevent American tax dollars from being used to help foreign countries who are trying to take our export markets" (Barjon, 2011: 19). The Bumpers Amendment was the response to a protest by the American Soybeans Association (ASA) over the USAID research project INTSOY, which was developing soybean varieties in competing countries such as Argentina and Brazil (Barjon, 2011). ASA, with the help of Senator Bumper, demanded the termination of research and technical assistance to foreign nations that compete with the USA and asked for a redirection of research funds (Manicad, 1995). This example highlights the importance of a systematic understanding of intervention-level trends in providing deeper insight into the factors driving project and policy design. Such insights can lead to a more informed understanding of the drivers of aid effectiveness, improving decision-making in the field of international development.

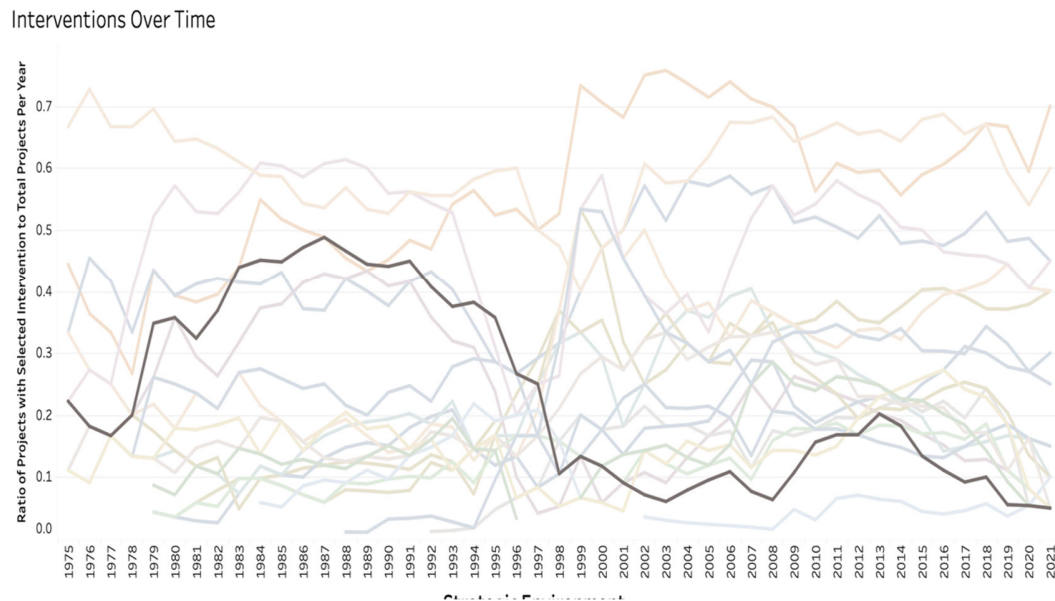


Figure 5: Research and Development Over Time

5. Conclusions and Limitations

The current body of research on individual agricultural development projects has been limited by a lack of a standardized taxonomy for benchmarking, hindering the systematic scrutiny of agricultural development projects across temporal or spatial dimensions. To address this gap, this paper introduced a novel methodology for the quantitative analysis of rich qualitative information within project evaluations based on the creation of a comprehensive taxonomy that enables the comparison of project interventions and outcomes across spatial and temporal dimensions. The result of this research is a new dataset on the multi-layer composition of development projects that serves as the foundation for a machine learning algorithm that expedites the process of synthesizing qualitative evidence and measuring impact at a systems level. The model can be used to automatically identify and classify interventions and outcomes in new documents, with the results compared to the existing taxonomy for accuracy. This new approach addresses critical knowledge gaps and provides a valuable tool for the development community to enhance project and policy implementation.

The methodology employed in the present research has certain limitations that must be acknowledged. Firstly, the identification of variables and assignment of weights on a Likert scale from one to four is to some extent subjective and may be influenced by the individual researcher's perspective. To mitigate this subjectivity, the research methodology employs experts in agriculture and international development to code the reports and employs a clear taxonomy with precise

definitions. Additionally, multiple reviewers are tasked with revising the same report to further enhance the reliability of the coding process.

Secondly, evaluations are often conducted by independent evaluators, but it is possible that the authors of evaluations may be inclined to overstate the successes of a project and underplay its weaknesses or failures to secure future funding from donors. As a result, the coding of interventions may be skewed towards the positive side. Nevertheless, the research methodology attempts to control for this potential bias by considering the ratings as a relative measure of the impact of an intervention compared to that of another intervention, rather than an absolute measure. Notably, the research has found that while a positive bias may exist in summary statements, reporting at the individual intervention level is frequently more critical.

This new approach provides a valuable tool for the development community to enhance project and policy implementation, not only in the agriculture sector but potentially in other sectors as well. Further research can expand the methodology to new sectors, enabling a deeper understanding of the determinants and effects of development assistance. Overall, this study contributes to a more rigorous and systematic evaluation of development projects, providing a foundation for evidence-based decision-making in the field of international development.

References

- Addison, T., Mavrotas, G., & McGillivray, M. (2005). Development assistance and development finance: Evidence and global policy agendas. *Journal of International Development*, 17(6), 819–836. <https://doi.org/10.1002/jid.1243>
- Alesina, A., & Dollar, D. (2000). Who gives foreign aid to whom and why? *Journal of Economic Growth*, 5(1), 33–63. <https://doi.org/10.1023/A:1009874203400>
- Barjon, R. (2011). Effectiveness of Aid in Haiti and How Private Investment Can Facilitate the Reconstruction. Statement to the U.S Senate Subcommittees of Foreign Relations on International Development and Foreign Assistance and Western Hemisphere. <https://www.foreign.senate.gov/imo/media/doc/Simon-Barjon%20testimony.pdf>
- Boone, P. (1996). Politics and the effectiveness of foreign aid. *European Economic Review*, 40(2), 289–329. [https://doi.org/10.1016/0014-2921\(95\)00127-1](https://doi.org/10.1016/0014-2921(95)00127-1)
- Bulman, D., Kolkma, W., & Kraay, A. (2017). Good countries or good projects? Comparing macro and micro correlates of World Bank and Asian Development Bank project performance. *Review of International Organizations*, 12, 335–363. <https://doi.org/10.1007/s11558-016-9256-x>
- Department of State, USAID. (2022). Methodology, Foreignassistance.gov. <https://www.foreignassistance.gov/about#tab-methodology>
- Denizer, C., Kaufmann, D., & Kraay, A. (2013). Good countries or good projects? Macro and micro correlates of World Bank project performance. *Journal of Development Economics*, 105, 288–302. <https://doi.org/10.1016/j.jdevco.2013.06.003>
- Dreher, A., Fuchs, A., Parks, B., Strange, A., & Tierney, M. J. (2021). Aid, China, and growth: Evidence from a new global development finance dataset. *American Economic Journal: Economic Policy*, 13(2), 135–174. <https://doi.org/10.1257/pol.20180631>
- Fløgstad, C., & Hagen, R. J. (2017). Aid dispersion: Measurement in principle and practice. *World Development*, 97, 232–250. <https://doi.org/10.1016/j.worlddev.2017.04.022>
- ForeignAssistance.gov. U.S. Government. <https://www.foreignassistance.gov/>. Accessed: March 24, 2023.
- Headey, D. (2008). Geopolitics and the effect of foreign aid on economic growth: 1970–2001. *Journal of International Development*, 20(2), 161–180. <https://doi.org/10.1002/jid.1395>
- Honig, D. 2014. More Autonomy for Donor Organizations and Their Agents (Sometimes): Bringing Organizational Behavior and Management Theory to Foreign aid Delivery. New Delhi: Global Development Network.

Honig, D., & Gulrajani, N. (2018). Making good on donors' desire to do development differently. *Third World Quarterly*, 39(1), 68-84. <https://doi.org/10.1080/01436597.2017.1369030>

Honig, D., Lall, R., & Parks, B. C. (2022) When does transparency improve institutional performance? Evidence from 20,000 projects in 183 countries. *American Journal of Political Science*. E-pub ahead of print. <https://doi.org/10.1111/ajps.12698>

Manicad, G. (1995) Agricultural biotechnology projects within USAID. *Biotechnology and Development Monitor*, n.24, p.8-10.

Office of the Director of U.S. Foreign Assistance of the State Department and USAID (2009). Glossary of Evaluation Terms. Washington DC. https://pdf.usaid.gov/pdf_docs/Pnado820.pdf

Oliver, K., Innvar, S., Lorenc, T., Woodman, J., & Thomas, J. (2014). A systematic review of barriers to and facilitators of the use of evidence by policymakers. *BMC Health Services Research*, 14, 2. <https://doi.org/10.1186/1472-6963-14-2>

Porciello, J., Ivanina, M., Islam, M., Einarson, S., & Hirsh, H. (2020). Accelerating evidence-informed decision-making for the Sustainable Development Goals using machine learning. *Nature Machine Intelligence*, 2(10), 559–565. <https://doi.org/10.1038/s42256-020-00235-5>

Ricciardi, V., Wane, A., Sidhu, B. S., Godde, C., Solomon, D., McCullough, E., Diekmann, F., Porciello, J., Jain, M., Randall, N., & Mehrabi, Z. (2020). A scoping review of research funding for small-scale farmers in water scarce regions. *Nature Sustainability*, 3(10), Article 10. <https://doi.org/10.1038/s41893-020-00623-0>

Strydom, W. F., Funke, N., Nienaber, S., Nortje, K., & Steyn, M. (2010). Evidence-based policymaking: A review. *South African Journal of Science*, 106(5–6), 17–24.

Takaaki, M., Custer, S., Eskenzai, A., Stern, A. & Latourell, R. (2017). Decoding Data Use: How do leaders source data and use it to accelerate development? AidData at the College of William & Mary.

Tierney, M. J., Nielson, D. L., Hawkins, D. G., Roberts, J. T., Findley, M. G., Powers, R. M., Parks, B., Wilson, S. E., & Hicks, R. L. (2011). More dollars than sense: Refining our knowledge of development finance using AidData. *World Development*, 39(11), 1891–1906. <https://doi.org/10.1016/j.worlddev.2011.07.029>

U.S. Government. (2023). Foreign Assistance Database. ForeignAssistance.Gov. <https://www.foreignassistance.gov>

USAID. (2004). USAID Agriculture Strategy: Linking Producers to Markets. Washington DC.

USAID. (2015). ADS Chapter 579: USAID Development Data. Washington DC.

USAID. (2016a). Synthesis of Evaluations Related to the Feed the Future Learning Agenda.

USAID. (2016b). USAID Evaluation Policy.

<https://www.usaid.gov/sites/default/files/documents/1870/USAIDEvaluationPolicy.pdf>

Wazir et. al (2010) Final Performance Evaluation Accelerating Sustainable Agriculture Program (ASAP) Checchi and Company Consulting, Inc.

The UNU-MERIT WORKING Paper Series

- 2023-01 *Can international mobility shape students' attitudes toward inequality? The Brazilian case* by Cintia Denise Granja, Fabiana Visentin and Ana Maria Carneiro
- 2023-02 *Demand-led industrialisation policy in a dual-sector small open economy* by Önder Nomaler, Danilo Spinola and Bart Verspagen
- 2023-03 *Reshoring, nearshoring and developing countries: Readiness and implications for Latin America* by Carlo Pietrobelli and Cecilia Seri
- 2023-04 *The role of product digitization for productivity: Evidence from web-scraping European high-tech company websites* by Torben Schubert, Sajad Ashouri, Matthias Deschryvere, Angela Jäger, Fabiana Visentin, Scott Cunningham, Arash Hajikhani, Lukas Pukelis and Arho Suominen
- 2023-05 *More than a feeling: A global economic valuation of subjective wellbeing damages resulting from rising temperatures* by Stephan Dietrich and Stafford Nichols
- 2023-06 *Was Robert Gibrat right? A test based on the graphical model methodology* by Marco Guerzoni, Luigi Riso and Marco Vivarelli
- 2023-07 *Predicting social assistance beneficiaries: On the social welfare damage of data biases* by Stephan Dietrich, Daniele Malerba and Franziska Gassmann
- 2023-08 *Related or unrelated diversification: What is smart specialization?* By Önder Nomaler and Bart Verspagen
- 2023-09 *Breach of academic values and digital deviant behaviour: The case of Sci-Hub* by Giulia Rossello and Arianna Martinelli
- 2023-10 *The effect of lobbies' narratives on academics' perceptions of scientific publishing: An information provision experiment* by Giulia Rossello and Arianna Martinelli
- 2023-11 *Making impact with agricultural development projects: The use of innovative machine learning methodology to understand the development aid field* by Lindsey Moore, Mindel van de Laar, Pui Hang Wong and Cathal O'Donoghue