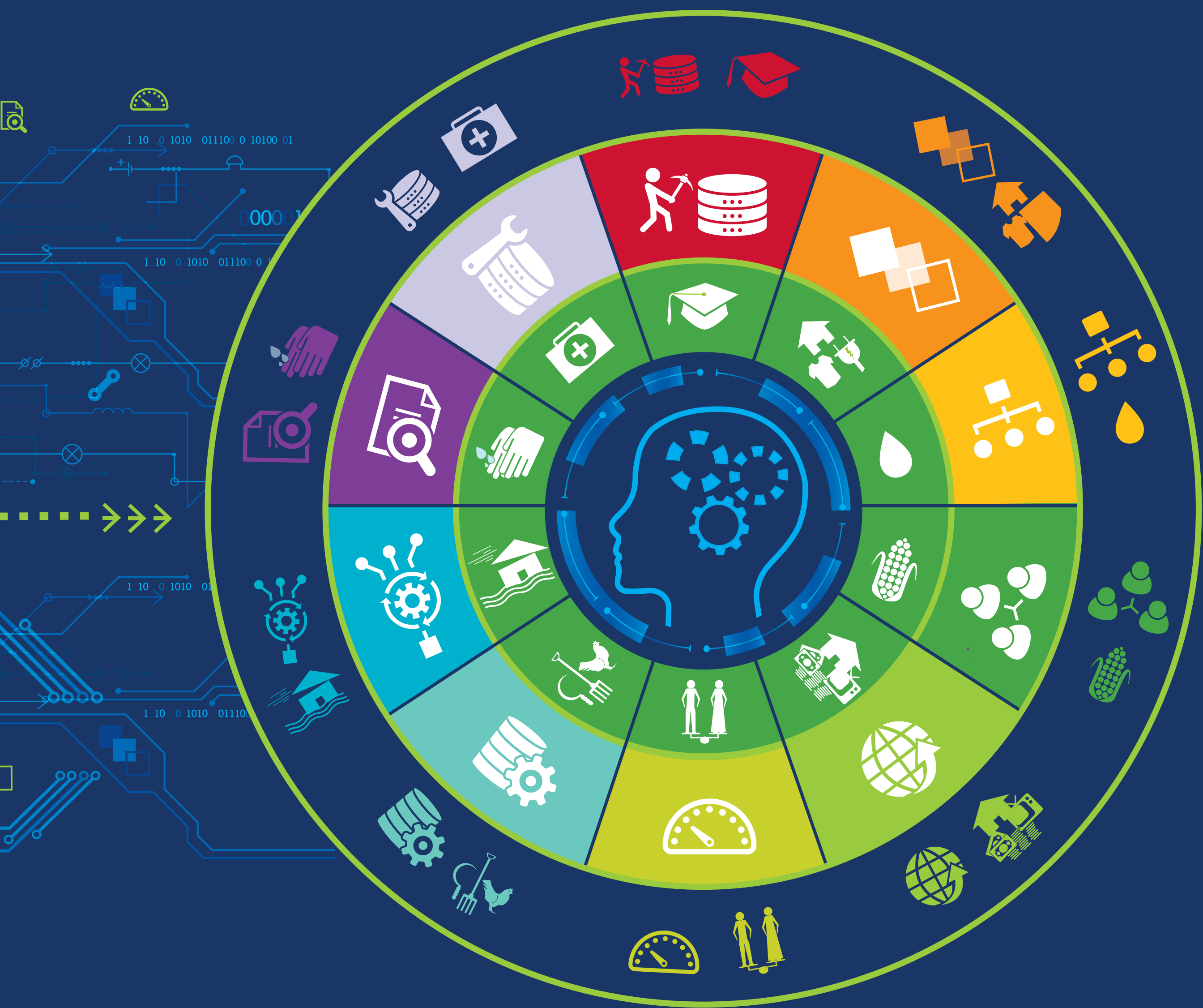


Accelerating knowledge generation for data-driven decision making

Leveraging Artificial Intelligence and Big Data for **IFAD 2.0** Innovation Challenge 2019

IFAD
Investing in rural people

Accelerating knowledge generation for data-driven decision making

Leveraging Artificial Intelligence and Big Data for **IFAD 2.0**
Innovation Challenge 2019

Final Report

February 2020

IFAD
Investing in rural people

Contents

List of tables and figures

Acknowledgements

This report is the output of the Innovation Challenge 2019 project "Leveraging Artificial Intelligence and Big Data for IFAD 2.0", led by Alessandra Garbero, Senior Economist in the Research and Impact Assessment Division (RIA). The project was the result of a collaborative effort with other Divisions, and most prominently the Operations and Results Division (OPR), where strong support was given by Lauren Phillips, Lead Results Specialist. The project was implemented within the timeframe of four months, from September to December 2019. The following consultants worked towards the success of this project: Bia Carneiro, Giuliano Resce, Giovanni Cerulli, Giosue Rescica, Jaron Porciello, Liliana Andriano, Luis de Los Santos, Davide Bellucci, Veronica Notaro. The work benefited from technical inputs from the Sustainable Production, Markets and Institutions Division (PMI), notably from the Director Thouraya Triki, and PMI staff: Richard Abila, Jean-Philippe Audinet, Mawira Chitima, Robert Delve, Michael Hamp, Enrique Hennings, Mylene Kherallah, Rikke Grand Olivera, Antonio Rota, and Kathy Zissimopoulos. Consultations were also held with Programme Management Department (PMD) Front Office (Tim Balint) and OPR staff (Ursula Wieland). Support was also provided by the Information Communication Technology (ICT) and Administrative Services Divisions (ADM). Additional people contributed to this project, notably Giovanni Giunta, Luisa Basile and Tuuli Sauren. A special acknowledgement therefore goes to all those who worked on this project as well as to those who made this possible, notably the Change Delivery and Innovation Unit (CDI), Edward Gallagher and Rebecca Slocum. Lastly, we are grateful to the RIA Director, Sara Savastano, and Associate Vice-President, Strategy and Knowledge Department, Paul Winters, for supporting this project.

Executive Summary

In the context of the Innovation Challenge initiative, sponsored by the Change Delivery and Innovation Unit (CDI) within IFAD, this project sought to bring innovation by unlocking the potential of artificial intelligence to accelerate knowledge generation and strengthen data-driven decision making in IFAD. A multi-disciplinary team of economists, data scientists and social scientists employed Machine Learning techniques to extract insights from IFAD investments, globally, across the entire portfolio. This enabled a global overview of types of investments and outcomes, the completion of systematic reviews to document impact of key interventions, and the development of models that are able to predict performance at project level and quantify the extent of positive impacts given certain targeting and project-level features. As IFAD12 reinforces the trend towards fewer, more focused, and larger investments in each country, as well as a focus on doubling impact and sustainability, gaining a comprehensive picture of the portfolio can support the achievement of strategic objectives and the Sustainable Development Goals (SDGs).

The project had three main aims – the first, to use machine learning for historical portfolio identification or systematization of almost 40 years of project implementation, and gather a global overview of the distribution of investments and outcomes; the second, to enhance and accelerate knowledge management, by extracting evidence on impact of key IFAD investments as well as improve targeting of impact assessments of IFAD-supported projects to under-evaluated areas; and the third, to allow for predictive analytics. Here, the project tried to predict two items: 1) the likelihood of project success (e.g. performance) based on portfolio features; 2) the likelihood of positive impacts at the household level given the type of interventions delivered on the ground and key household-level features.

A critical mass of structured (e.g. quantitative) as well as unstructured (e.g. text) data was gathered as part of this project to allow the application of a mixed methods approach. Using techniques such as text mining, natural language processing, systematic reviews, meta-analysis, and predictive analytics, some of the main findings of this initiative include the following:

- As IFAD11 works towards mainstreaming climate change, gender, youth and nutrition, text mining on more than 2000 documents uncovered an upward trend in reporting against them, especially with regards to climate change adaptation.
- Further text mining based on key terms related to Agenda 2030's Sustainable Development Goals indicated an increase in the presence of SDG-related content in project documentation across all 17 goals.
- A collaboration with Cornell University leveraged an advanced machine learning model designed specifically for agriculture to examine project reports for interventions, topics and outcomes. Three main classes of interventions were detected: socioeconomic, technology, and ecosystem.

- Socioeconomic interventions represented 37% of the dataset, within which finance and government-related interventions were the most frequently reported. Technology interventions followed very closely (36%) and comprised primarily of crop and irrigation-focused activities.

- Ten outcome classes were also detected. Expected outcomes relating to livelihoods (e.g. economic mobility) and production were the most mentioned within the documentation, followed by water use and resilience.

This project enabled leveraging existing data to explore new questions and gain additional knowledge. For IFAD to improve its focus on results, strengthen mechanisms for successful project design, and become a leader on measuring and attributing impact against the SDGs, machine learning and other artificial intelligence approaches must be used not only in an experimental fashion through venues such as the Innovation Challenge, but mainstreamed into IFAD's everyday work.

1. Introduction

1.1. How can artificial intelligence and Big Data support IFAD's mission?

Developments around artificial intelligence and Big Data offer great potential for international finance institutions to improve evidence-based decision-making. Big Data research is increasingly considered the new gold standard. While Kitchin (2014) defines Big Data as huge in volume and generated continuously – which gives it depth, flexibility and scalability – boyd and Crawford (2012) point out that Big Data is not so much about massive amounts of data as it is about the "capacity to search, aggregate, and cross-reference large data sets" (2012:663). Scholars have shown that Big Data contains information for complex phenomena that may be difficult to observe using traditional methods (di Bella *et al.* 2016; Einav, Levin, 2014).

In particular, Big Data are available at a larger scale, on novel types of variables (created from text), for a low cost (no new surveys needed to collect information) and in real-time, thus closing the time gap between observation and analysis, which is a typical challenge of studies based on traditional approaches (Giannone *et al.* 2008). Of course, some important issues remain unsolved for the use of Big Data in the social sciences. As with any other kind of data, Big Data is in reality socially created, thus having a social life of its own that requires caution and a critical perspective when engaging with it (Lupton, 2012).

A few of the great challenges of Big Data research include coping with heterogeneity, its dynamic yet time-sensitive information, filtering through noisy data, coming to terms with the fact that much of what is generated is a by-product of other activities not related to the researcher's specific question, and problems of privacy and transparency (di Bella *et al.* 2016; Kitchin, 2014). While sheer volume makes Big Data seem like it can answer any question, boyd and Crawford highlight the danger of encouraging apophenia: *"the practice of seeing patterns where there are none, simply because enormous quantities of data can offer connections that radiate in all directions"* (2012: 668). Undoubtedly, identifying patterns is hugely different from explaining them, and Big Data only has any purpose when properly analysed. As affirmed by Kitchin, detecting a pattern should not be the endpoint of research, but rather the starting point for additional analysis (2014).

Big Data shares the limelight with Artificial Intelligence as the vanguard of knowledge creation. While the terms machine learning and artificial intelligence are often used interchangeably, ***"machine learning is part of artificial intelligence, which in turn is a discipline in computer science. Machine learning aims to learn from data using statistical methods"*** (Storm et al 2019: 1). To put it simply, artificial intelligence uses computers to automate decision-making processes. Within the much larger scope of artificial intelligence, machine learning comprises various methods that get computers to recognize patterns in data, and then uses these patterns to make future predictions.

Framed within this promising, yet challenging context of recent technological developments, this project sought to unlock the potential of artificial intelligence to accelerate knowledge generation and strengthen data driven decision making in IFAD. A multi-disciplinary team of economists, data scientists and social scientists employed Machine Learning techniques to extract insights from IFAD investments. This enabled the completion of systematic reviews to uncover impact of key interventions, and the development of models to predict project performance and likelihood of impact given certain targeting features. As IFAD12 reinforces the trend towards fewer, more focused, and larger investments in each country, as well as a focus on doubling impact and sustainability, gaining a comprehensive picture of the portfolio can support the achievement of strategic objectives and the SDGs.

2. Project aims and objectives

The project had three main aims – the first, to use machine learning for historical portfolio identification or systematization of almost 40 years of project implementation, to facilitate a global overview of type of investments and outcomes; the second, to enhance and accelerate knowledge management and improve targeting of impact assessments of IFAD-supported projects to under-evaluated areas; and the third, to allow for predictive analytics. Here, the project tried to predict two items: 1) the likelihood of project success based on portfolio features; 2) the likelihood of positive impacts at the household level given the type of interventions delivered on the ground and key household-level features. Predictive analytics has the goal of informing project design about features of successful projects as well as the likelihood of achieving positive impacts given certain types of interventions and targeting characteristics.

2.1. Systematize IFAD's investment portfolio

As IFAD's Development Effectiveness Framework and accountability mechanisms contribute to the generation of thousands of portfolio-related documents, as well as data from project reports to ratings and cost tables, the institution can be considered a "walking data generator" (McAfee et al. 2012). The unstructured mass of information continuously generated through IFAD's routine activities belongs to the family of Big Data, as documents contain a wealth of information about interventions, activities, costs, outcomes, populations targeted, and more. However, while various tools and data sources are available within IFAD to support portfolio analyses – such as project reports, impact assessments, cost tables (COSTABs), Flexcube, the Grants and Investment Projects System (GRIPS), the Independent Office of Evaluation's independent evaluation ratings database (ARRI Database), and Oracle Business Intelligence (BI) – currently, there is no centralized system that integrates all the information in a way that provides quickly and confidently information about what IFAD invests in, where, and how much (IFAD, 2019a). Consequently, extracting knowledge from these various sources through conventional approaches such as manual keyword searches is a challenge.

Hence, the starting point for this project's data analytics approach was to first gather a critical mass of data, structured and unstructured (e.g. text), to then try to uncover trends in IFAD's investment portfolio globally, over the span of 38 years: specifically, what thematic areas IFAD invests in, through what types of intervention, for whom and where. By understanding the heterogeneity of projects over time, the intention was to establish a **taxonomy** of project topics, interventions and outcomes. To ensure the corporate relevance of this innovation, IFAD technical specialists and domain experts were consulted to identify key concepts and establish a thesaurus of keywords, as well as to have a sense of which knowledge gaps in the portfolio they wanted the analyses to uncover.

2.2. Enhance knowledge management

Systematic reviews and meta-analysis are powerful techniques for results synthesis and can complement IFAD project-level impact assessments in domains where IFAD-specific evidence is lacking. Systematic reviews and meta-analysis are also instrumental to identify under-evaluated areas where IFAD should conduct more impact assessments to gain targeted knowledge about the impact of its investments. As such, this project intended to apply data-driven decision-making to first define key interventions within the IFAD portfolio, to then use AI to speed up systematic reviews and meta-analyses that measure the extent of evidence and impact across such domains and eventually provide guidance to where impact assessment would be most needed across the portfolio.

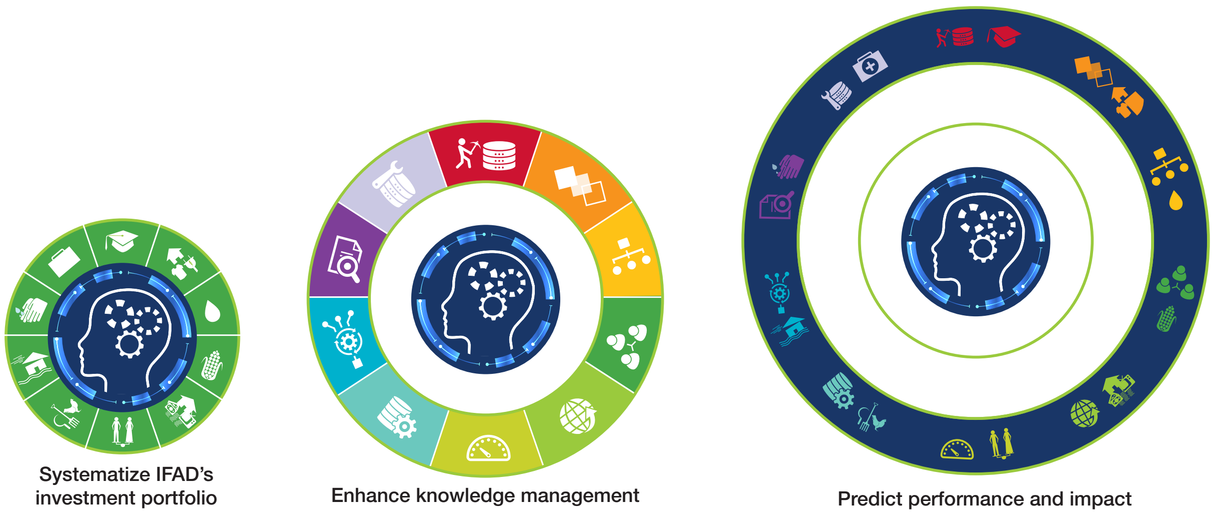
To this end, a collaboration with Cornell University was established to apply machine learning algorithms to speed up evidentiary searches for systematic reviews and meta-analyses, leveraging on the Ceres2030 initiative, a partnership between Cornell University, the International Food Policy Research Institute (IFPRI) and the International Institute for Sustainable Development (IISD) funded by BMZ (Germany's Federal Ministry of Economic Cooperation and Development) and the Bill and Melinda Gates Foundation (BMGF). Ceres2030 evaluated agricultural interventions that could contribute to *SDG 2: Zero Hunger* and transform the lives and incomes of the world's poorest farmers while preserving the environment. The initiative created a database of more than 500,000 publications in the domain of agriculture and rural development interventions and thanks to the partnership with IFAD, the database was used and extended to extract IFAD-relevant interventions that contribute to IFAD strategic objectives and overarching goal[1].

Initially, it was decided to test the approach on the domain of livestock interventions, given its relevance in the IFAD portfolio. As a second step, the results from the portfolio systematization informed a data-driven decision on other key interventions for which similar reviews will be carried out to assess impact and inform about effectiveness.

2.3. Predict performance and impact

The third step in this project was to harness and integrate the evidence generated through machine learning to develop models that support the project cycle through ex-ante predictions of project performance, as well as ex-ante predictions of the likelihood of positive impact of IFAD-supported policies, given a specific set of portfolio and beneficiary features.

Two main exercises were carried out. First, a framework was devised to combine the topics detected in IFAD documents through text mining and natural language processing with projects' objective features, such disbursement data, ratings, and country-level characteristics from World Development Indicators, in order to create a model that predicts the probability of project success and performance. Second, household-level data from IFAD10 impact assessments[2] were used to develop a prediction model to map the probability of project success (positive outcome), based on a set of household and project characteristics that affect impact. While the first exercise, can inform about successful features for portfolio performance, the second could guide targeting at the project-design level by determining the beneficiary features that drive positive impact.



1 IFAD Strategic Objectives (SOs) are the following: 1) increase the productive capacity of small-scale producers; 2) increase their benefits from market participation; 3) strengthen the environmental sustainability and climate resilience of their economic activities. IFAD's overarching goal is defined as increasing economic mobility of its beneficiaries. See IFAD Strategic Framework 2016-2025 (https://www.ifad.org/en/strategic-framework)

2 IFAD10 Impact Assessment Report: Draft (https://webapps.ifad.org/members/eb/126/docs/EB-2019-126-R-4.pdf)
IFAD10 Impact Assessment Report (https://webapps.ifad.org/members/eb/127/docs/EB-2019-127-R-5.pdf)

3. Methodology

In a machine-learning environment, computers use various algorithms to stimulate human learning and perform tasks, and the performance of these tasks is subsequently improved based on acquiring new knowledge from human experts. Machine-learning models are most successful when they have clearly stated communication and classification tasks and a plethora of available data. Figure 1 shows a simplified look at the machine learning process, where data is gathered, processed, and statistical models called "machine learning models" are applied. They are tested for accuracy and eventually turned into analytics and visualisations.

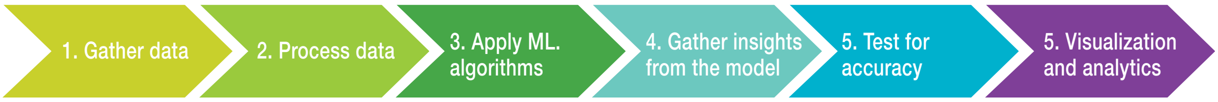


Figure 1 The Machine Learning process (Source: Ceres2030)

3.1. Gathering the data

As mentioned above, various resources are available within IFAD to support portfolio analyses. The research pooled together quantitative and qualitative data (e.g text) for all investment projects in IFAD's portfolio, totalling almost 900 projects, as well as data from external sources. At the project level, all available performance management reports were collected, namely: 1. Project Design Reports (PDR), 2. Mid-Term Review Reports (MTR), 3. the last Project Supervision Report available for each project (PSR); and Project Completion Reports (PCR). In total, 2302 documents for 856 projects with approval dates ranging from 1981 to 2019 were collected, from which 1462 were in English; 531 in French; 288 in Spanish; and 21 in Portuguese.

Disbursement data was also compiled, as well as internal and external ratings from the Programme Management Department (PMD) and the Independent Office for Evaluation (IOE), respectively, and detailed cost tables (COSTABs). At the household level, impact assessments for projects from IFAD10 were used. Lastly, World Development Indicators (WDI) were collected for all the countries included in the dataset. Table 1 below details the various datasets.

Moreover, consultations conducted with IFAD technical specialists identified key terminology for each domain so that an IFAD-specific taxonomy could be devised. Through these consultations and by referring to the newly approved Categorisation Framework (IFAD, 2019a), keywords for each category, sub-category and cross-cutting themes were identified.

Table 1 Datasets

Total IFAD portfolio	**1313**
Investments projects considered (grants excluded)	894
Projects with at least one report available	856
Reports collected (PDR, MTR, PSR, PCR)	2302
Projects with disbursement data	705
Projects with internal evaluation data	562
Projects with external evaluation data	251
Projects with COSTABs	361
Projects with impact assessments	19
Countries with World Development Indicators	124

3.2. Processing the data

The first step in data preparation was making suitable for analysis all the non-suitable documents present in the database. This included converting all project report files into searchable documents, in order to enable data extraction. Customised scripts in Python and Bash were developed using the open source engine Tesseract to find and convert hundreds of image and MS Word documents from the database into searchable PDFs. While seemingly simple, this was a very resource-intensive task, as detecting language and page orientation to provide better results demands computer processing power and time.

Once all documents were ready, algorithms were applied to automate text extraction and prepare a corpus with the entire set of project reports (removal of figures, punctuation, stop words, and numbers; conversion of words to lowercase; stemming[3]). Outputs from this step include a merged corpus containing more 180,000 pages of text, and a Term Document Matrix, which identified more than 197,000 unique terms distributed across the documents. Table 2 below shows a small part of the Term Document Matrix as an example.

Table 2 Part of the Term Document Matrix considering the whole text extracted from 2302 IFAD documents

Terms \ Docs	albania_1100001339_design-report.pdf	albania_1100001339_supervision-report.pdf	albania_1100001452_completion-report.pdf
aminur	0	0	0
amip	0	0	0
amir	0	0	0
amira	0	0	0
amis	0	0	0
amiss	0	0	0
amissah	0	0	0
amistad	0	0	0
amit	0	0	0
amiti	0	0	0
amiv	0	0	0
amiz	0	0	0
amizmiz	0	0	0
amjad	0	0	0
amk	0	0	0
amka	0	0	0

3.3. Applying machine learning

3.3.1. Descriptive text mining

Descriptive text mining is an approach that seeks to automate the retrieval of high-quality information from text, usually by finding patterns and trends through machine learning, statistics and linguistics. The aim of this technique is to enable the analysis of large amounts of unstructured text to derive insights. Text mining was applied to 2302 project reports from 856 projects, with the objective of identifying IFAD's activities as reported through performance management mechanisms. This approach focused on two main questions:

1. As IFAD focuses on mainstreaming climate change, gender, youth and nutrition, how do projects report against these and other cross-cutting issues?
2. How are IFAD projects working towards Agenda 2030's Sustainable Development Goals?

3.3.1.1. Presence of cross-cutting issues and mainstreaming themes

To answer the first question, a two-part supervised model was devised. First, the consultation with IFAD technical experts resulted in the identification of 118 keywords for eight cross-cutting issues, as proposed by the new categorization framework (IFAD, 2019a): climate change adaptation, climate change mitigation, gender, youth, indigenous peoples, nutrition, natural resource management and land tenure, and emerging issues. IFAD's four mainstreaming themes of climate change, gender, nutrition and youth (IFAD, 2019b) are found within these issues. The terms were then translated to the four languages detected in the documents (English, French, Spanish, and Portuguese).

However, as there are often several ways to describe a particular concept (i.e. women, female, girl, etc.), the terms specified by the experts may be represented differently in the documentation. As such, an important element of textual analysis is semantic association. Identifying synonyms enhances the ability to uncover patterns in the corpus. These potential term variations were tackled through a machine learning procedure that expanded each

3 Stemming is the process of reducing each word to its word stem that affixes to suffixes and prefixes or to the roots of words. Stemming is an important step in natural language understanding and natural language processing.

manually collected keyword using Word2vec[4], a model used to produce a distributed representation of similar words (Resce, Maynard 2018; Mikolov *et al.* 2013). In the interest of time, multi-language, pre-trained algorithms trained on Wikipedia by Bojanowski *et al.* (2017) were used. To prevent over-generation, the expansion was limited to 1000 words for each cross-cutting theme, per language (totalling 32,472 keywords).

Next, words were stemmed and detected in the corpus. Taking only unique words for each cross-cutting issue[5], the presence of a particular issue in a document was estimated by the share of words – as a percentage – associated to that issue within the document. Documents were then aggregated by project ID to calculate the presence of issues at the project-level (see equation 1 in Appendix I).

As a further step, the relationship among projects and the cross-cutting issues was explored through network analysis. This technique enables the visualization of relational data organised as matrices, where entities are the nodes – in this case, projects and the cross-cutting themes – and their relations are the lines connecting pairs of nodes – here calculated by the weighted share of words for each cross-cutting theme, for each project. This means that a project is "connected" to an issue if associated words are present according to the criteria detailed in equations 2-4 in Appendix I. The strength (or weight) of this connection is based on the percentage of words related to that particular issue, which captures both the extensive and the intensive margins of the spread of issues among projects – that is, not just the presence of an issue, but also the share of words as a measure of significance.

Another way to represent the importance of a theme or an issue is by weighting its presence against the amount invested by IFAD in a particular project. The final analysis done in the corpus was to calculate the cost equivalent of the share of words for each of the eight cross-cutting issues based on IFAD investment per project (see equation 5 in Appendix I)

3.3.1.2. Presence of Sustainable Development Goals

In 2015, more than 190 world leaders committed to Agenda 2030's 17 Sustainable Development Goals (SGDs). In contrast to their predecessors, the Millennium Development Goals that expired in 2015, the SDGs were designed to tackle the root causes of poverty and other deprivations, taking into account their interconnectedness (Hák *et al.* 2016). Hence, a key feature is their focus on the mobilization of financial resources, as well infrastructure and technology (Zilberman *et al.* 2018).

In order to explore how IFAD projects are contributing towards the SDGs, in terms of objectives and semantics in the documents, text mining was employed to determine the frequency of terms associated to the official SDGs definitions in the four languages detected in project reports. Similar to the cross-cutting issues, an expansion for each term in the definitions was performed through the Word2vec model (Mikolov *et al.* 2013) trained on Wikipedia by Bojanowski *et al.* (2017). Expansion was also limited to 1000 words per SDG, per language (a total of 68.000 keywords).

Words were then stemmed and detected in the corpus. Taking only unique words for each SDG[6], the presence of a particular goal in a document was estimated by the share of words – as a percentage – associated to that goal within the document. Documents were then aggregated by project ID to calculate the presence of the SDGs at the project-level (see equation 6 in Appendix I).

The relationship among projects and the SDGs was also explored through network analysis, where projects and SDGs are represented as nodes and their relations are the lines connecting pairs of nodes calculated by the weighted share of words for each SDG, for each project. A project is "connected" to a SDG if associated words are present according to the criteria detailed in equations 7-9 in Appendix I.

4 Word2vec is a suite of models used to find words semantically similar to others by means of similarity measures. The similarity is determined on a large unlabelled corpus, based on the notion that semantically similar words have a similar context. The main output of Word2vec is a fixed-length vector for each word. These vectors are then used to find similarity among words with standard similarity measures. This study applied the approach proposed by Bojanowski (2017), which is based on the skip-gram model, where each word is represented as a bag of character -grams. This method outperforms the baseline (Mikolov *et al.* 2013), since it takes into account sub-word information, rare words, and morphologically rich languages (Bojanowski, 2017).

5 Words present in more than one category were excluded.

6 Words present in more than one SDG were excluded.

As before, to represent the importance of a SDG, its presence was weighted against the amount of funds IFAD invested in a particular project, where the cost equivalent of the share of words for each of the SDGs was calculated based on IFAD investment per project (see equation 10 in Appendix I).

3.3.2. Natural language processing and topic modelling

The partnership with Cornell University enabled the application of a machine-learning model developed by Porciello (forthcoming, 2020) for the *Ceres2030: Sustainable Solutions to End Hunger project*. As discussed previously, this initiative has evaluated agricultural interventions that could contribute to *SDG 2: Zero Hunger*, and has successively supported 77 researchers from 23 countries together in eight teams to produce evidence syntheses on agricultural interventions that have the potential to improve the livelihoods of small-scale farmers while minimizing the impact on the environment. The eight manuscripts will be published (pending peer-review) in Nature Research Journals, expected by June 2020.

This approach aimed to explore how machine learning could be employed to detect intervention types, topics and outcome details from IFAD project documentation. A dataset containing one report for each of around 800 projects was analysed. As the focus was on detecting and discerning patterns of interventions from an IFAD-specific perspective, the dataset prioritised Project Design Reports (PDR) whenever they were available. The criteria was because PDRs are written by IFAD staff and approved by the IFAD Board, making them the documents most consistently aligned with the institution's strategic objectives, whereas the other documents are often written by external consultants and as such do not always follow the same standards. If a PDR was not available for a particular project, an alternative document was selected in the following order of priority: Project Completion Report (PCR), Mid-Term Review Report (MTR) and Project Supervision Report (PSR). The corpus comprised 170,000 pages of text in four languages: English, Spanish, French and Portuguese. To narrow down the analysis and reduce noise, a keyword detection process was used to find relevant pages containing intervention or implementation data, including first searching for words such as "intervention," "project objective", "component", "subcomponent", etc., in all four languages. In the end 56,000 pages from 743 reports were extracted and processed.

The model uses open-source models and algorithms including Bidirectional Encoder Representations from Transformers (BERT), Support Vector Machines (SVM)- KNN-Stochastic Gradient Boosting Machines, Word2Vec with applied Heart patterns, and Linear Discriminant Analysis (LDA). Each algorithm is designed to perform a different task, such as topic modelling, intervention detection, and measurement detection. The model has been 'pre-trained' to understand language and syntax using Google News and Wikipedia. It has also been trained using summary and citation-level data from 500,000 materials in agriculture (peer-reviewed articles and grey literature). This makes it very suitable to detect and classify agricultural-specific details such as populations, geographies, study design types, numerical measurements and more. The model was trained to detect and classify "interventions" first by looking to understand the concept of an intervention, its various meanings and syntax, before creating an intervention taxonomy. The intervention taxonomy is comprised of both broad and narrow classes. Table 3 is an overview of the classification taxonomy for interventions.

Table 3 Example of taxonomy of interventions

Broad class category	Examples
Technology	Molecular markers, zero tillage
Socioeconomic	Input price, agricultural extension services
Ecosystem	Alternative fuel, agroforestry
Storage	Jute bags, storage pits
Mechanization	Threshers, tractors
Non intervention	Associative mapping, overviews

The text is fed through the model and the algorithms work together to isolate certain parts of the text, such as where terminology that is relevant for outcomes and interventions are detected. Those are then extracted and fed through the model again for further analysis and parsing. The model was applied to the IFAD dataset to detect and label interventions, outcomes, topics, normalized keywords, populations, crops and plants and animals. The data was analysed in three tiers:

1. At the individual document level, in this case, the "page level." This is useful to see the breadth and depth of granular data across the dataset.

2. At the report level, where all detected pages and relevant data are collapsed into one 'field' with all of the data. This is the gold-standard in terms of wanting to detect what is happening at the project level and to be able to provide general conclusions of what is happening across countries, regions, interventions, and more.
3. At the dataset level. The dataset level relies on page counts, primarily, to give a sense of overall robustness of the corpus of documents.

The data was first gathered and then processed in the machine model environment. All of the raw data was provided back to IFAD as tabular data files. As this was the first time the Ceres2030 model was applied to project documentation, there were challenges related to the extraction and analysis of the results, primarily because the short timeline of this project meant there was little time to explore pre-test pages as well as alternative models to support the analysis. Pre-testing pages would have reduced the overall noise in the dataset, improved the precision of thematic classification, and reduced the amount of resources required to run the model.

3.3.3. Systematic reviews and meta-analyses

While systematic reviews and meta-analyses are powerful techniques for results synthesis, the style of systematic reviews presently used in academia can take 18 months to three years to reach consensus on a single intervention, by which time the research risks being out-of-date. Search functions, based on keywords and meta tagging, are inadequate, and especially likely to miss important research in agriculture. The access to dozens of databases needed to find relevant research represents a major obstacle for many researchers, universities, and institutions, and duplication in literature searches wastes valuable resources. In this context, a machine-learning approach can dramatically shorten the time to perform an evidence synthesis.

The partnership with Cornell University also employed machine learning to speed up data collection (e.g. evidentiary searches) for a systematic review and meta-analysis on IFAD-supported interventions. Initially, it was decided to test the approach on the domain of livestock interventions, given its relevance in the IFAD portfolio. As a second step, the results from the portfolio systematization informed a data-driven decision on other key interventions for which similar reviews will be carried out to assess impact and inform about effectiveness. Further thematic systematic reviews will be carried out from among the following broader classes of interventions: finance and government-related socioeconomic interventions; crops and irrigation-related technology interventions; and conservation-related ecosystem interventions.

3.3.3.1. Systematic review on Livestock

Livestock can contribute to preventing people from falling into poverty and, eventually, to ending hunger and all forms of malnutrition (FAO 2018). Previous empirical evidence finds that livestock interventions can have positive effects on socio-economic outcomes in low- and middle-income countries, but there is a limited understanding of the causal links between them and contextual factors that shape their effectiveness. This type of work is especially important in light of the increasing investments IFAD has made in this sector. As such, a systematic review of the evidence for IFAD-supported livestock interventions was carried out to answer two questions:

1. What is the causal impact of IFAD-supported livestock interventions on agricultural production, agricultural productivity, market access, resilience, mitigation, adaptation, nutrition, gender and youth empowerment, business profit, income, assets, expenditure, employment, and poverty?
2. What are the most important factors that shape the extent to which the causal impacts exist?

First, searching the database of all IFAD projects, the specific livestock activities that IFAD invests in were identified, as well as the countries and the timeframes. Based on commonly used guidelines, a specific set of study inclusion criteria is formulated. This protocol is a key requirement of the systematic review methodology, as it provides a clear guide to be used during the search, bolsters the transparency of the review, and enables replication. The inclusion criteria are the following:

- Study focus includes population of smallholder farmers, micro small and medium-sized enterprises, youth, and women;
- Study involves a micro-level impact evaluation with a valid identification strategy and a counterfactual, i.e. experimental or quasi-experimental designs which include randomized control trials, propensity score matching,

instrumental variable approaches, endogenous switching regression models, differences in differences, regression-discontinuity design;

- Study describes the livestock intervention implemented and the intervention is also present within the IFAD portfolio; in addition to this, the impact on agricultural production, agricultural productivity, market access, resilience, mitigation, adaptation, nutrition, gender and youth empowerment, business profit, income, assets, expenditure, employment, and poverty is estimated;
- Study area of focus includes target populations in low- and middle-income countries (See Appendix I for a full list of included countries);
- Study is in English, French, German, Spanish or Portuguese;
- Academic papers from journals and PhD thesis are acceptable, Masters or Undergraduate theses are ineligible;
- Study considers the following animals: cattle, cows, goats, poultry, chickens, broilers, ducks, rabbits, yaks, pigs, sheep, camelids, bees.

Conducting a broad and thorough search as part of a systematic review is imperative in order to maximise the insights produced by the review. One advantage of considering different study designs is that the search expands the number of robust quantitative studies around a topic. This also helps avoid potential bias in the research identified, such as the inclusion of only published or widely cited studies. Accordingly, the aim of the search is to provide a comprehensive list of all relevant published and unpublished studies. Cornell University provided data from their licensed databases, including the preparation and execution of an exhaustive literature search that resulted in almost 20,000 citations. Additional data was added into the review based on recommendations from IFAD's technical expert on livestock. The Cornell team also processed the underlying data (citations) to provide labelled spreadsheets of machine-detected metadata for more than 32 metrics, including study design, interventions, geographies, and more, to accelerate the research process. The IFAD team used the data to review and remove irrelevant data.

The search also involved the use of "snowballing", where new studies are sourced from the reference sections of relevant studies already identified through the search, and the systematic probing of the websites of key practitioners, universities, and think-tanks (See Appendix II for the list of the 133 sources searched for this review). The latter was particularly time-consuming, especially as more than 100 websites were used. While it would have been possible to diminish the number of sources for the search to speed up the process, this could have also led to potential bias in the studies identified. Thus, a software was developed in Python to automate a Google search and download studies that met the criteria.

The databases and websites were searched using pre-defined search terms, with efforts being made to keep them as uniform as possible across sources. The main search terms used for this review were the following:

livestock AND intervention*
AND ("impact evaluation"
OR "impact assessment")

The process of identifying eligible studies through the search results is to first screen the title and abstract and finally the full text according to the review's inclusion criteria. The use of text mining tools and ML algorithms speeds up the first screening process, and the outcome is validated by manual (human) inspection. If a study passes these three rounds, they are added to the final set of studies to be included in the review. A total of 19,197 studies were identified through the search, from which the number of studies deemed relevant from title and abstract was 198.

After these were screened full text, 50 studies met all criteria and were thus selected for this review. A critical appraisal of the selected studies was then performed. Particularly following Waddington *et al.* (2014) and Garbero *et al.* (2018) the studies are evaluated based on the risk of bias (due to internal validity or causal identification). Each study is assigned a bias score based on the following questions:

1. Attrition bias (only relevant for panel datasets): is there any evidence that there is systematic attrition between the survey rounds?[7]
2. Selection bias: is there a randomization factor or do participants self-select into the programmes?

7 This criterion is relevant only for panel data studies and is not considered in the calculation of the final bias score for each study.

3. Hawthorne effect bias: is there monitoring of the participants that might have changed their behaviours and the final outcomes?
4. Spillover effects: is there a large geographical distance between treatment and control groups to guarantee that the benefits of the treatment are not received by the control groups?
5. Selective reporting bias: is there any evidence of selective reporting? Are there gaps in the analyses that seemed purposefully omitted?
6. If the authors use propensity score matching/ instrumental variables, do they provide diagnostic statistics to ensure that the necessary assumptions are met?

Terciles of the bias score were constructed and each study was assigned to a risk of bias category (with studies in the lowest tercile classified as "High risk of bias studies", those in the second tercile as "Medium risk of bias studies" and those in the highest tercile as "Low risk of bias studies"). The higher the bias score, the better the paper meets the internal validity criterion. When multiple papers (and estimates) from the same study exist (e.g., a working paper and a journal article with the same authors), the study associated with the lowest risk of bias was preferred.

In addition to assessing the studies' internal validity, their external validity was also critically appraised, and a generalizability score was assigned based on the following screening questions:

1. Motivation of the research: is the context explained? Is there an adequate literature review?
2. Sampling descriptions: are the descriptive statistics provided? Is the data collection process described? Is the sampling strategy appropriate?
3. Completeness of analysis.
4. Presence of triangulation methods: do the authors use several robustness checks in the estimations?
5. Quality of conclusions and discussions.

The higher the bias and generalizability scores, the better the paper meets the internal and external validity criteria.

The next steps currently in course involve performing a quantitative meta-analysis and a meta-regression analysis. The meta-analysis is a method for synthesising the findings of the studies identified by a systematic review by producing aggregate estimates of the impact of an intervention on a given outcome, based on the quantitative findings of the identified studies. To perform the meta-analysis, standardized estimates for all included studies have to be computed and the final estimates of the meta-analysis are expressed in terms of effect sizes.

Because of the greater ease of interpretation and comparability across different contexts, the response ratio (RR) was chosen as the appropriate effect size metric as opposed to the standardized mean differences. The computation of the RR was not so straight forward and was time-consuming for various reasons. In the instance where additional information was required to calculate the effect size, the corresponding authors were contacted, and this slowed down the process. In other cases, studies reported multiple dependent effect sizes, for example according to different follow-up periods or when multiple variables measure the outcome of interest. In these cases, it was possible to calculate a "synthetic effect size" based on the sample-weighted average, using appropriate formulas to recalculate variances according to Borenstein *et al.* (2009, chapter 24).

After standardising the effects as RRs, the summary statistics (impact estimates or effect sizes) for each study were combined using a variety of meta-analytic methods; these can be classified as fixed-effect models or random-effects models. In a fixed-effect model, the main assumption is that the true effects are the same across studies. Following this approach, larger studies are given more weight because they lead to more representative results. In a random-effects model, the true effects are assumed to differ across studies, and the differences between the true effects and the observed effects are due to differences in the true effects (in addition to differences in sampling error). Because livestock interventions may have different impacts in different settings, a random-effects model was chosen to derive the final estimate.

Finally, a meta-regression will be performed to determine the reason for the heterogeneity of estimates between studies. A meta-regression is a linear estimation of the effect size on study characteristics, including the risk of bias and generalizability scores, as well as other moderator variables such as whether the study is published or the region of analysis. This step makes it possible to identify whether or not there is a linear relationship between effect sizes and study characteristics and thus has the potential to highlight the main determinants driving the magnitude of the impact of livestock interventions.

3.3.4. Predictive analytics

Predictive analytics comprise a variety of statistical techniques from data mining, predictive modelling, and machine learning, that analyse current and historical facts to make predictions about future outcomes or trends. Insights generated are intended to guide discussion and critical inquiry, rather than taken as 'absolutes' to apply to a decision-making framework. In this case, the project aimed to develop algorithms to support the project cycle through ex-ante predictions of performance and likelihood of positive impact of a certain policy (treatment effects) based on set of features.

3.3.4.1. Project performance prediction

From a policy perspective, predicting the performance of a new project based on its characteristics can support IFAD's decision-making processes. On a practical level, for instance, together with the PDR, the Board can receive information about the predicted chance of success/failure of a particular project as estimated by a prediction tool.

Project performance can be influenced by both internal factors, which are those that IFAD can control directly and that it can potentially alter to improve, and external factors, which are within the purview of governments or outside of the overall context of IFAD's interventions (e.g. conflict situations). As such, the envisioned model considered the following two definitions of performance: 1) disbursement performance; and 2) implementation performance based on ratings (self-evaluation). Three main types of predictors were employed: 1) macro-economic country-level indicators; 2) project features concerning financial attributes and extent of outreach (quantitative data from ORMS[8], Oracle BI – GRIPS[9] and Flexcube); 3) projects features determined through text mining and classification algorithms obtained within this project.

Based on a previous study (Balint et.al., 2019), disbursement performance can be measured by "disbursement readiness" (average time from approval to effectiveness and first/second disbursement) and "disbursement effectiveness" (cumulative disbursement rates during the project's lifetime and at financial closure). Projects with serious delays for a first disbursement indicate a lack of project readiness and the likely occurrence of subsequent implementation problems. In order to prepare the dataset for developing the prediction model, disbursement data was collected from Flexcube, excluding start-up advances as well as pre-financing for Botswana, Mexico and Morocco. The outcomes extracted were:

1. Time from approval to the first disbursement
2. Time from approval to entry into force
3. Time from entry into force to first disbursement
4. Time from first disbursement to second disbursement
5. Time from approval to second disbursement

This data was then integrated and augmented with the topics and interventions detected through the Cornell University model to identify, through word-based statistics, any significant characteristics that influence project performance, as well as with the internal and external ratings attributed to projects by IOE and PMD. Lastly, the macro-economic indicators considered were the World Development Indicators (WDI), which is the World Bank's compilation of cross-country comparable data on development. These were collected for all countries for which projects have been found.

8 The Operational Results Management System (ORMS) is an IFAD internal system for the management and tracking of project related data – including Logframe, Performance, Action tracker and Lessons Learned.

9 The GRIPS reports provide data and information from the corporate to the project level thus supporting management in decision-making. GRIPS reports allow reporting on grants and investment projects.

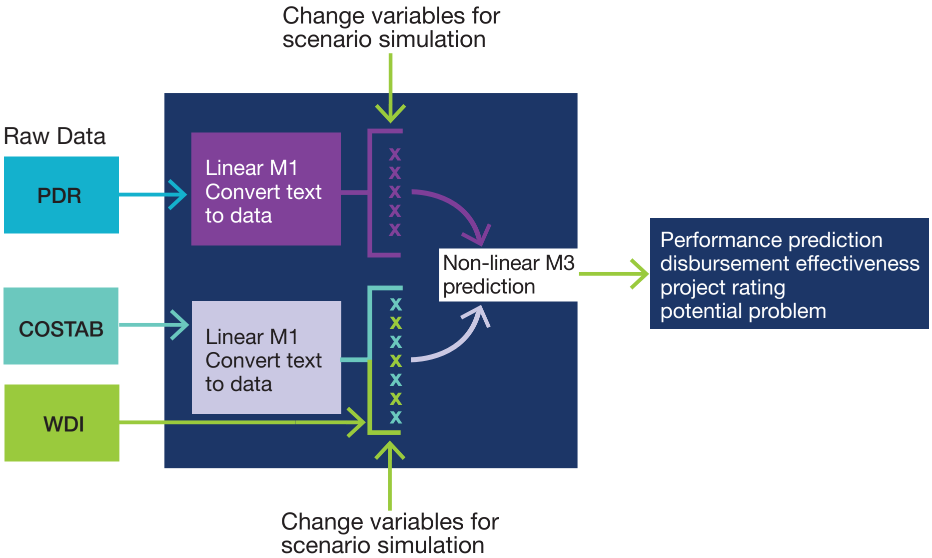


Figure 2 Conceptual framework for project performance prediction model

3.3.4.2. Predicting likelihood of positive impact from household-level impact assessments characteristics

Analysis of the heterogeneity of impacts (e.g. treatment effects) helps understand the conditions under which policies or treatments are effective or ineffective, as well as inform the design and implementation of interventions, so as to maximize their effectiveness. In this project, household heterogeneous effects, defined as impacts contingent on household characteristics, were analysed for 19 projects from IFAD10 in order to develop a prediction model to map the probability of project success (i.e. a positive outcome), as well as household and project characteristics that affect positive impact. In lay terms, the percentage of households exhibiting positive impacts were identified in the IFAD10 impact assessment studies.

Analytically, the household *i* Treatment Effect (*TE*) is defined as the difference in the outcome interest if the household is treated (i.e. receives the policy) minus what the outcome would have been in the absence of the policy (Garbero, 2016):

$$TE = y_{1i} - y_{0i}$$

As it refers to the same household at the same time, just one of the two quantities, not both (missing observation problem) can be observed at the same time for the same household. The main population parameters can be defined as

$$\text{Average Treatment Effect} = ATE = \mathrm{E}(y_{1i} - y_{0i})$$

$$\text{Average Treatment Effect on Treated} = ATET = \mathrm{E}(y_{1i} - y_{0i} \mid w_i = 1)$$

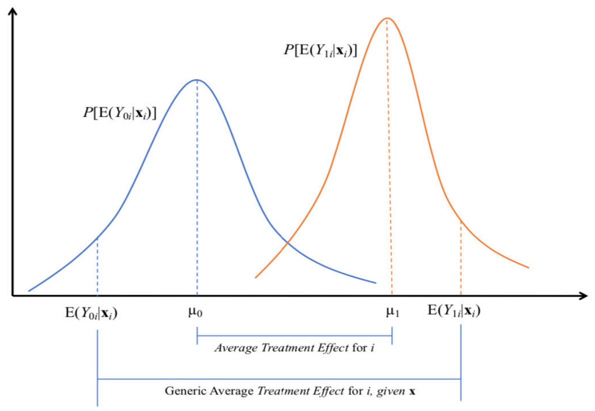
$$\text{Conditional Average Treatment Effect} = (\boldsymbol{x}_1) = \mathrm{E}(y_{1i} - y_{0i} \mid \boldsymbol{x}_1)$$

$$\text{Conditional Average Treatment Effect on Treated} = \tau_1(\boldsymbol{x}_1) = \mathrm{E}(y_{1i} - y_{0i} \mid \boldsymbol{x}_i, w_i = 1)$$

The heterogeneous effect is measured by $\tau(X)$ or $\tau_1(X)$ and **indicates the effect of the policy for each household given its characteristics**. Under conditional mean independence, it is found that:

$$(\boldsymbol{x}_i) = \mathrm{E}(y_{1i} - y_{0i} \mid \boldsymbol{x}_i) = \mathrm{E}(y_i \mid \boldsymbol{x}_i, w_i = 1) - \mathrm{E}(y_i \mid \boldsymbol{x}_i, w_i = 0)$$

which is an observable quantity. The same result holds for $\tau_1(\boldsymbol{x}_i)$. Figure 2 shows the distribution of $\mathrm{E}(y_{1i} \mid \boldsymbol{x}_i)$ and $\mathrm{E}(y_{0i} \mid \boldsymbol{x}_i)$, whereas Figure 3 presents the distribution of $\tau(\boldsymbol{x}_i)$. In the latter, the coloured part represents the number of households getting a positive effect. Identifying the percentage of these households was the main task of this analysis.



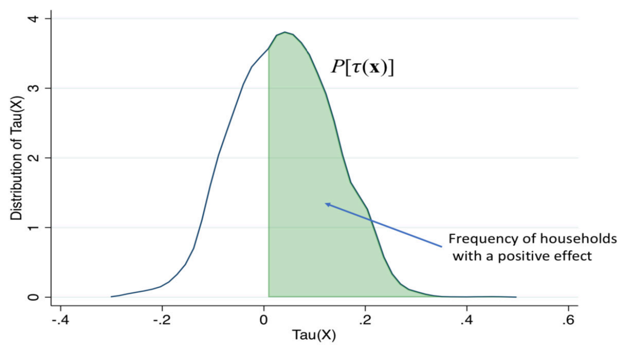


Figure 3 Distribution of E(y1y | $\boldsymbol{x}$i) and E(yi | $\boldsymbol{x}$i)

Figure 4 distribution of ($\boldsymbol{x}$i)

Table 4 shows the outcome and control variables of the model specification employed in this analysis. There are nine outcome variables and eleven control variables, plus the binary treatment variable w, which represents the IFAD-specific intervention evaluated as part of the impact assessment. It is important to note that these initial analyses only considered outcomes pertaining to economic mobility – e.g. income, both total, gross and net, from either all sources, crops and livestock – as well as IFAD's strategic objective of improving resilience (proxied by the indicator Ability to recover from shocks). Table 5 shows the 15 countries considered in this study and their respective number of observations (Sao Tomé & Principe and Madagascar were excluded from the analysis due to time constraints).

Table 4 Outcome and control variables used for heterogenous effect policy evaluation.

Outcome variables	Control variables
Total Gross Income from all Sources	Household size
Total Net Income from all Sources	Square of the household size
Gross Income from Crops and By-products	Age of the household head
Net Income from Crops and By-products	Square of the age of the household head
Gross Income from Livestock and By-products	Number of children
Net Income from Livestock and By-products	Number of adults
Net income from fisheries (Bangladesh & Indonesia)	Years of education
Ability to recover from shocks – 1	Education level
Ability to recover from shocks – 2	Gender of the household head (1=female)
	Area of land owned by household
	Dependency ratio

Table 5 Countries considered and number of observations.

Country	N. of observations
Bangladesh	2743
Bolivia	2751
Brazil	1939
Chad	2198
China	1801
Ethiopia	2924
Indonesia	2956
Kenya	2562
Mexico	2230
Nepal	2874
Philippines	1803
Rwanda	4355
Senegal	2233
Tajikistan	2262
Tanzania	1986
TOTAL	**38011**

4. Main findings

4.1. Reporting against IFAD's mainstreaming themes and cross-cutting issues

Overall, descriptive text mining uncovered an upward trend in reporting against IFAD's mainstreaming themes and cross-cutting issues, especially with regards to climate change adaptation. The first explorative exercise was to identify the share of words associated to IFAD's mainstreaming themes and cross-cutting issues in each document, namely: climate change adaptation, climate change mitigation, gender, youth, indigenous peoples, nutrition, natural resource management (NRM) and land tenure, and emerging issues. A matrix was devised with the eight themes by column and the 2302 documents by row, where the percentage of words in a document that were also present in a theme was calculated. Table 6 below shows an extract of the table, with the head of the matrix and the share of words in each IFAD document that are associated to themes, where yellow indicates a high frequency and green indicates low frequency.

Table 6 Head of the matrix with the share of words in each project that is associated to a theme (keywords for themes manually collected and expanded by word2vec).

Country	Project.Short.Name	Climate Change Adaptation	Climate Change Mitigation	Emerging Issues	Gender	Indigenous Peoples	NRM and Tenure Security	Nutrition	Youth
Afghanistar	RMLSP - AF	3,998	0,515	1,198	0,684	0,384	1,136	1,621	2,131
Afghanistar	CLAP - AF	5,947	0,606	1,598	0,652	0,455	1,561	1,598	0,561
Afghanistar	SNaPP2 - AF	4,410	0,301	1,300	0,127	0,251	1,360	0,833	0,744
Albania	Northeastern Districts	2,175	0,448	0,467	0,051	0,066	1,484	0,955	0,499
Albania	Small-scale Irrig. Rehab.	1,976	0,023	0,362	0,000	0,021	1,267	0,982	0,901
Albania	Mountain Areas Develop.	3,929	0,501	1,847	0,430	0,279	2,377	0,813	0,569
Albania	SDRMA	3,164	0,734	1,503	0,150	0,162	1,674	1,091	0,724
Albania	MMP	3,844	0,491	1,467	0,390	0,179	1,462	1,061	1,102
Algeria	Artisanal Fisheries Pilot	1,575	0,127	0,250	0,068	0,079	0,409	0,055	0,482
Algeria	Pilot Proj. Oued Saf Saf	0,898	0,103	0,082	0,155	0,158	1,091	0,715	0,341
Algeria	Wilaya of M'Sila	0,944	0,285	0,410	0,285	0,051	3,468	0,307	0,798
Angola	Malanje Smallholder Rehab	0,691	0,054	0,755	0,064	0,047	0,221	0,815	0,171
Angola	Northern Region Foodcrops	3,119	0,175	0,667	0,076	0,183	0,758	0,832	0,316
Angola	Northern Fishing	1,314	0,175	0,713	0,146	0,191	0,364	0,281	0,658
Angola	MOSAP	3,692	0,381	1,463	0,190	0,053	1,665	0,616	0,580
Angola	AFAP	3,138	0,639	1,376	0,545	0,500	1,424	1,860	0,852
Angola	SADCP-C&H-SAMAP	3,712	0,279	2,433	0,103	0,246	1,133	1,309	0,736
Angola	ARP	2,301	2,277	1,101	0,137	0,095	1,861	1,430	0,857
Argentina	PRODERNEA	1,006	0,027	0,358	0,250	0,403	0,771	1,299	0,054
Argentina	PRODERNOA	1,746	0,118	0,575	0,036	0,320	0,383	0,732	0,407
Argentina	PRODERPA	1,969	0,064	0,795	0,038	0,725	0,489	0,562	0,171
Argentina	PRODEAR	2,950	0,148	0,492	0,069	0,136	0,216	1,532	0,125
Argentina	PRODERI	1,459	0,138	0,533	0,051	0,226	2,688	0,604	0,245

Figure 5 shows the distribution of the eight themes. Overall, three main trends emerge. The first is characterised by issues with a wide distribution of percentages across the portfolio, such as Climate Change Adaptation and Natural Resource Management and Tenure Security. The second trend is characterised by themes that are still well spread, but for which the share of words is less distributed, namely Climate Change Mitigation, Youth, Nutrition, and Emerging Issues. The third group is composed by issues with low presence in projects and a narrow distribution of percentages: Gender and Indigenous Peoples.

The presence of cross-cutting issues in IFAD documentation can be further visualised through the network graph in Figure 6. As discussed previously, the graph shows the connections between projects and cross-cutting issues. The force-directed algorithm used in constructing the network displays the spatialization of nodes, which maps the proximity and the authority of themes in relation to each other, (Jacomy *et al.* 2014). This means that linked nodes are drawn closer while unrelated nodes are pushed farther apart, thus allowing for a visual interpretation of the relationships between projects and themes. A modularity algorithm (Blondel et al., 2008) was applied to identify "communities", or clusters – as represented by nodes that are more densely connected together than to the rest of the network, and which are coloured accordingly. The resulting network indicates that **Climate Change Adaptation is the theme most widely spread across the documentation from both intensive and extensive margins** (i.e. it is present in more documents and with a higher share of words in relation to the other issues), placing it as the most central node in the network and closely related to Natural Resource Management and Tenure Security. The modularity analysis also identifies a cluster composed by Climate Change Adaptation, Climate Change Mitigation and Emerging Issues, indicating that these three are often present together in the documentation. Likewise, "Gender" and "Youth" are grouped into a cluster, albeit located more peripherally in the network. The average degree of connectivity of the network is 3.7, which means that, on average, each project addresses 3 to 4 mainstreaming themes or cross-cutting issues.

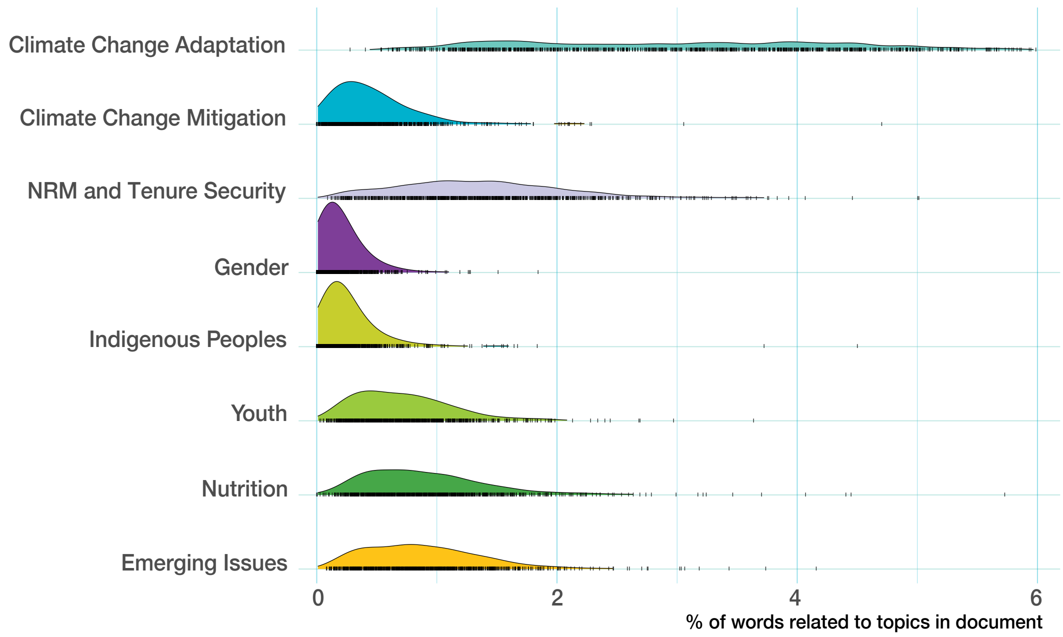


Figure 5 Distribution of mainstreaming issues and cross-cutting themes in project documentation. Share of words associated to IFAD's mainstreaming themes present in project documentation. Key terms collected manually and expanded by word2vec (2302 reports from 856 projects analysed).

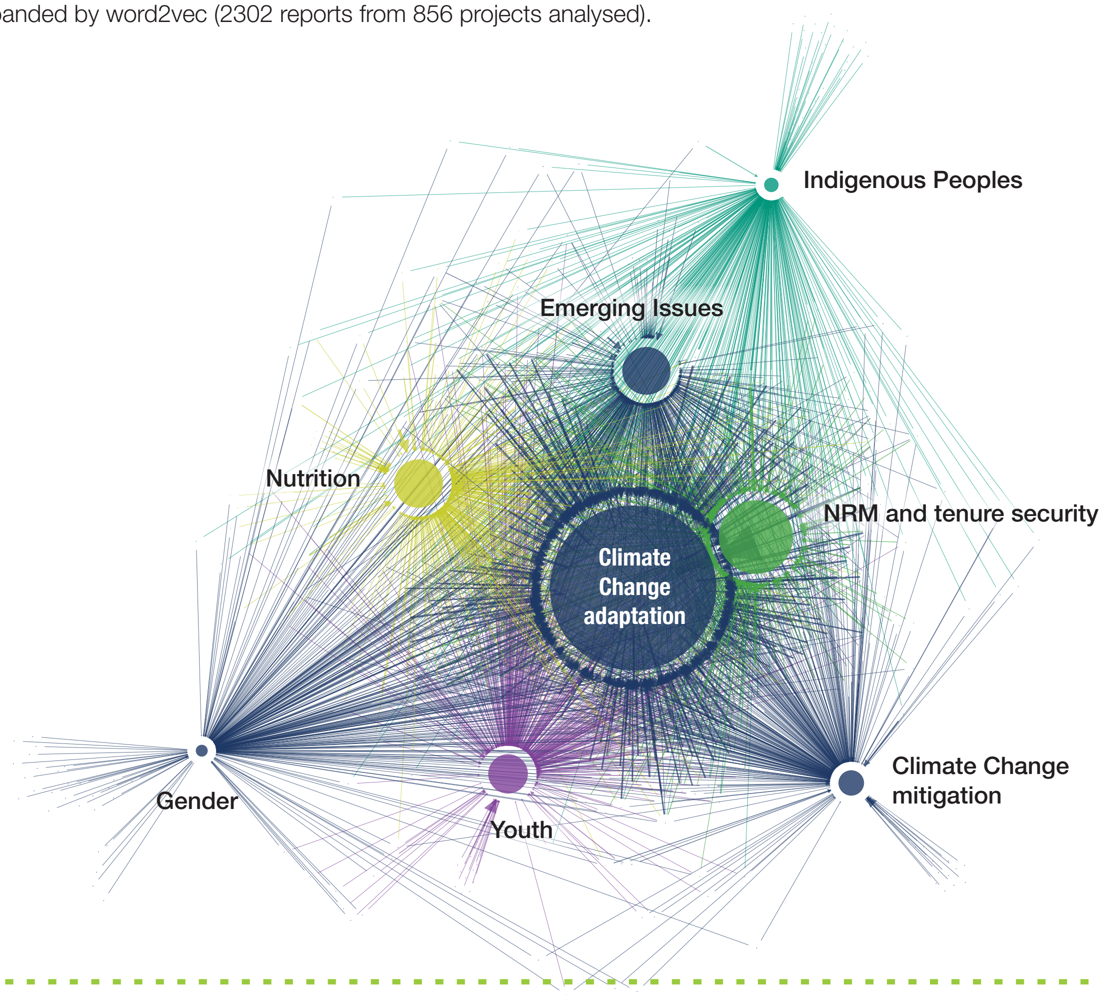


Figure 6 Network of cross-cutting themes detected in IFAD documentation. Parameters: force-directed graph, with node size partitioned as Weighed in-Degree, coloured by modularity class. 780 nodes (project IDs + themes) and 2890 edges (weighed by share of words). Average degree: 3.7.

Taking the date of approval for each project, Figure 7 shows time trends for the eight cross-cutting issues. All of them present significant increases, with the exception of Indigenous Peoples, which features a flat trend. This indicates that, in terms of addressing these themes across projects, IFAD interest in Climate Change Adaptation, Climate Change Mitigation, Natural Resource Management and Tenure Security, Gender, Youth, and Nutrition has increased between 1981 (the year of approval of the first document considered in this analysis) and 2019. The whole distribution of the eight issues per decade is shown in Figures A1-A8 in Appendix IV.

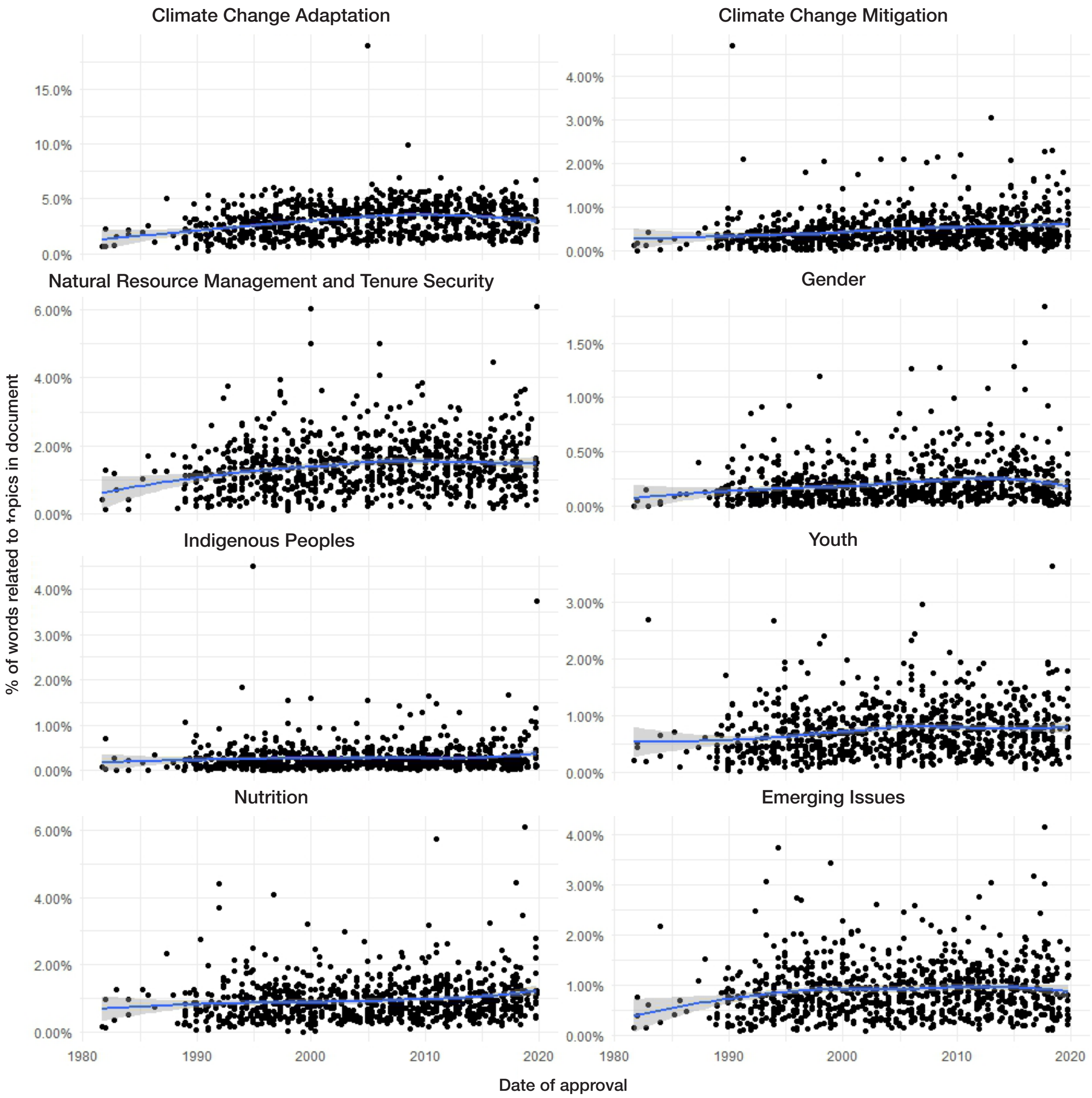


Figure 7 Time trend for the presence of mainstreaming themes and cross-cutting issues in IFAD documents, by project approval date (1981-2019). Share of words associated to IFAD's mainstreaming themes present in project documentation. Key terms collected manually and expanded by word2vec (2302 reports from 856 projects analysed).

Figure 8 shows the distribution of the themes across countries in which projects have been implemented. It is possible to discern some regional patterns, such as a stronger focus on nutrition in Latin America and the Caribbean, and a prevalence of projects covering Climate Change Adaptation and Natural Resource Management and Tenure Security issues in Africa and Asia.

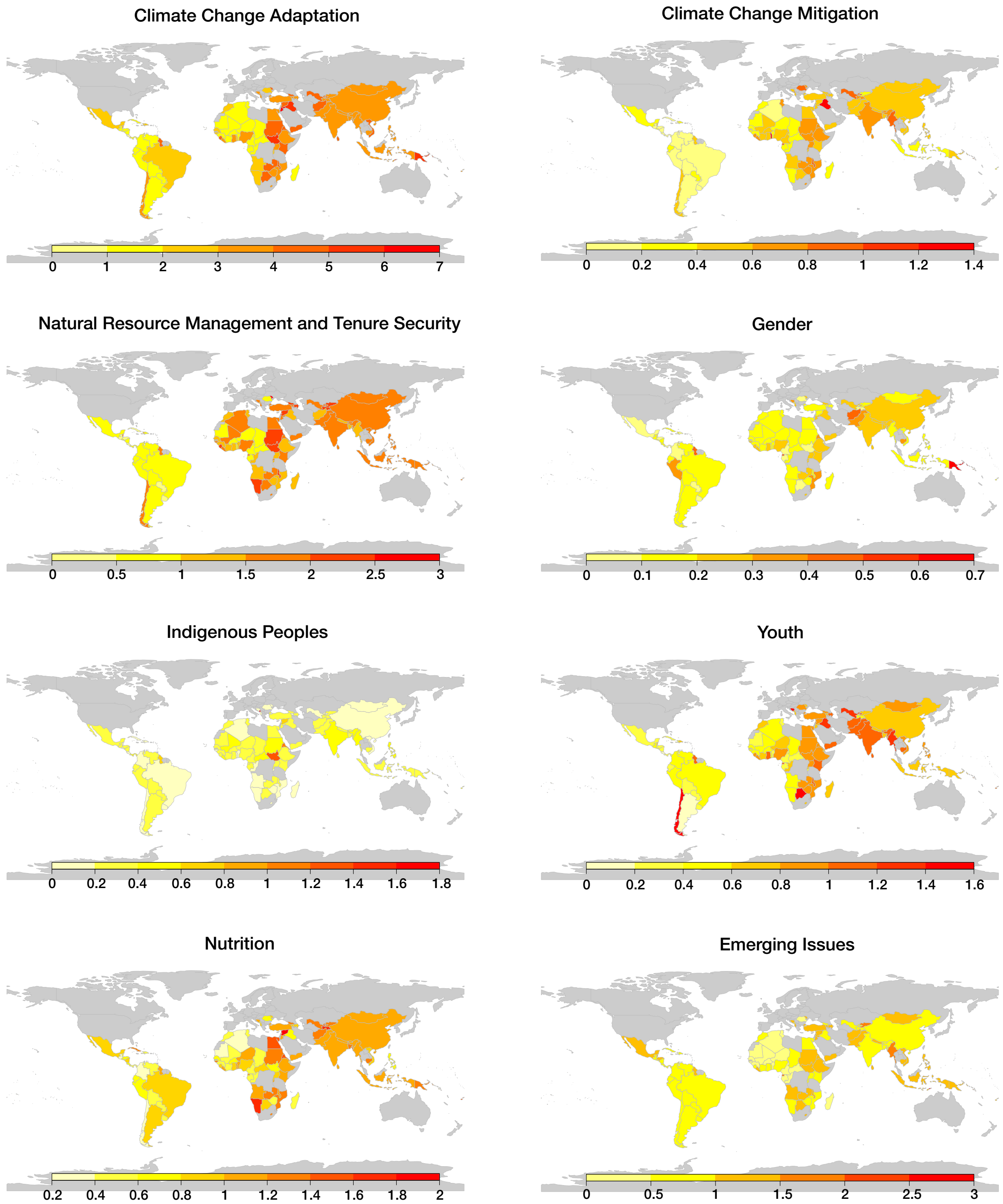


Figure 8 Country level analysis of the presence of mainstreaming themes and cross-cutting issues in IFAD documents. Average by country, share of words associated to IFAD's mainstreaming themes present in project documentation. Key terms collected manually and expanded by word2vec (2302 reports from 856 projects analysed).

In order to reflect on the monetary significance of the mainstreaming themes and cross-cutting issues, the share of words was weighted against IFAD project financing, as presented in Figure 9. Once again, an upward trend is detected, with Climate Change Adaptation dominating the other themes.

Text mining analysis can also help identify whether policies are put into "practice", as reflected through coverage in project documentation. Taking IFAD's Policy on Gender Equality and Women's Empowerment as an example, since its approval by the Executive Board in 2012, the policy has played a central role in supporting the overall goal of IFAD's Strategic Framework 2011-2015 of enabling poor rural women and men to improve their food security and nutrition, raise their incomes and strengthen their resilience. Figure 10 shows a slight increase in the presence of gender-related terms in documents for projects approved after 2012.

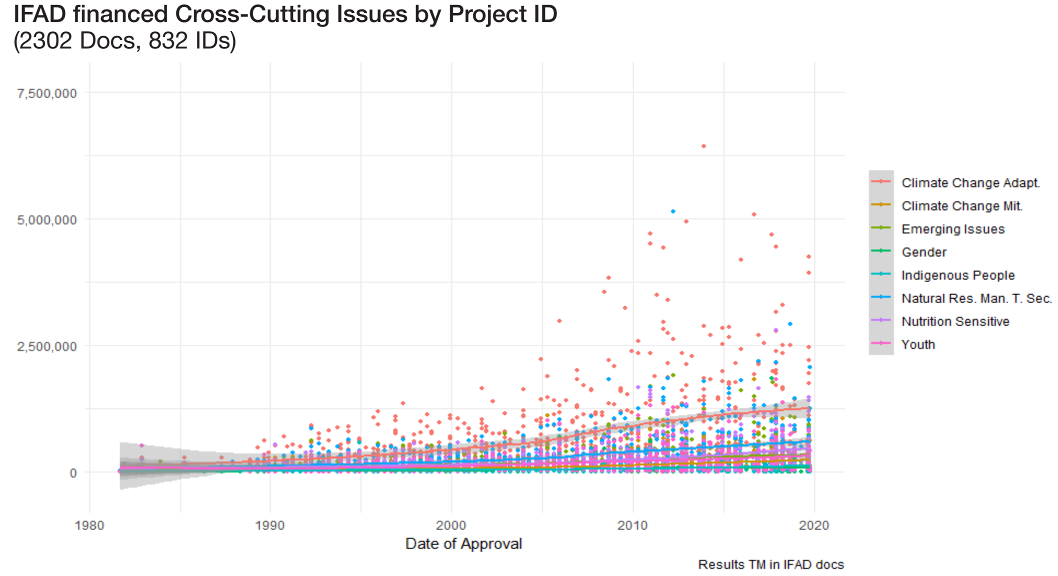


Figure 9 Time trend of the presence of mainstreaming themes and cross-cutting issues in IFAD documents, weighted for total financed by project. Share of words associated to IFAD's mainstreaming themes present in project documentation. Key terms collected manually and expanded by word2vec (2302 reports from 856 projects analysed).

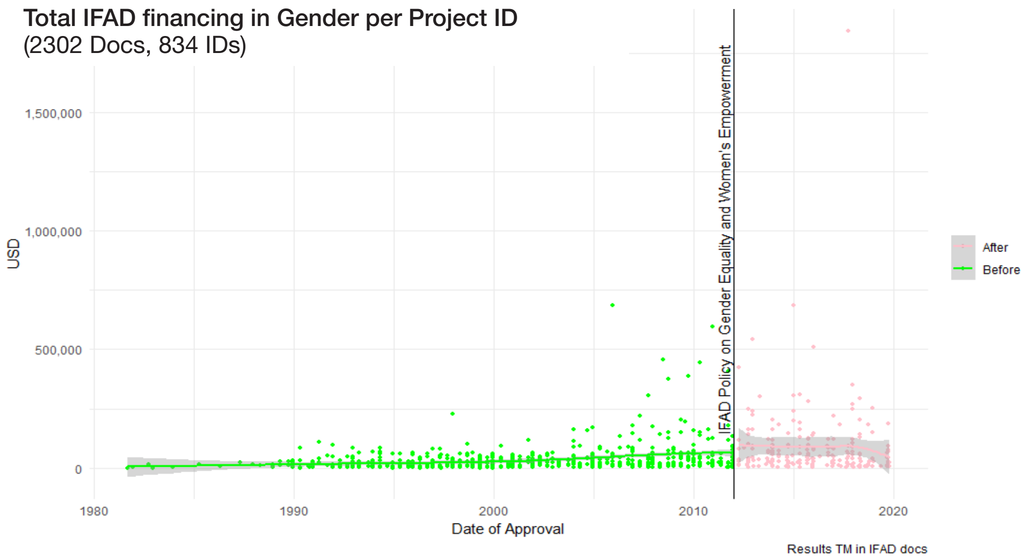


Figure 10 Time trend of the presence of gender associated terms in IFAD documents, weighted for total financed by project. Share of words associated to IFAD's mainstreaming themes present in project documentation. Key terms collected manually and expanded by word2vec (2302 reports from 856 projects analysed).

The analysis was augmented by exploring the association between the most frequent term for each issue (stemmed) and other words. This association translates into the correlation between words that appear in the same documents and is a starting point to analysing concepts and topics. Figure 11 shows the first 20 words for each theme. Some interesting associations include the most frequent stem word for the mainstreaming theme Youth, which is related to education, and for Gender and Nutrition, both of which are strongly correlated with a term indicative of "sensitive", denoting that IFAD's commitment to gender and nutrition sensitive interventions is reflected through the reports.

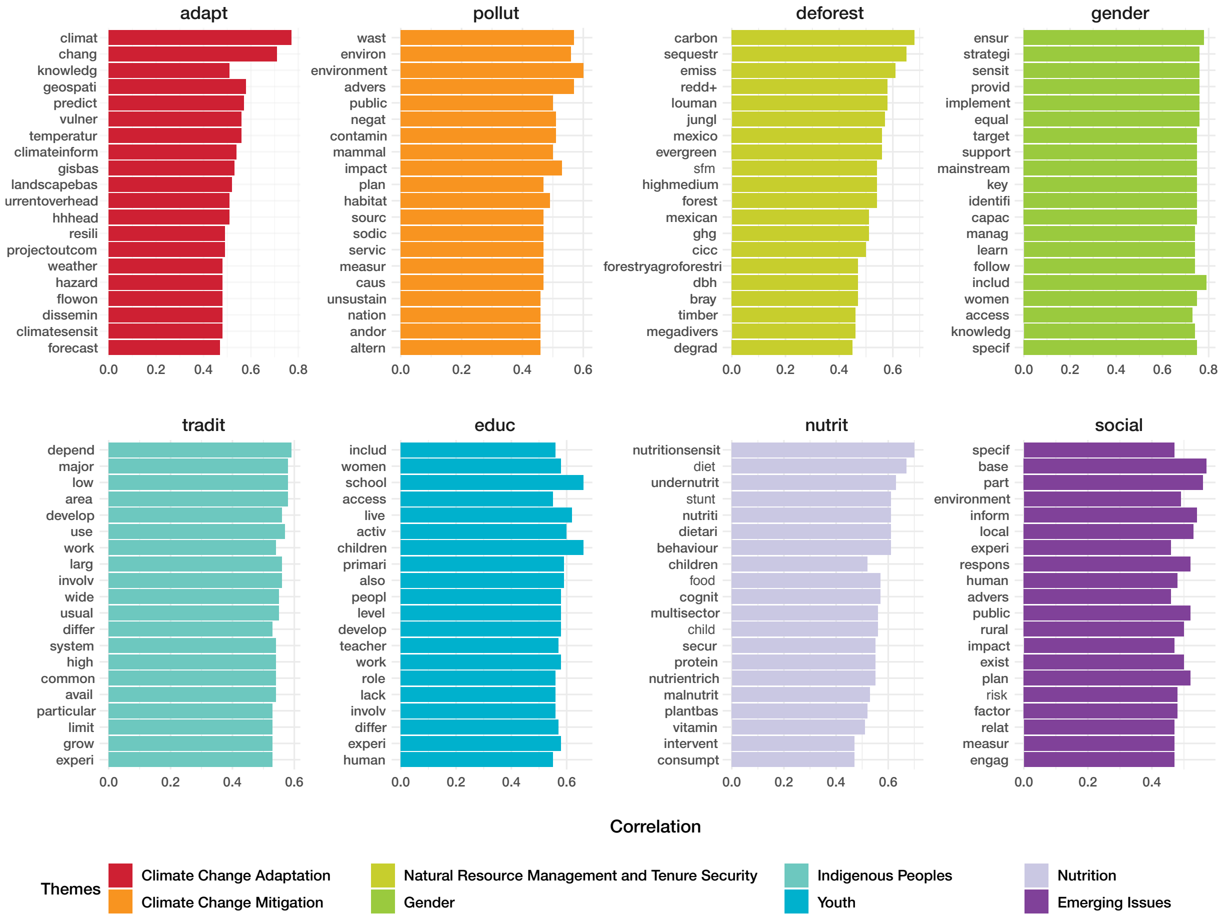


Figure 11 The first 20 words associated to the most frequent word for each of the cross-cutting issues present in IFAD documents. Share of words associated to IFAD's mainstreaming themes present in project documentation. Key terms collected manually and expanded by word2vec (2302 reports from 856 projects analysed).

4.2. Reporting against the Sustainable Development Goals

Following the same approach as the analysis of cross-cutting issues, in order to map the presence of terms associated to Agenda 2030's Sustainable Development Goals in IFAD documentation, a matrix was developed with the 17 SDGs by column and the 2302 documents by row. The percentage of words in a document that was also associated to a SDG was calculated. Table 7 shows an extract of the table, with the head of the matrix and the share of words in each project that are associated to SDGs, where yellow indicates a high frequency and red indicates low frequency.

Table 7 Head of the matrix with the share of words in each project that is associated to a Sustainable Development Goal (association is taken from public definition and expanded by word2vec).
Yellow = high; Red = low

Country	Project.Short.Name	1 NO POVERTY	2 ZERO HUNGER	3 GOOD HEALTH AND WELL-BEING	4 QUALITY EDUCATION	5 GENDER EQUALITY	6 CLEAN WATER AND SANITATION	7 AFFORDABLE AI
Afghanistan	RMLSP - AF	2,316	0,572	0,835	1,599	0,613	0,205	0,462
Afghanistan	CLAP - AF	0,788	0,030	1,568	1,924	1,462	0,083	0,833
Afghanistan	SNaPP2 - AF	1,143	0,305	0,714	0,837	0,792	0,308	0,229
Albania	Northeastern Districts	0,352	0,340	0,948	0,482	0,983	0,116	0,742
Albania	Small-scale Irrig. Rehab.	2,088	0,594	0,610	1,645	0,203	0,086	0,116
Albania	Mountain Areas Develop.	0,631	0,284	1,805	1,408	1,130	0,542	0,725
Albania	SDRMA	1,200	0,241	0,634	1,087	0,733	0,072	0,333
Albania	MMP	0,668	0,377	0,842	1,390	0,930	0,165	0,694
Algeria	Artisanal Fisheries Pilot	0,842	0,093	0,300	0,356	0,216	0,042	0,275
Algeria	Pilot Proj. Oued Saf Saf	0,550	0,365	0,512	0,198	0,264	0,164	0,301
Algeria	Wilaya of M'Sila	0,819	0,161	0,315	0,893	0,439	0,578	0,102
Angola	Malanje Smallholder Rehab	0,500	0,044	0,295	0,178	0,231	0,111	0,255
Angola	Northern Region Foodcrops	1,130	0,476	0,825	1,087	0,484	0,168	0,294
Angola	Northern Fishing	0,408	0,495	0,910	0,398	0,535	0,158	0,287
Angola	MOSAP	0,832	0,148	1,151	1,043	0,241	0,539	0,059
Angola	AFAP	1,404	0,445	1,269	1,246	1,122	0,407	0,183
Angola	SADCP-C&H-SAMAP	2,474	0,470	0,798	1,139	0,587	0,219	0,379
Angola	ARP	1,232	0,314	2,061	0,626	0,452	0,119	0,163
Argentina	PRODERNEA	0,317	0,816	0,574	0,091	0,355	0,587	0,267
Argentina	PRODERNOA	0,406	0,555	0,508	0,675	0,267	0,240	0,235
Argentina	PRODERPA	0,707	0,124	0,272	0,373	1,677	0,156	0,206
Argentina	PRODEAR	1,314	0,337	0,573	0,446	0,349	0,310	0,108

The overall distribution of the Sustainable Development Goals within project documentation is shown in Figure 12. A more distributed curve means that the share of words associated to each SDGs is more spread across documents. For instance, in the case of *Goal 1: No Poverty*, the percentage of words related to this SDG in the documents was concentrated between 0% and 3%. Conversely, the distribution of *Goal 14: Life Below Water*, was concentrated between 0% and 1%. Hence, besides *Goal 1: No Poverty*, the other goals that presented a wider spread of the share of words include *Goal 3: Good Health and Well-Being; Goal 4: Quality Education; Goal 10: Reduced Inequalities; Goal 12: Responsible Consumption and Production; Goal 13: Climate Action; Goal 15: Life On Land;* and *Goal 17: Partnerships.*

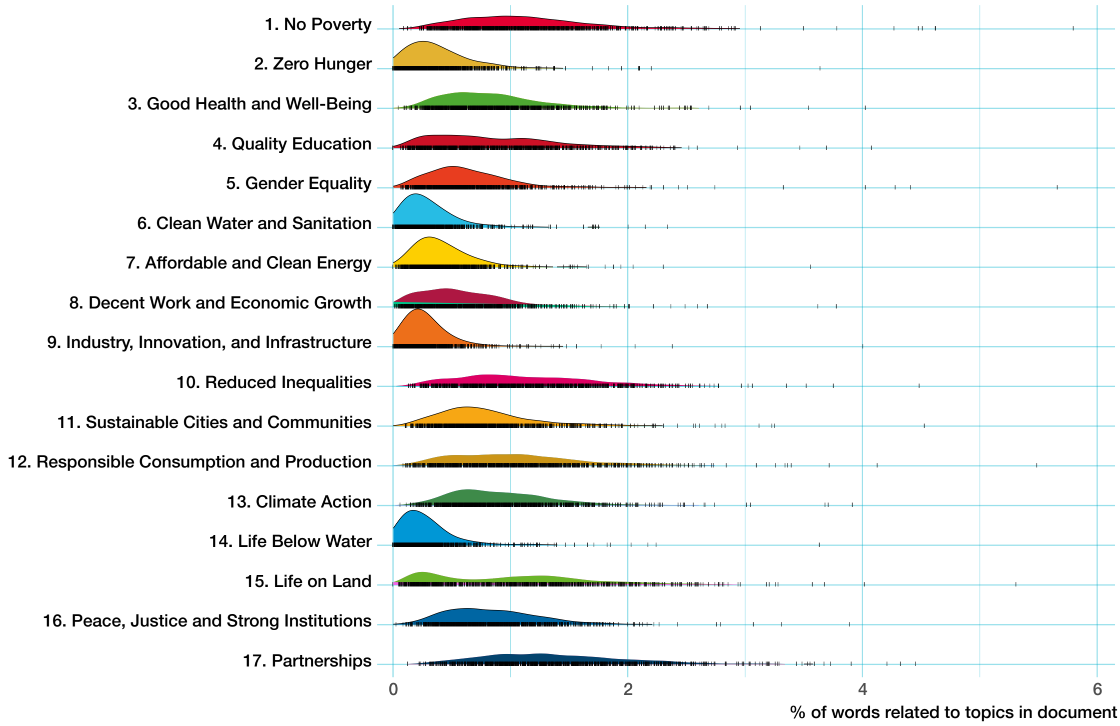


Figure 12 Distribution of Sustainable Development Goals in project documentation. Share of words associated to SDGs present in project documentation. Key terms taken from public definition and expanded by word2vec (2302 reports from 856 projects analysed).

The network graph presented in Figure 13 illustrates the relationship between projects and the SDGs. The graph shows a network where the projects and the SDGs are the nodes, with connections weighted by the share of words for each document in relation to the SDGs. Again, the force-directed algorithm used in constructing the network displays the spatialization of nodes, which maps the proximity and authority of themes in relation to each other (Jacomy *et al.* 2014). **The core of the network is composed of *Goal 2: Zero Hunger, Goal 16: Peace, Justice and Strong Institutions, Goal 11: Sustainable Cities and Communities, Goal 13: Climate Action and Goal 1: No Poverty.***

In the case of Goal 2, which is particularly relevant for IFAD given its experience in transforming rural areas, it is observed that while the intensive margins, as measured by the share of words in the documents, is not the largest, Zero Hunger is the widest spread among documents (ie. extensive margins) and is present in rich documents, which are those that cover many Goals. This means that when analysed in relation to the other SDGs, Zero Hunger is the goal covered most frequently in projects. The fact that it is the largest node in the network, placed in the core and close to other SDGs indicates that Goal 2 is often reported in projects that also address other issues, which reinforces the integration between the economic, social and environmental dimensions recognised by Agenda 2030.

Six clusters have also been identified by the modularity algorithm. Notably, the largest cluster is composed by Goal 3 – Good Health and Wellbeing, Goal 4 – Quality Education, Goal 5 – Gender Equality, and Goal 16 – Peace, Justice and Institutions. Analysis of these groups can indicate how strategies adopted by projects address the SDGs and the interconnectedness between them. The average degree of connectivity of the network is 6.64, which means that, on average, each project reports against 6 to 7 SDGs.

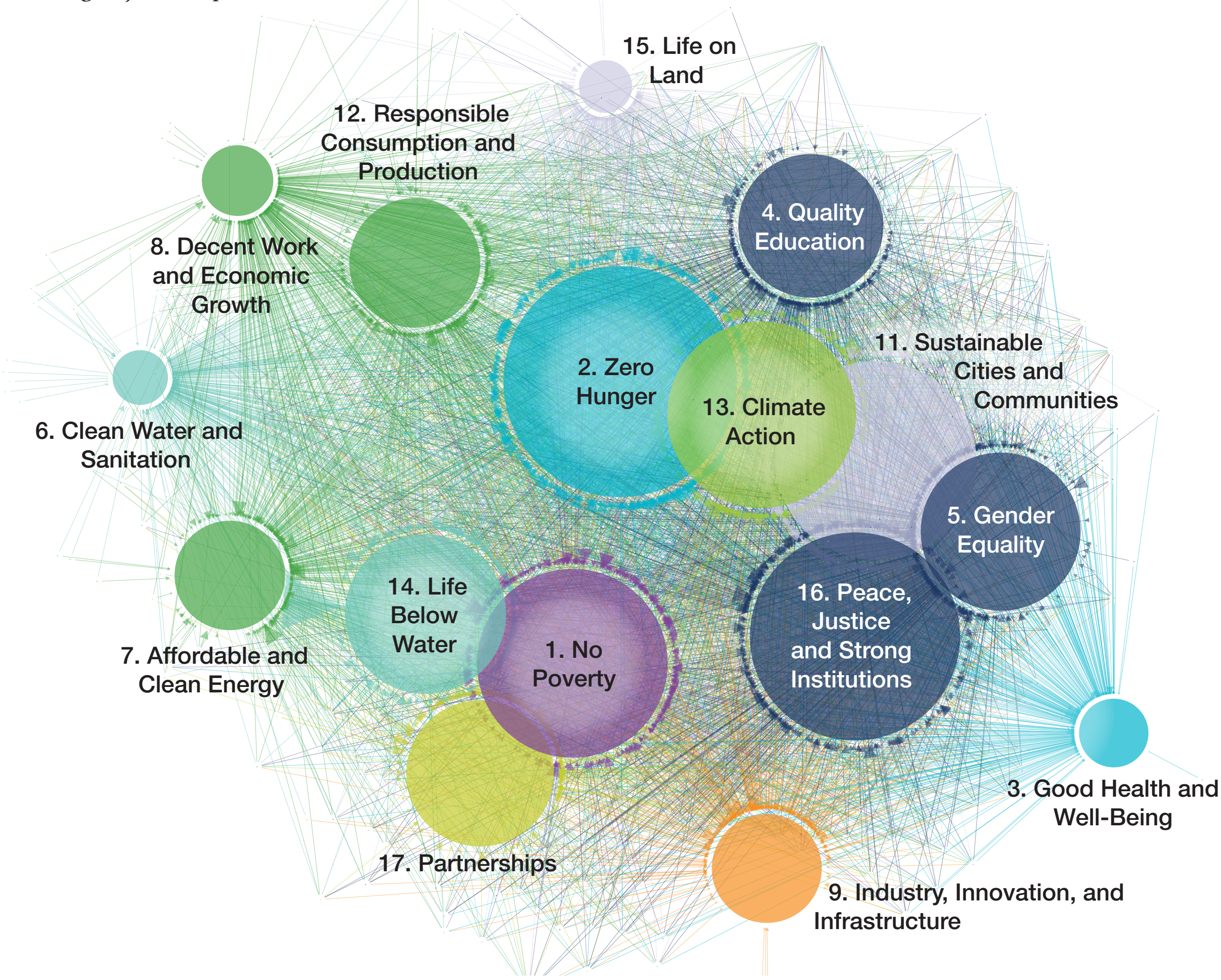


Figure 13 Network of Sustainable Development Goals detected in IFAD documentation. Parameters: force-directed graph, with node size partitioned as Weighed in-Degree, coloured by modularity class. 875 nodes (project IDs + SDGs), 5817 edges (weighed by share of words). Average degree: 6.64.

Taking the date of approval for each project, Figures 14 and 15 show the time trends for the Sustainable Development Goals. The trends indicate significant increases for all SDGs, which means that, in terms of the average presence of SDG-related content in project documentation, interest in addressing them has grown from 1981 to 2019. The whole distribution of the 17 SDGs across projects per decades is presented in Figures A9-A25 in Appendix V.

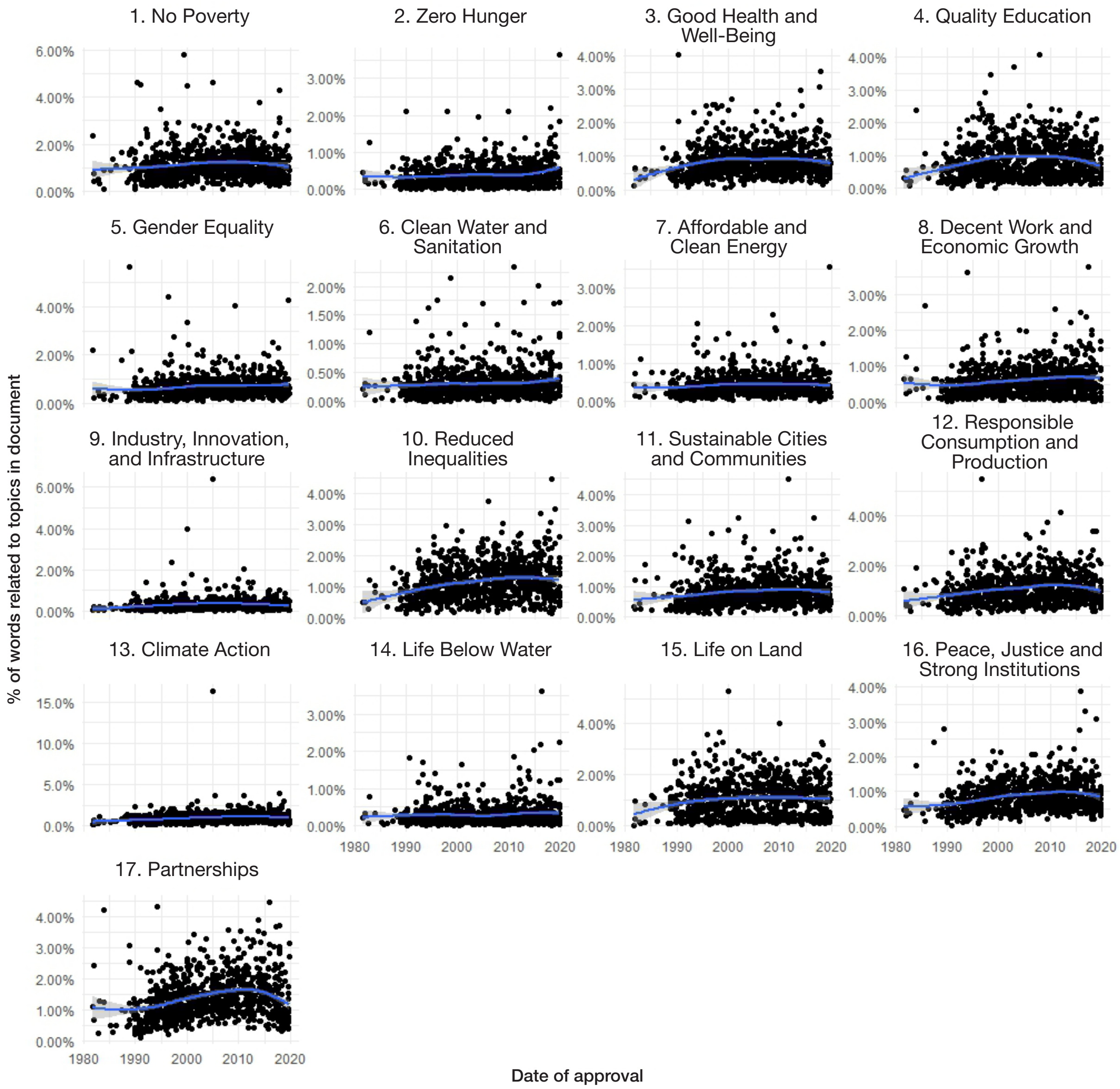


Figure 14 Time trend for the presence of Sustainable Development Goals in the IFAD documents by project approval date (1981-2019). The share of words associated to SDGs present in project documentation. Key terms taken from public definition and expanded by word2vec (2302 reports from 856 projects analysed).

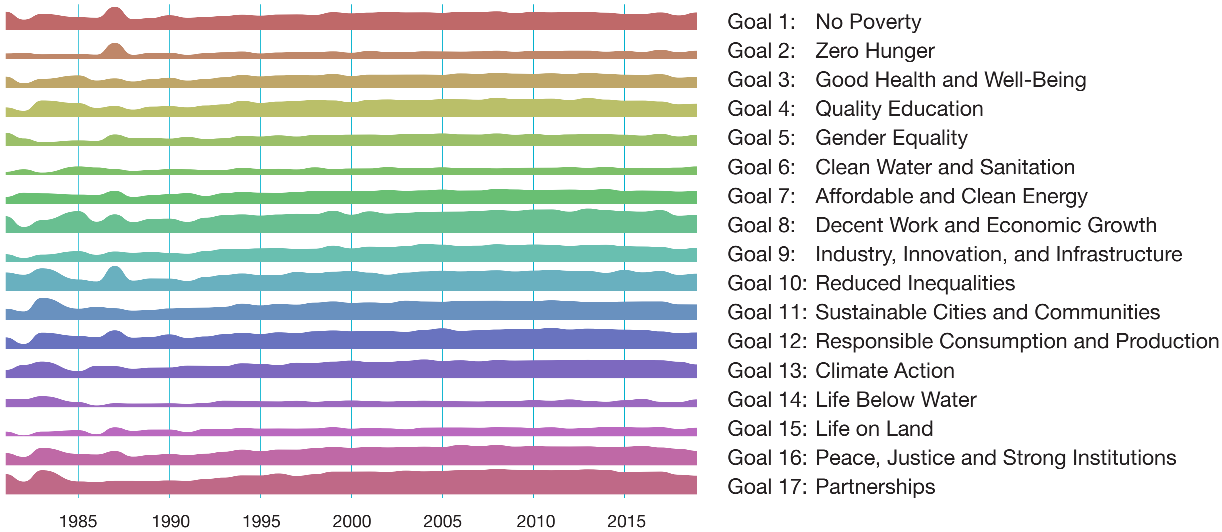


Figure 15 Time series for the presence of Sustainable Development Goals in IFAD documents. Share of words associated to SDGs present in project documentation, by year of approval (1981-2019). Key terms collected manually and expanded by word2vec (2302 reports from 856 projects analysed).

The presence of Sustainable Development Goals in IFAD documents weighed by project financing as estimated by equation is presented in Figure 16. Overall, the graph shows that a focus on addressing the SDGs in IFAD activities is increasing over time, both before and after 2015, the year Agenda 2030 entered into force.

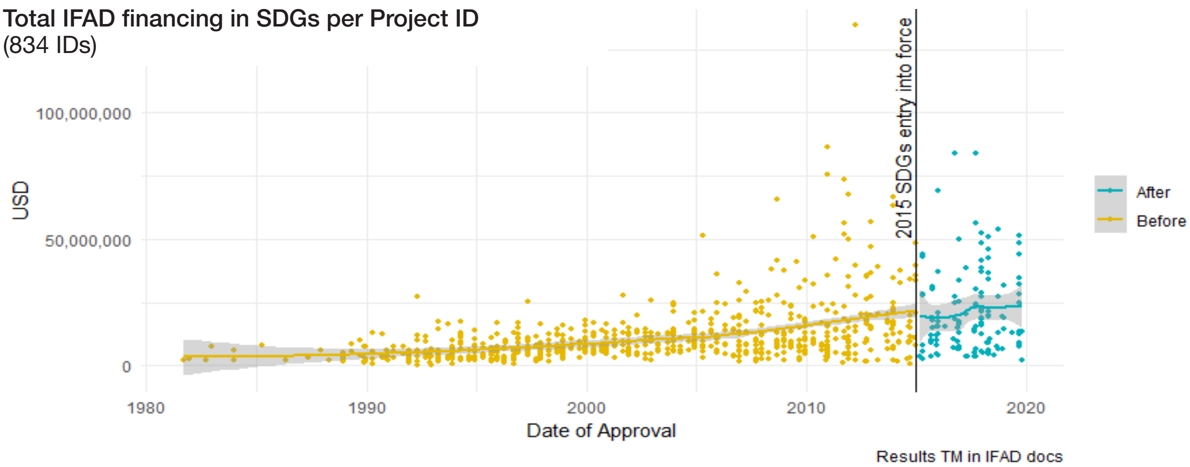


Figure 16 Time trend of the presence of Sustainable Development Goals in IFAD documents, weighted for total financed by project. Share of words associated to SDGs present in project documentation. Key terms taken from public definition and expanded by word2vec (2302 reports from 856 projects analysed).

Analysis of the association between the most frequent term for each goal (stemmed) is shown in Figure 17 and indicates some of the strategies adopted in IFAD projects to tackle the issues addressed by the SDGs. The first word representing Goal 1 – No Poverty is "income", which is then correlated to stem words that point to income generation. In the case of Goal 11 – Sustainable Cities and Communities, the stem word for "community" is correlated to several terms indicating involvement at the local level, which also reflects IFAD's participatory approach.

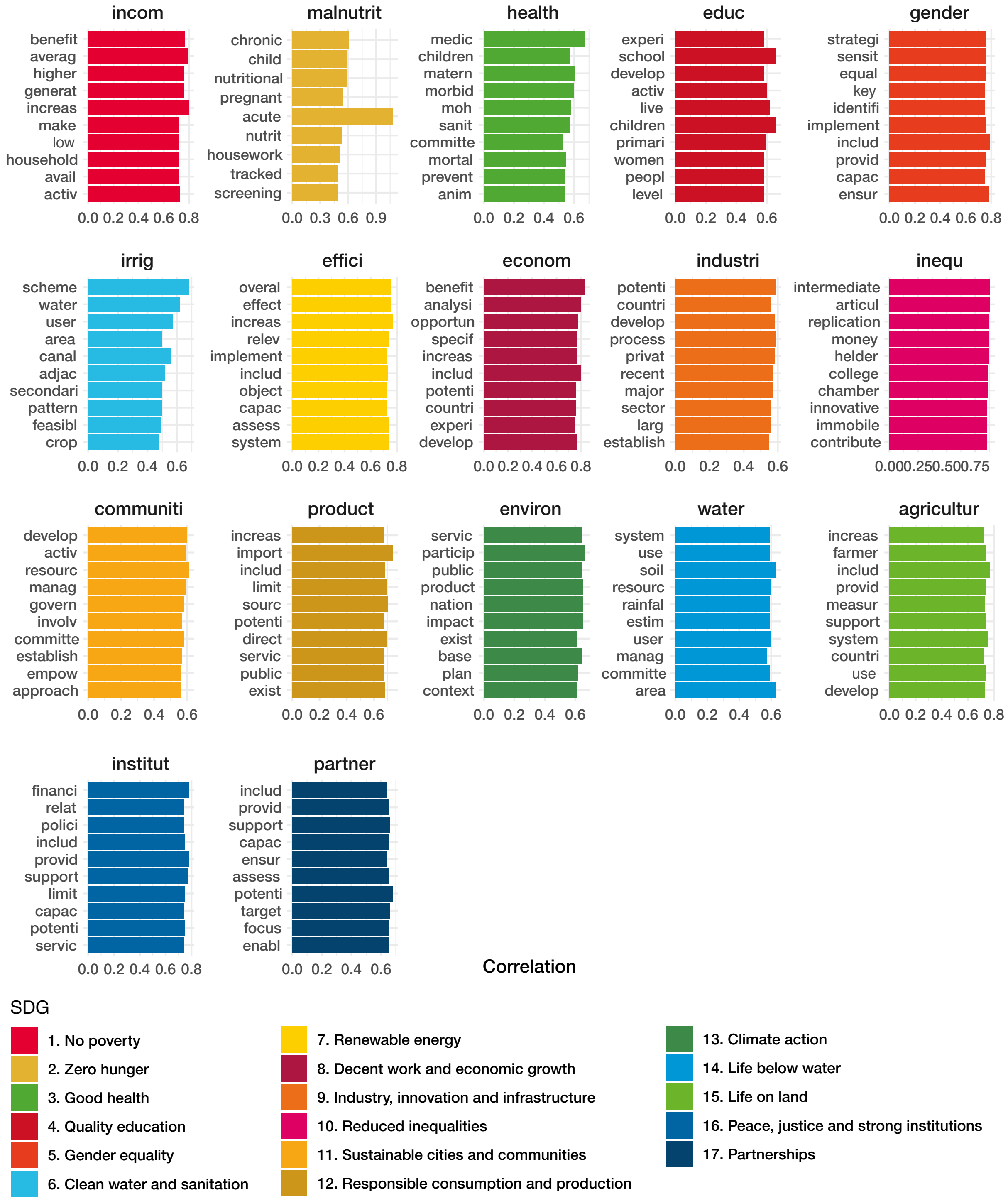


Figure 17 The first 10 words associated to the most frequent word for each of the 17 Sustainable Development Goals present in IFAD documents. Share of words associated to SDGs present in project documentation.
Key terms taken from public definition and expanded by word2vec (2302 reports from 856 projects analysed).

4.3. Detecting topics, interventions and outcomes of IFAD projects

The starting point for the analysis through Cornell University's agriculture-specific model was to find convergence between IFAD project classification and the Ceres2030 keyword and intervention mapping. To this extent, the sector classification[10] for IFAD projects was mapped against the Ceres2030 dataset in order to identify topics and interventions for each of the eight sectors, based on the trained model. Figure 18 presents the machine-driven characterization of IFAD-specific sectors of IFAD-specific sectors, which takes into account the existing IFAD sector classification and assigns interventions detected in the text through the existing algorithms.

Project documents were then processed through the model to classify them according to this typology, determined by weighting of keywords. While it is important to acknowledge the multi-disciplinary nature of IFAD projects, they were each characterized as falling into one primary sector based on the underlying interventions and topics. The distribution of sectors is shown in Figure 19, where "Rural Development" is the most prevalent sector, followed by "Agricultural Development" and "Credit and Financial Services".

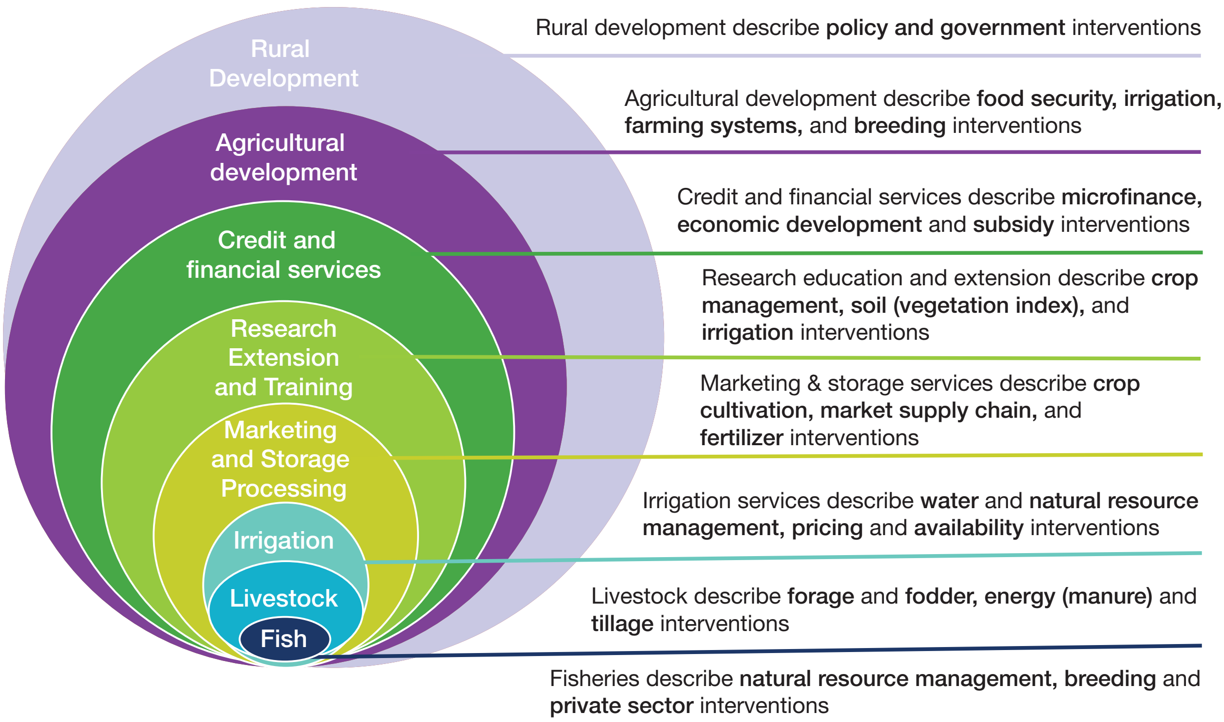


Figure 18 Project sector categories as identified against Ceres2030 model

10 Sectors are based on the first layer of IFAD's current project classification framework in the Grants and Investment Projects System (GRIPS) and include: Rural Development, Agricultural development, Credit and Financial Services, Research Extension and Training, Marketing and Storage Processing, Irrigation, Livestock and Fisheries.

Based on the sector classification, it was possible to determine related interventions, topics and outcomes. Regarding interventions, the model first detects 'synonyms' of interventions before classifying into three levels: general, common and specialized. The general and common categories were then re-evaluated based on the taxonomy established by the model, which is trained on more than half-million texts in agriculture from peer-review and grey literature. Overall, the model detected **2200 interventions** across project documentation. While more than 50% of them occur with a frequency of less than >1, making them null for this analysis, identifying them is still useful to track trends within the organisation. This analysis considered only interventions where the prevalence was greater than >50. Overall, the interventions were grouped into high level classes as shown in Figure 20. Socioeconomic interventions represented 37% of the dataset, within which finance and government-related interventions were the most frequently reported (Figure 21). Technology interventions followed very closely (36%) and comprised primarily of crop and irrigation-focused activities (Figure 22).

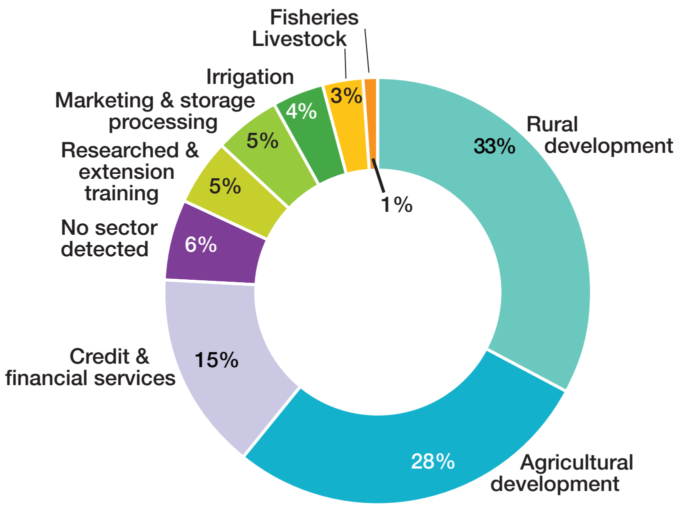


Figure 19 Project sector distribution detected by Ceres2030 model (743 reports from 743 projects analysed).

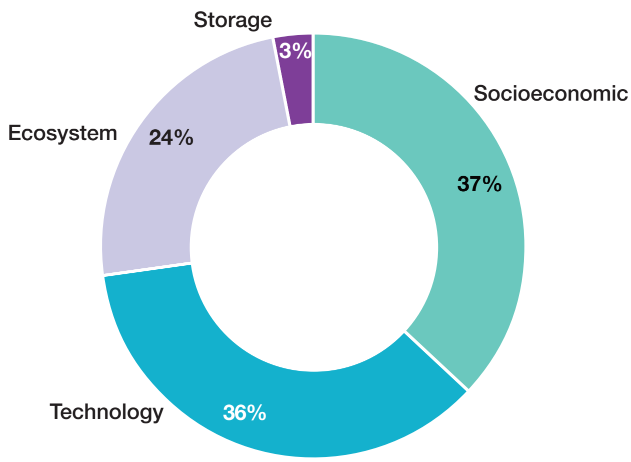


Figure 20 Intervention classes detected through the Ceres2030 model (743 reports from 743 projects analysed).

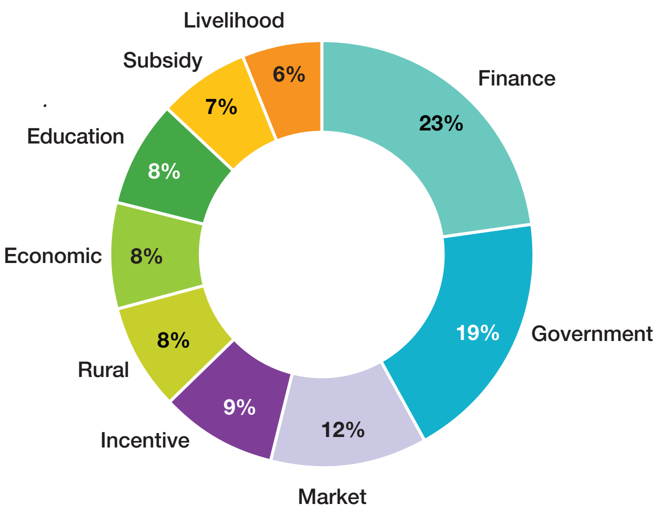


Figure 21 Overall socioeconomic interventions detected through the Ceres2030 model (743 reports from 743 projects analysed).

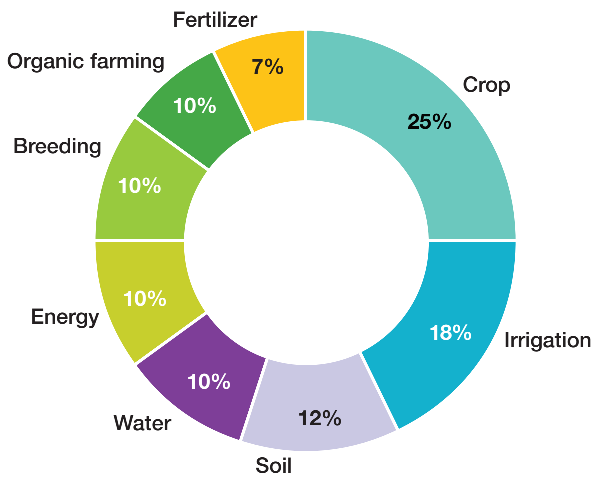


Figure 22 Overall Technology interventions detected through the Ceres2030 model (743 reports from 743 projects analysed).

Bringing the analysis to the sector level, figure 23 shows the main interventions detected for the top five sectors identified by the model, as well as whether any numerical data was associated with those interventions, which offer an indication of how they are being measured or monitored in the documentation. As the largest sector, Rural Development interventions were mostly focused on extension services, value chain integration and private sector development. Fertilizer use was the prevalent intervention detected in Agricultural Development projects. Projects within the Credit and Financial Services sector included interventions intended to reach smallholder farmers, small businesses and to provide access to credit. Marketing and Storage Processing projects were primarily focused on crop storage, whereas the top intervention for the Research and Extension Training sector reinforces IFAD's aim to increase market access by raising capacity of beneficiaries to market their products.

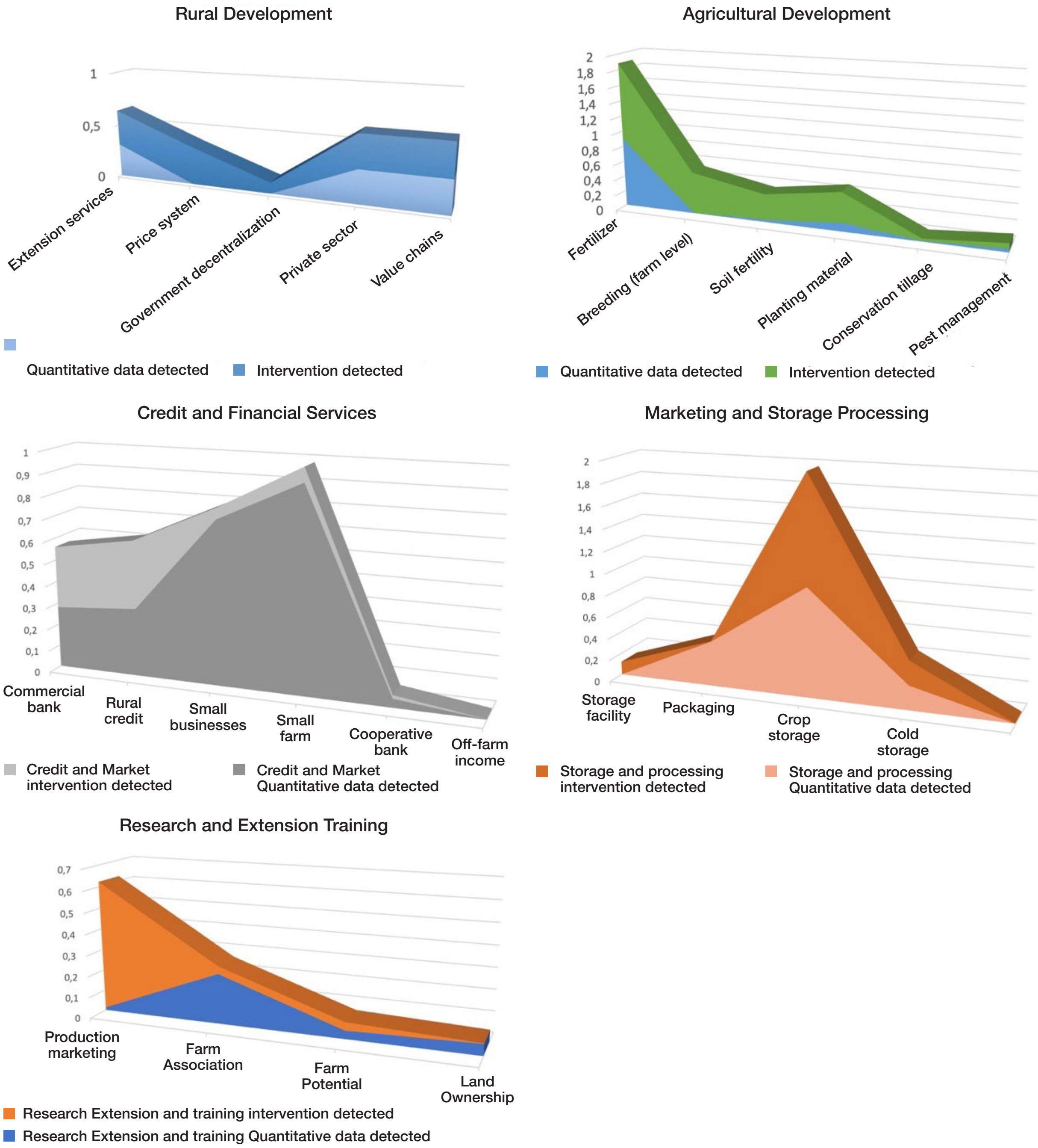


Figure 23 Main interventions detected through the Ceres2030 model for the top five sectors (743 reports from 743 projects analysed).

With regards to outcomes, unlike interventions, no synonyms are included. Once an outcome is detected, it is evaluated against a narrowly defined dictionary of outcomes based on the pre-trained Ceres2030 model. Ten outcome classes were detected in the documentation, as shown in Figures 24 and 25. **Expected outcomes relating to economic mobility (e.g livelihoods, such as income generation, employment and assets) and production were the most prominent overall and across the sectors, followed by improved water use, resilience and nutrition**. These findings are aligned with IFAD's strategic objectives of increasing small scale producers' productive capacities, benefits from market participation and economic mobility as well as strengthening the environmental sustainability and climate resilience of their economic activities.

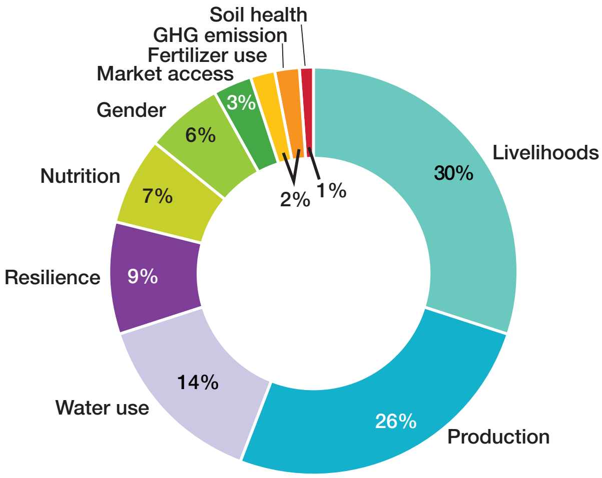


Figure 24 Overall outcomes detected through the Ceres2030 model (743 reports from 743 projects analysed).

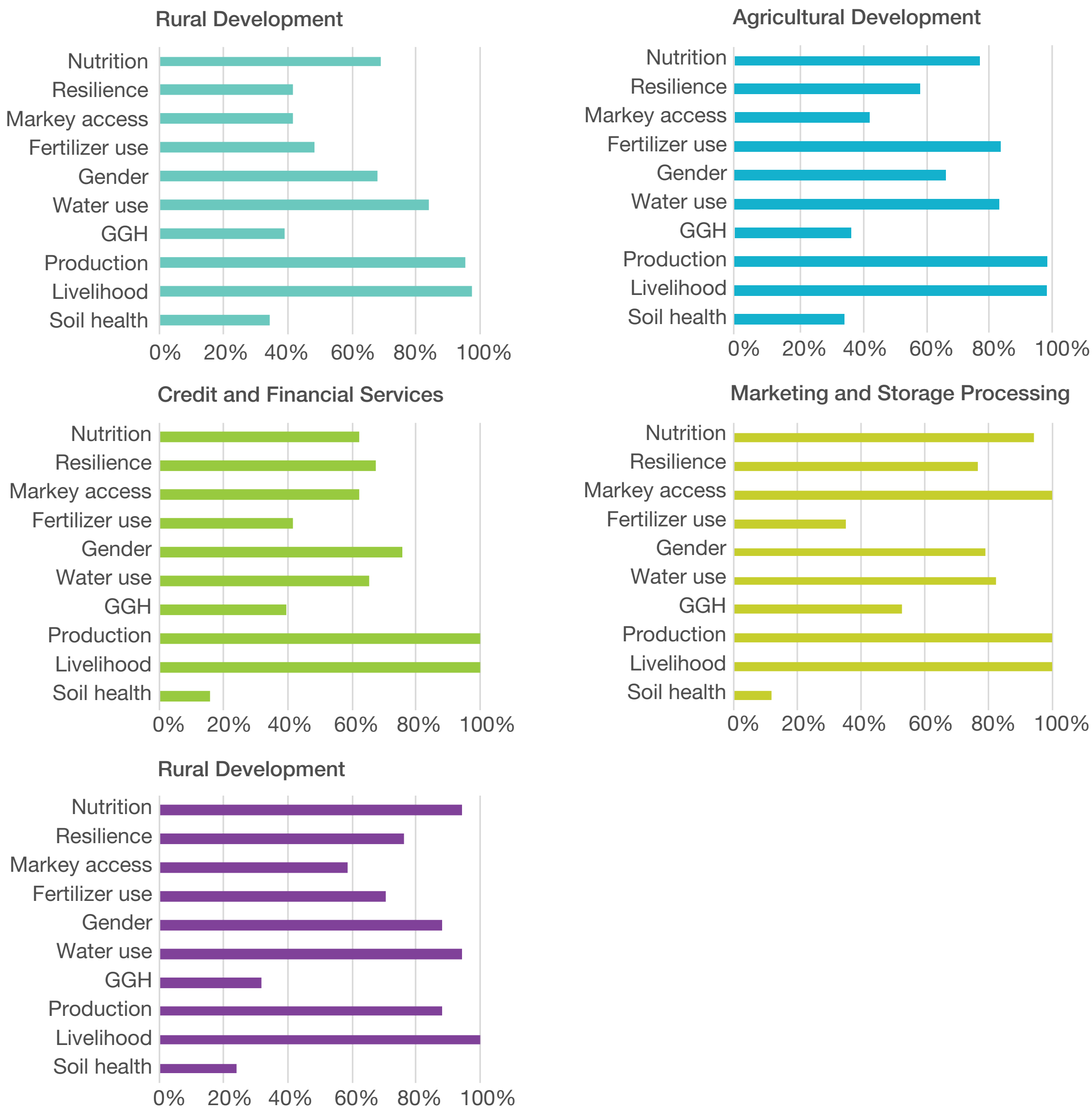


Figure 25 Outcomes detected through the Ceres2030 model for the top five sectors (743 reports from 743 projects analysed).

Lastly, themes are determined using a co-frequency analysis of words in the dataset. The words are then organized into topic models, which provide a sense of what is being discussed in the dataset. Figure 26 presents the main topics detected. Notably, the prevalent topic is closely related to IFAD's mission of working with governments to improve the economic mobility and livelihoods of rural people.

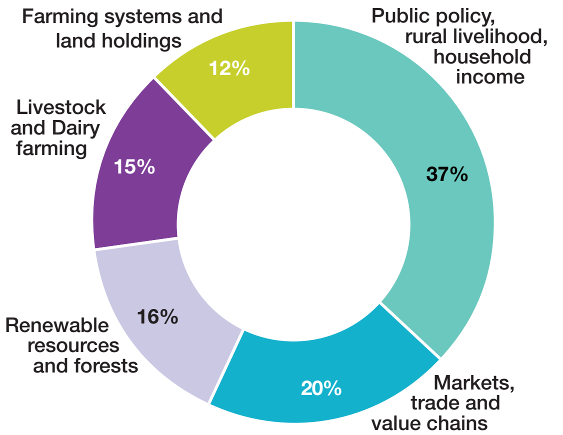


Figure 26 Topics detected through the Ceres2030 model across the documentation.

4.4. Predicting positive effects and determinants of project success

Based on the 17 impact assessments conducted on 19 IFAD10 projects, machine learning was used to predict the percentage of households obtaining a positive effect from the specific interventions implemented within the projects, conditional on their characteristics or features *X* (notably demographic characteristics including education and land ownership). The probability of positive impact of the policy/IFAD-supported intervention was estimated for each country and each outcome variable (income and resilience-related indicators). As the outcome variables were not all available for all countries, some countries present fewer outputs. Country-by-country results in terms of the estimation of the conditional distribution of the impact given a vector of features and percentage of households exhibiting a positive effect are presented in Appendix VI. Figures 27 and 28 show the results of this aggregation by country for $(\boldsymbol{x})$ and $\tau_1(\boldsymbol{x})$, respectively. They clearly show that Brazil, Indonesia, Kenya, Rwanda, and Senegal had a percentage of households obtaining a positive effect below 50%, whereas Tajikistan had the highest percentage of households obtaining a positive effect from the intervention, at almost 100%.

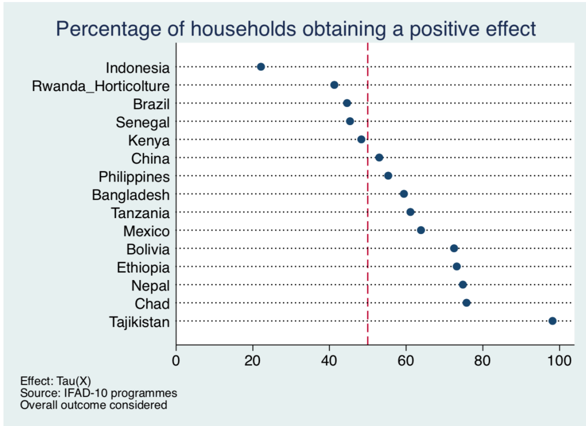


Figure 27 Overall effect aggregation by country. Results for τ ($\boldsymbol{x}$i).

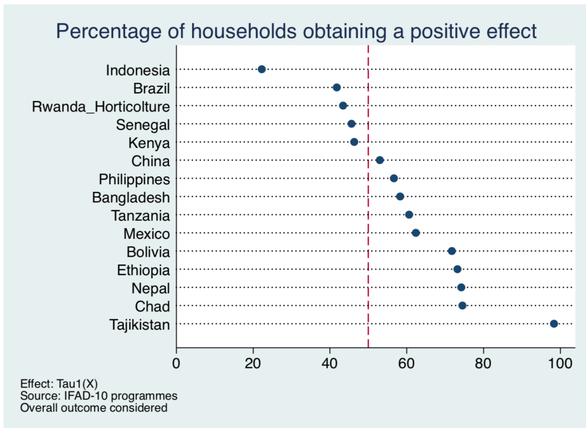


Figure 28 Overall effect aggregation by country. Results for $\tau_i(\boldsymbol{x}_i)$.

In order to identify the country- or project-specific determinants (characteristics) of the percentage of positive effect on country by country basis, a regression of the percentage of positive effect among the treated households at country level on project and country variables was performed. The project variables considered were "approval effectiveness" (days from approval to project entry into force), "Days from approval to first disbursement", the ratio between the financed and the approved amount, the total amount financed by IFAD divided by the country's population, and lastly, the project sector. Other country-level variables included in the model were the log of population size and the per capita GDP (PPP).

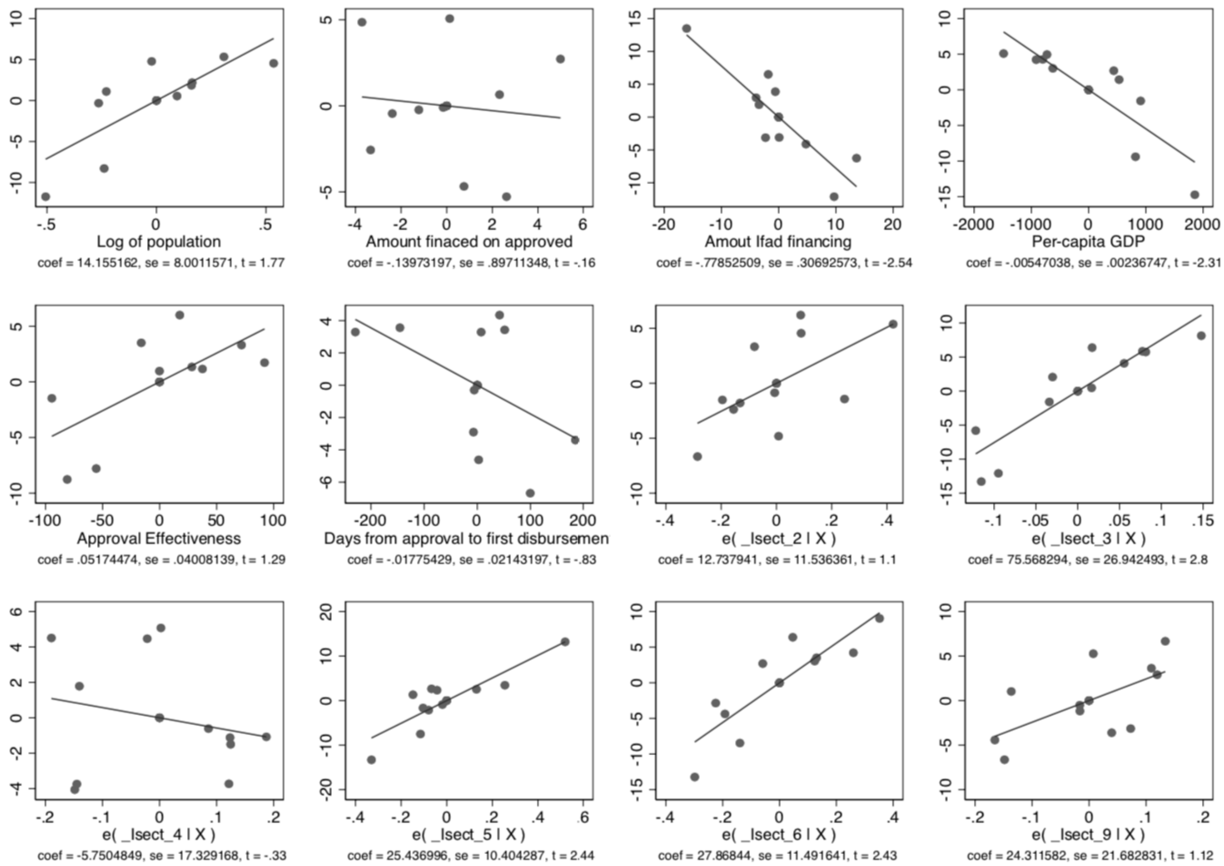


Figure 29 Project and country-level determinants of treated percentage of positive effect.

Figure 30 shows the results using data from 15 projects. The partial regression coefficients of each determinant are mildly significant (around 10%). However, it makes sense to comment on the slopes, which are reported in beta-coefficients. As shown in Table 8, significant variables having negative beta-coefficients are, by order of magnitude: "GDP (PPP) per capita" (-2.41), and "Ratio between the financed and the approved amounts" (-0.63). Variables that have a positive effect are "Days from approval to implementation" (3.00), and "Total country population" (1.81). Finally, among the intervention sectors, "Livestock" (2.31) and "Rural Development" (1.48) show the largest impact on percentage of positive effect, also significant at 5%. In lay terms, this means that IFAD is having a more positive impact in the least developed countries, and that cost-effective interventions seem to generate more impact on the ground.

Table 8 Beta coefficients of the determinants of percentage of positive effect for the treated.

	Beta
Population	1.15
Ratio of amount financed on amount approved	-0.04
Amount finance by Ifad	-0.93
Per-capita GDP (PPP)	-1.28
Days from approval to implementation	0.77
Days from approval to first disbursement	-0.13
Irrigation 2	*0.18*
Livestock 3,	*1.07*
Marketing/Storage/Processing 4	*-0.08*
Research/Extension/Training 5	*0.36*
Rural Development 6	*0.77*
Livestock and Rural Development	*0.34*

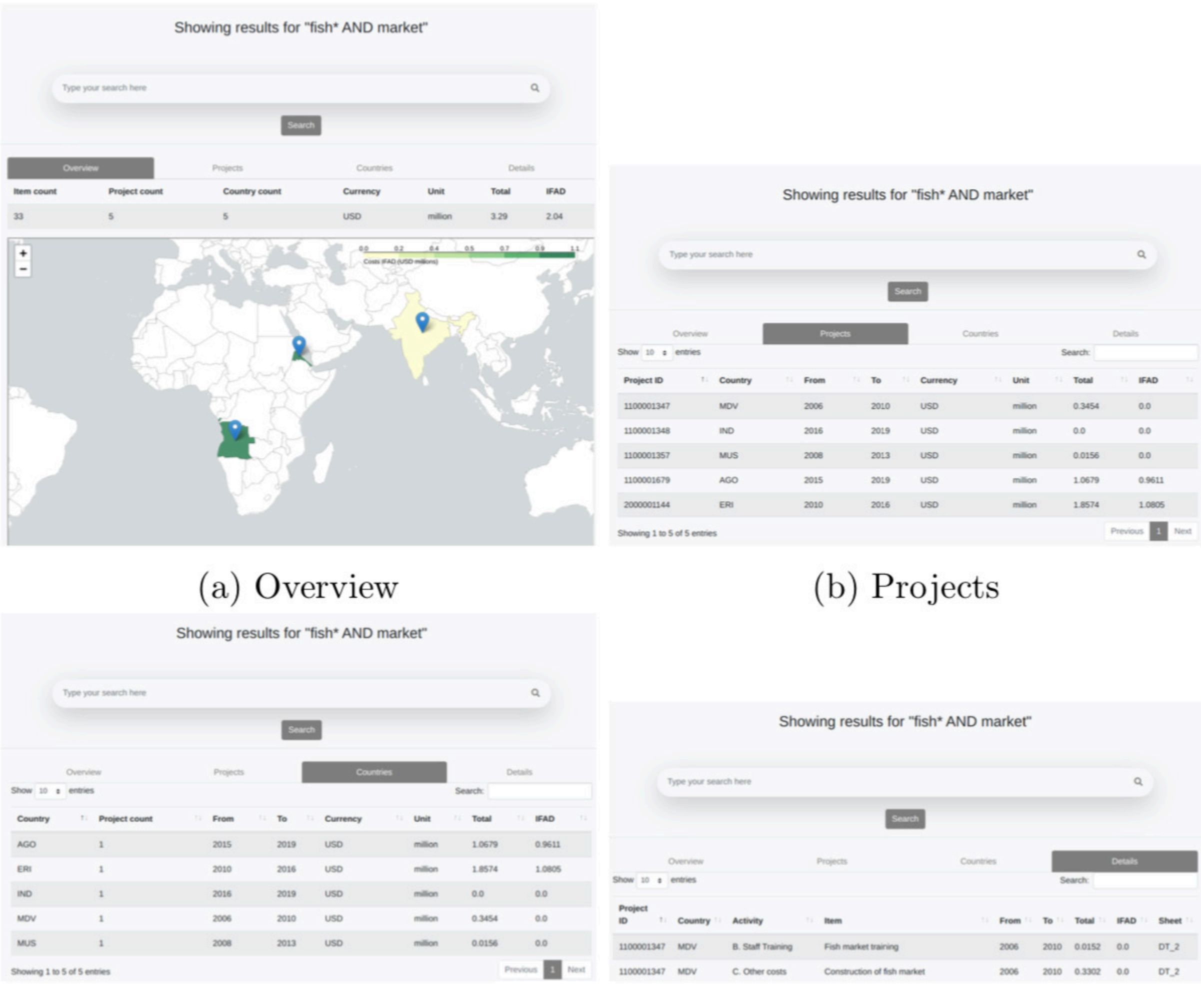
In sum, preliminary results from prediction algorithms show that, in the case of predicting the likelihood of positive impact, IFAD-supported interventions can generate positive impact for the large majority of beneficiaries. In the case of the project evaluated in Tajikistan, it was found that almost all beneficiaries had a positive impact across income and resilience outcomes. Percentage of households exhibiting positive impact ranged from 20% in the Indonesian sample to 100% in the Tajikistan sample across all income and resilience outcomes together.

4.5. Exploratory tools for project level cost-analysis

The consultations with IFAD technical experts identified a recurring need for a mechanism that could aggregate project expenditure to answer questions about what IFAD invests in, whether on more generalist themes –such as "how much does IFAD invest in extension training?" – or more specific issues – such as "how much does IFAD spend on rice-related activities?". Currently, experts rely on manually searching through project documents to take stock of the themes they are interested in investigating. Within this context, and also considering that the cost of project activities can represent a measure of importance of different interventions – sometimes even identifying the "real" focus of a particular project, as opposed to its intended objective, a part of this initiative's efforts concentrated on exploring ways to analyse investments at the project-level. The outputs include two experimental tools: a parser app built to standardise data extracted from project cost tables, and a user-friendly search engine created to display aggregated analysis and visualizations of project costs.

Based on the approach proposed in the thematic paper for the new Categorization Framework (IFAD, 2019a), project cost tables (COSTABs) were selected as the primary source for this exercise. COSTABs are semi-standardised excel files containing all the investment information about a project, and in particular all the activities involved within it. The first challenge was creating a standardised database from these files: despite them being generated with the aid of a dedicated software, resulting tables can be quite different from each other, thus requiring an extraction algorithm that not only recognised the information within the files, but also provided a standardised output from all documents. A Python command line program – the "Costaparse" – was developed to extract relevant data from COSTAB excel files, such as activity descriptions and their associated cost by financier. Currently, the program has only been tested under Linux OS and extracts the following information:

- file and sheet
- country
- project number
- starting year
- ending year
- activity
- item
- unit
- quantities
- cost type (investment, recurrent, financial)
- base costs
- costs including contingencies
- IFAD costs (loan, grants)
- GEF costs
- ASAP costs
- account of the item (code-like)
- currency and unit



(a) Overview (b) Projects

(c) Countries (d) Details

Figure 30 Example of results for query “fish* AND market”

A total of 335 COSTABs were processed through this program, related to investments on 101 different countries, from 1997 to 2020. The output was a spreadsheet containing more than 70,000 standardised project activities/items as described above. In order to leverage on the information compiled through the parser, a user-friendly web app was created to operate as a search engine within COSTAB items. It allows for text-based queries into the "activities" and "items" described in the COSTAB database, calculating the related costs. Figure 31 shows an example. Search operators include "AND" and "OR" searches, stem word search "*" and exact phrase search as exemplified below:

- fish* = will show results for any words that begin with "fish" (fisheries, fishing)
- fish AND equipment = will show results that contain the two words
- fish OR livestock = will show results that contain either word
- "fish market" = will show results with the exact phrase within the quotations

Results of the queries are presented through various tabs, including:

- An overview with the sum of investments and a map for easy visualisation
- A list of the sum of investments by projects
- A list of the sum of investments by country
- And the detailed costs item by item

A few limitations of this tool should be highlighted. First, that as a text-based search, the specific term must be present in the body of text to appear in the results. This means that activities related to a term, but that do not directly reference it may be overlooked. Another key limitation is that reports are written in four languages, which requires that multi-language dictionaries/translators be incorporated into the program. Finally, the scope of the results is also dependent on the availability of COSTABs for all projects. Nevertheless, if these limitations are resolved and the search parameters are refined, this tool presents a potential to leverage on project cost tables as a source of aggregated investment-related information.

5. Policy implications

From a corporate perspective, as IFAD's Strategic Framework 2016-2025 highlights the centrality of ICT and encourages expanding the uptake of new technologies, this innovation boosted the creation of tools to enhance knowledge management and supports IFAD's ICT4D strategy by proposing an **integrated, machine-driven approach to analyse project documentation and predict impact.** It also addressed concerns from the knowledge management strategy regarding the fragmentation of information, by showing how new technologies help leverage existing data sources to answer new questions. Text mining turned masses of unstructured text from IFAD documents into structured data, which was then analysed for trends, integrated with other data sources, and incorporated into machine learning models. Household data from impact assessments were also repurposed for the development of a prediction model that is able to detect the extent of positive impact based on project and household level features. Such an approach not only enables IFAD to gain further insight into its data, but also delivers added value as it does not require collecting new information.

Additionally, machine learning can play a key role in supporting the advancement of IFAD11's commitments and targets by accelerating project-level data analysis, thus **enabling more regular reporting on the mainstreaming themes and the SDGs against the strategic outcomes.** Regarding impact, machine learning dramatically shortened the time necessary to carry out a systematic review and meta-analysis. Within the timeframe of the project – four months – the systematic review of livestock interventions assessed more than 20,000 citations. By harnessing internal and external evidence of which interventions deliver the most effective results, not only can IFAD strengthen its own project design, but also disseminate learning to other international institutions.

6. Lessons learned

A key lesson from this initiative is that effective machine learning is an iterative process that requires large amounts of data and time to explore various approaches.

Data collection was a major challenge, as there is not a central location from which to extract all the documentation and the documents were inconsistent. Much more time was spent in this initial task, compromising the timeframe of subsequent deliverables. Once files were downloaded, they were checked to ensure they were in text-searchable PDF format (i.e. files were not images) – and to this end, a program was developed to search the project's database to recognize and digitize text present in hundreds of image files, which demanded high processing power and time.

With regards to preparing the data for analysis, project reports were found to contain a lot of extraneous information, which makes the data noisy and sometimes difficult to isolate what has really taken place. Additionally, as reports are written in English, French, Spanish and Portuguese, all analysis had to consider multi-language models. This had implications for the data processing at Cornell University, as the 'real' server costs for this analysis exceeded the budget and amounted to more than $15,000 USD due to the size of the dataset and the multiple re-processing of documents using cloud-based models. Some of this cost could be curtailed with better methods of detection.

While for some aspects of machine learning the size of our dataset was sufficient, namely for text mining and topic modelling, the predictive analytics can benefit from more data, particularly if it is extended to the entire programme of work of IFAD. On the other hand, predicting likelihood of positive impact at the household level relied on impact assessment data from 19 projects encompassing around 40,000 household-level observations, hence providing enough data to achieve the required prediction performance.

Nevertheless, the applications for machine learning employed in this project can greatly support IFAD's development effectiveness framework, especially regarding accelerating knowledge generation, improving efficiency in corporate reporting, and building an evidence base to inform policy and the design of successful projects.

7. Recommended next steps

The data has been parsed, cleaned and is available for further analysis, and as such this initiative should be refined and scaled up. In the first case, more time and further data collection are necessary in order to improve the categorization of interventions and the prediction algorithms. More questions can also be explored, such as trends in financing, and other strategic areas that projects report on and other strategic areas that projects report on. The Ceres2030 model can also be customized further to better suit IFAD's project classification framework.

The short span of the project meant there was little time to collect and prepare the datasets, as well as to carry out pre- and post- tests, and to explore different models. Consequently, while a strong conceptual framework for project performance prediction was established, more time is required to clean the data and to improve the technicalities of the model, as well as to integrate the various sources and tools the project has created.

However, as the methods employed can (and should) process large amounts of data, this effort can be refined by adding more project-level documentation, eventually extending it to the entire IFAD programme of work (including grants) and to IFAD9 impact assessments, which would not only deliver more consistent insights, but also help calibrate the prediction models. The envisioned final product is a user-friendly dashboard that integrates the predictive analytics with data search and visualization features.

This was an exciting project that capitalised on existing data to uncover new patterns and gain additional knowledge. For IFAD to improve its focus on results, strengthen mechanisms for successful project design, and become a leader on measuring and attributing impact against the SDGs, machine learning and other artificial intelligence approaches must be used not only in an experimental fashion through venues such as the Innovation Challenge, but mainstreamed into IFAD's everyday work. More staff resources, centralised in a dedicated unit at IFAD headquarters, with the right skill set to develop and employ these technologies according to IFAD's corporate objectives and to analyse and systematize findings in ways that are useful for project teams, would further place IFAD as a world class institution capable of learning from its work and ensuring excellent and efficient outcomes.

References

Balint, T., Higgins, D., Mallia, P., Scalzo, S., Winters, P. (2017) Disbursement performance of the International Fund for Agricultural Development (IFAD): An in-depth analysis of drivers and trends. *IFAD Research Series*, *14*.

Blondel, V.D., Guillaume, J., Lambiotte, R., Lefebvre, E. (2008) Fast Unfolding of Communities in Large Networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(P10008), 1-12.

Bojanowski, P., Grave, E., Joulin, A., Mikolov, T. (2017) Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, 5, 135-146.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., Rothstein, H. R. (2009) *Introduction to Meta-Analysis*. Chichester, UK: John Wiley & Sons, Ltd.

boyd, d., Crawford, K. (2012) Critical Questions for Big Data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15(5), 662–679.

Einav, L., Levin, J. (2014) The data revolution and economic analysis. *Innovation Policy and the Economy*, 14(1), 1-24.

FAO (2009) *The state of food and agriculture: Livestock in the balance*. Rome, Italy: FAO.

FAO (2018) *World Livestock: Transforming the livestock sector through the Sustainable Development Goals.* Rome, Italy: FAO.

Feinerer, I., Hornik, K., Meyer, D. (2008) Text Mining Infrastructure in R. *Journal of Statistical Software,* 25(5): 1-54*.*

Feinerer, I., Hornik, K. (2018) tm: Text Mining Package. *R package version 0.7-5.* https://CRAN.R-project.org/package=tm

Garbero, A. (2016). Measuring IFAD's impact: Background paper to the IFAD9 Impact Assessment Initiative, *IFAD Research Series*, *07*.

Garbero, A., Marion, P., Brailovskaya, V. (2018) *The impact of the adoption of CGIAR's improved varieties on poverty and welfare outcomes – A systematic review*. Rome, Italy: International Fund for Agricultural Development.

Giannone, D., Reichlin, L., Small, D. (2008) Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, 55(4), 665-676.

Hák, T., Janoušková, S., Moldan, B. (2016) Sustainable Development Goals: A need for relevant indicators. *Ecological Indicators*, 60, 565-573.

Hornik,. K, Mair, P., Rauch, J., Geiger, W., Buchta, C., Feinerer, I. (2013) The textcat Package for n-Gram Based Text Categorization in R." *Journal of Statistical Software*, 52(6), 1-17.

IFAD (2019a), *A new categorization framework for IFAD-supported project interventions*, Draft Thematic Paper, Strategy and Knowledge Department. Rome, Italy: International Fund for Agricultural Development.

IFAD (2019b), *IFAD 2.0: The Way Forward,* Discussion paper, Follow-up session to the fifth Executive Board Retreat Rome, 24-25 October 2019. Rome, Italy: International Fund for Agricultural Development.

Jacomy, M., Venturini, T., Heymann, S., Bastian, M. (2014). ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software. *PLoS ONE* 9(6):e98679.

Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, 1(1), 1-12.

Laney, D. (2001) 3D data management: Controlling data volume, velocity and variety. *META group research note*, 6(70), 1.

Livestock in Development (1999) *Livestock in poverty-focused development*. Crewkerne: Livestock In Development.

Lupton, D. (2012) *Digital sociology: An introduction*. Working Paper. Sydney: University of Sydney.

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., Barton, D. (2012) Big data: the management revolution. *Harvard business review*, 90(10), 60-68.

Mikolov, T., Chen, K., Corrado, G., Dean, J. (2013) Efficient estimation of word representations in vector space. *arXiv preprint,* arXiv:1301.3781.

Miola, A., Schiltz, F. (2019) Measuring sustainable development goals performance: How to monitor policy action in the 2030 Agenda implementation? *Ecological Economics*, 164, 106373.

Ooms, J. (2018). pdftools: Text Extraction, Rendering and Converting of PDF Documents. *R package version 1.8*. https://CRAN.R-project.org/package=pdftools

Polley, Eric C. and van der Laan, Mark J. (2010) "Super Learner In Prediction" *U.C. Berkeley Division of Biostatistics Working Paper Series,* 266.

Python Software Foundation (2019) *Python Language Reference, version 2.7.* Available at http://www.python.org

R Core Team (2019). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria.

Randolph, T. F., Schelling, E., Grace, D., Nicholson, C. F., Leroy, J. L., Cole, D. C., et al. (2007). Role of livestock in human nutrition and health for poverty reduction in developing countries. *Journal of Animal Science*, *85*(11), 2788–2800.

Resce, G., & Maynard, D. (2018). What matters most to people around the world? Retrieving Better Life Index priorities on Twitter. *Technological Forecasting and Social Change,* 137, 61-75.

Waddington, H., Snilstveit, B., Hombrados, J. G., Vojtkova, M., Anderson, J., White, H. (2014). *Farmer Field Schools for Improving Farming Practices and Farmer Outcomes in Low- and Middle-income Countries: A Systematic Review*. London, UK: International Initiative for Impact Evaluation.

Wickham, H., Henry, L. (2018). tidyr: Easily Tidy Data with 'spread()' and 'gather()' Functions. *R package version 0.8.1.* https://CRAN.R-project.org/package=tidyr

Zilberman, D., Gordon, B., Hochman, G., Wesseler, J. (2018) Economics of sustainable development and the bioeconomy. *Applied Economic Perspectives and Policy*, 40(1), 22-37.

Appendices

Appendix I – Descriptive text mining equations

Equation 1 – Presence of a cross-cutting issue in a document

Taking only unique words for each $\mathbf{Issue_i}$ (where issue represents the cross-cutting issue)[11] the presence of a particular cross-cutting issue in a $\mathbf{Document_j}$ was estimated as follows:

$$(1) \qquad \%\,\mathrm{Issue}_{ij} = \left[\frac{\sum\left(\mathrm{Words}_{\mathrm{Document}_j} \in \mathrm{Words}_{\mathrm{Issue}_i}\right)}{\sum \mathrm{Words}_{\mathrm{Document}_j}}\right] \times 100\,,\ i = 1, \ldots, 8$$

Where $\sum \mathbf{Words_{Document_j}}$ is the total number of words in document $\mathbf{j}$ and $\sum\left(\mathbf{Words_{Document_j}} \in \mathbf{Words_{Issue_i}}\right)$ is the sum of words belonging to the cross-cutting issue $\mathbf{i}$ detected in document $\mathbf{j}$. By construction, $\mathbf{\%\,Issue_{ij}}$ is between 0 and 100 and it is a measure of the extent that $\mathbf{Issue_i}$ is present in $\mathbf{Document_j}$.

Equations 2-4 – Relational analysis of projects and cross-cutting issues

An issue is considered connected to a project as follows:

$$(2) \qquad \mathrm{Issue}_i \longmapsto \mathrm{Proj}_{\mathrm{ID}_j} \text{ if } \%\,\mathrm{Issue}_{i\mathrm{ID}_j} > \mathrm{Average}_{\mathrm{ID}}\left(\%\,\mathrm{Issue}_{i\mathrm{ID}_j}\right)$$

Where, $\mathbf{Issue_i}$ is connected to $\mathbf{Proj_{ID_j}}$ if it is present in more than its average presence among all project IDs. Secondly, the network is analysed by two matrices:

$$(3) \qquad \mathrm{M}^{\mathrm{A}}(\mathrm{i},\mathrm{j}) = \begin{Bmatrix} 1 & \text{if} & \mathrm{Issue}_i \rightarrow \mathrm{ID}_j \\ 0 & \text{if} & \mathrm{Issue}_i \nrightarrow \mathrm{ID}_j \end{Bmatrix}$$

which considers only the presence or absence (0/1) of an issue in a project, and:

$$(4) \qquad \mathrm{M}^{\mathrm{B}}(\mathrm{i},\mathrm{j}) = \begin{Bmatrix} \%\,\mathrm{Issue}_{i\mathrm{ID}_j} & \text{if} & \mathrm{Issue}_i \rightarrow \mathrm{ID}_j \\ 0 & \text{if} & \mathrm{Issue}_i \nrightarrow \mathrm{ID}_j \end{Bmatrix}$$

which also accounts for the weight of the connection ($\mathbf{\%\,Issue_{iID_j}}$). In sum, whereas in $\mathbf{M^A(i,j)}$ the presence or absence of an issue in a project is expressed by dummies, in $\mathbf{M^B(i,j)}$ the presence of an issue in a project is expressed by the percentage of words related to that particular issue. Therefore, $\mathbf{M^B(i,j)}$ allows to capture both the extensive and the intensive margins of the spread of issues among projects.

Equation 5 – Cross-cutting issues weighted for IFAD-financing per project

To summarise information across projects, the study considered both $\mathbf{\%\,Issue_{iID_j}}$, and $\mathbf{\%\,FWIssue_{iID_j}}$ which is $\mathbf{\%\,Issue_{iID_j}}$ weighted for the total amount of money that IFAD invests in a project, as follows:

$$(5) \qquad \%\,\mathrm{FWIssue}_{i\mathrm{ID}_j} = \%\,\mathrm{Issue}_{i\mathrm{ID}_j} \times \text{Total Financed}_{\mathrm{ID}}$$

$\mathbf{\%\,FWIssue_{iID_j}}$ is a measure that takes into account both the presence of an issue detected in the documentation for a particular project and the amount of money that IFAD invested in that project.

11 Words present in more than one category were excluded.

Equation 6 – Presence of SDGs in a document

Taking only unique words for each $\mathbf{SDG_i}$ (words present in more than one SDG were removed) the presence of the $\mathbf{SDG_i}$ – related words in the $\mathbf{Document_j}$ is estimated as follows:

$$(6) \qquad \% \, SDG_{ij} = \left[\frac{\sum \left(Words_{Document_j} \in Words_{SDG_i} \right)}{\sum Words_{Document_j}} \right] \times 100 \, , \; i = 1, \dots, 17$$

Where $\sum Words_{Document_j}$ is the total number of words present in document j and $\sum \left(Words_{Document_j} \in Words_{SDG_i} \right)$ is the sum of words associated to SDG_i detected in document j. $\% \, SDG_{ij}$ is between 0 and 100 and it is a measure of the extent that SDG_i is present in $Document_j$.

Equations 7-9 – Relational analysis of projects and SDGs

An SDG_i is considered connected to a project as follows:

$$(7) \qquad SDG_i \longmapsto Proj_{ID_j} \; if \;\; \% \, SDG_{iID_j} > Average_{ID} \left(\% \, SDG_{iID_j} \right)$$

Secondly, the network is analysed by two matrices:

$$(8) \qquad M^A(i,j) = \begin{Bmatrix} 1 & if & SDG_i \rightarrow ID_j \\ 0 & if & SDG_i \nrightarrow ID_j \end{Bmatrix}$$

which only consider the presence or absence (0/1) of SDGs in a project, and:

$$(9) \qquad M^B(i,j) = \begin{Bmatrix} \% \, SDG_{iID_j} & if & SDG_i \rightarrow ID_j \\ 0 & if & SDG_i \nrightarrow ID_j \end{Bmatrix}$$

which also takes into account the weight of the connection ($\% \, SDG_{iID_j}$).

Equation 10 – SDGs weighted for IFAD-financing per project

To summarise information across projects both $\% \, SDG_{iID_j}$, and $\% \, FWSDG_{iID_j}$ were used, where the latter is $\% \, SDG_{iID_j}$ weighted for the total amount of money that IFAD invests in a project, as follows:

$$(10) \qquad \% \, FWSDG_{iID_j} = \% \, Issue_{iID_j} \times Total \; Financed_{ID}$$

$\% \, FWSDG_{iID_j}$ is a measure that takes into account both the presence of a SDG in the documentation for a particular project and the amount of money that IFAD invested on that project.

Appendix II – Systematic review: list of countries included

Afghanistan, Albania, Algeria, Angola, Argentina, Armenia, Azerbaijan, Bangladesh, Belize, Benin, Bhutan, Bolivia (Plurinational State of), Bosnia and Herzegovina, Botswana, Brazil, Burkina, Faso, Burundi, Cabo Verde, Cambodia, Cameroon, Central African Republic, Chad, Chile, China, Colombia, Comoros, Congo, Costa Rica, Côte d'Ivoire, Cuba, Cyprus, Democratic People's Republic of Korea, Democratic Republic of the Congo, Djibouti, Dominica, Dominican Republic, Ecuador, Egypt, El Salvador, Equatorial Guinea, Eritrea, Eswatini, Ethiopia, Fiji, Gabon, Gambia (The), Georgia, Ghana, Grenada, Guatemala, Guinea, Guinea-Bissau, Guyana, Haiti, Honduras, India, Indonesia, Iraq, Jamaica, Jordan, Kenya, Kiribati, Kyrgyzstan, Lao People's Democratic Republic, Lebanon, Lesotho, Liberia, Madagascar, Malawi, Maldives, Mali, Mauritania, Mauritius, Mexico, Mongolia, Montenegro, Morocco, Mozambique, Myanmar, Namibia, Nepal, Nicaragua, Niger, Nigeria, North Macedonia, Pakistan, Palestine, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Republic of Moldova, Romania, Rwanda, Saint Lucia, Saint Vincent and the Grenadines, Samoa, Sao Tome and Principe, Senegal, Seychelles, Sierra Leone, Solomon Islands, Somalia, South Sudan, Sri Lanka, Sudan, Suriname, Syrian Arab Republic, Tajikistan, Thailand, Timor-Leste, Togo, Tonga, Tunisia, Turkey, Uganda, United Republic of Tanzania, Uruguay, Uzbekistan, Venezuela (Bolivarian Republic of), Viet Nam, Yemen, Zambia, Zimbabwe.

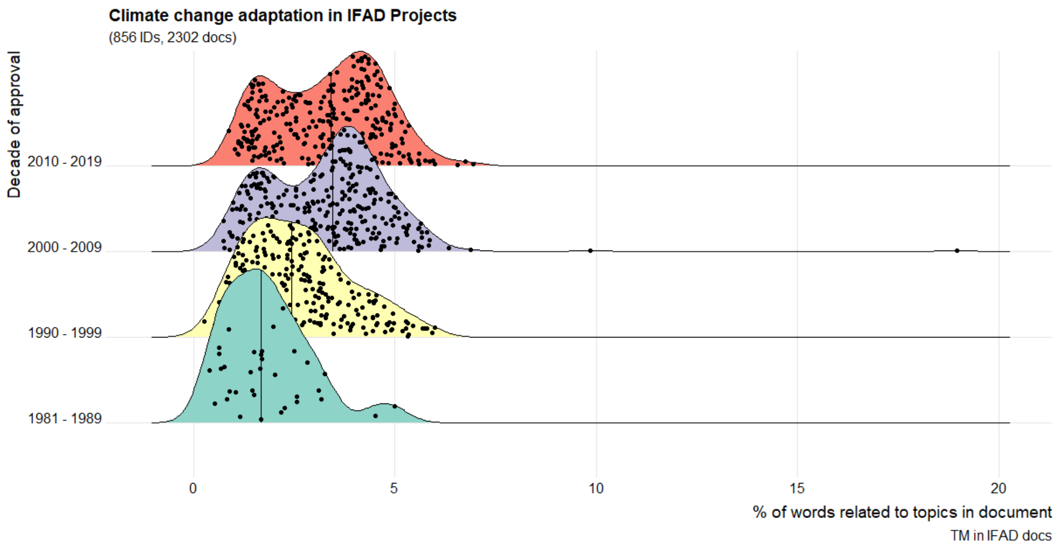
Appendix III – Systematic review: list of sources searched

CAB Abstracts, Scopus, EconLit search, AGRIS (FAO-consolidated search), International Food Policy Research Institute (IFPRI), World Bank, Collaboration for Environmental Evidence (CEE), Council for Agricultural Science and Technology (CAST), J-PAL/ATAI Impact evaluations (IPA), Overseas Development Institute (ODI), UK Department for International Development (DFID), International Fund for Agricultural Development (IFAD), International Institute for Environment and Development, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Agritrop (French Agricultural Research Centre for International Development [CIRAD]-consolidated search), AgEcon Search, EMBRAPA, WHO, UNEP, WFP, African Theses and Dissertations, Campbell Collaboration, CIRAD, Dissertations and Theses Global (access via ProQuest), Cochrane Collaboration, Gardian (searches 15 CGIAR websites), Cgspace, 3ie Impact Evaluation Database, BRAC, Center for Effective Global Action (CEGA), African-Asian Rural Development Organization (AARDO), African Development Bank (ADB), African Union Interafrican Bureau for Animal Resources (AU-IBAR), Agri Benchmark, AgResearch, Anthra, Association por la Promotion de l'Élevage au Sahel et en Savane, Leibniz-Institut für Agrartechnik und Bioökonomie e.V. (ATB), Animal Task Force (ATF), Bill & Melinda Gates Foudnation, Beam Beijing, Bern University of Applied Sciences School of Agricultural, Forest and Food Sciences HAFL, Agriculture and Agri-Food Canada, The Canadian Cattlemen's Association, Centro Agronómico Tropical de Investigación y Enseñanza (CATIE), Centro Brasileiro de Pecuaria Sustentavel (CBPS), Centro para la Investigación en Sistemas Sostenibles de Producción Agropecuaria (CIPAV), Compassion In World Farming (CIWF), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Confédération Nationale de l'Élevage, Country Carbon, Dairy Sustainability Framework, An Roinn Talmhaíochta, Bia agus Mara | Department of Agriculture, Food and the Marine (DAFM), República Dominicana: Dirección General de Ganadería, El Colegio de la Frontera Sur (ECOSUR), Ethiopian Society of Animal Production, The International Dairy Federation (IDF), Leibniz Institute for Farm Animal Biology, Fundación CoMunidad, Deutsche Gesellschaft für Internationale Zusammenarbeit (GIX) GmbH, Global Dairy Platform, Global Initiatives, The Global Roundtable for Sustainable Beef (GRSB), Heifer International, Human Society International, Kenya Dairy Board, Kyeema Foundation, Inter Eco Center, International Center for Tropical Agriculture (CIAT), International Poultry Council, International Farm Comparison Network (IFCN), International Feed Industry Federation (IFIF), Instituto Interamericano de Cooperación para la Agricultura (IICA), French National Institute for Agricultural Research (INRA), Instituto Nacional de Tecnología Agropecuaria (INTA), Institut de l'Élevage, Plan Agropecuario, International Meat Secretariat, Senaapathy Kangayam Cattle Research Foundation, College of Veterinary Medicine | Kansas State University, Kenya Livestock Producer's Association, League of Pastoralism, LIFE Network, Livestock Farming and Local Development (LiFLoD), Ministerio de Agricultura y Ganadería (MAG) | Costa Rica, Ministerio de Agricultura y Ganadería (MAG) | Ecuador, Ministerio de Agricultura y Ganadería (MAG) | Paraguay, Ministry of Livestock and Fisheries | Ethiopia, Ministry of Agriculture, Livestock and Fisheries | Kenya, Ministry of Agriculture and Animal Resources | Rwanda, Ministry of Food, Agriculture and Light Industry | Mongolia, Ministry of Economic Affairs of the Netherlands, Government of Panama | Ministry of Agricultural and Livestock Development, Ministerio de Medio Ambiente de El Salvador, Government of New Zealand | Ministry for Primary Industries, The Nature Conservancy, Novus International, World Organization for Animal Health, Fundación Produce Michoacán, PACTO COQUETA, Redes Chaco, France Ministère de l'Agriculture et de l'Alimentation, Republic of Keanya | County Government of Busia | Department of Agriculture and Animal Resources | County Livestock Production, Royal Veterinary College | University of London, Ranch 4 International Ltd | Canada, SAVES, the Swedish University of Agricultural Sciences (SLU), Savory Institute, Swiss Confederation – Federal Office for Agriculture (FOAG), Swissgenetics, The Donkey Sanctuary, Trust in Animals and Food

Safety (TAFS), Turkey Farmers of Canada (TFC), Uganda Bureau of Statistics (UBOS), European Livestock and Meat Trading Union (UECBV), University of Florida's (UF) Institute of Food and Agriculture Sciences (IFAS), Universidade Federal de São João del-Rei, Universidad Austral de la Patagonia | Argentina, Uganda National Farmers Federation, University of Melbourne, VanDrie Group, National Institute of Animal Sciences – Thuy Phuong, Tu Liem, Ha Noi, Viet Nam, VetEffecT Veterinary and Public Health, VSF International – Vétérinaires Sans Frontières, World Alliance of Mobile Indigenous Peoples (WAMIP), World Animal Protection, World Horse Welfare, World Wildlife Fund (WWF), The Yield Lab Institute, International Goat Association (IGA), World Agroforesty Center, International Centre for Integrated Mountain Development (ICIMOD).

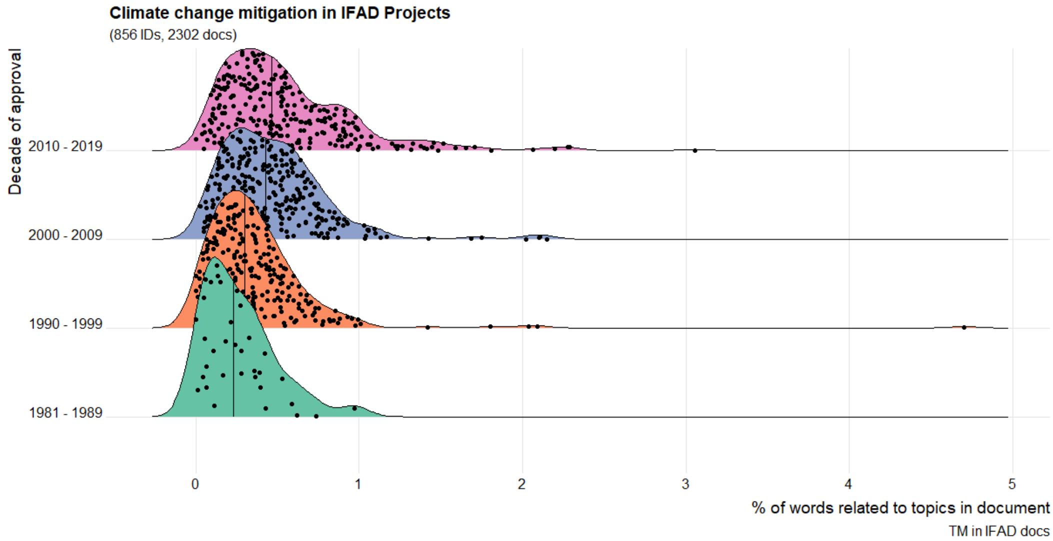
Appendix IV– Presence of mainstreaming themes and cross-cutting issues in IFAD documents over time

Figure A1 – Distribution of Climate Change Adaptation in project documentation per decade



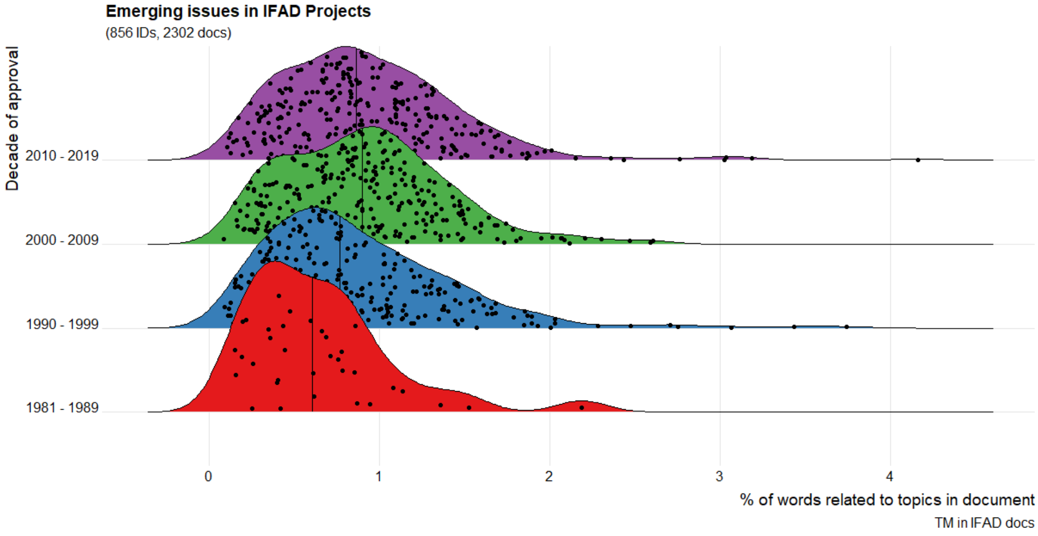
Share of words associated to IFAD's mainstreaming themes present in project documentation. Key terms collected manually and expanded by word2vec (2302 reports from 856 projects analysed).

Figure A2 – Distribution of Climate Change Mitigation in project documentation per decade



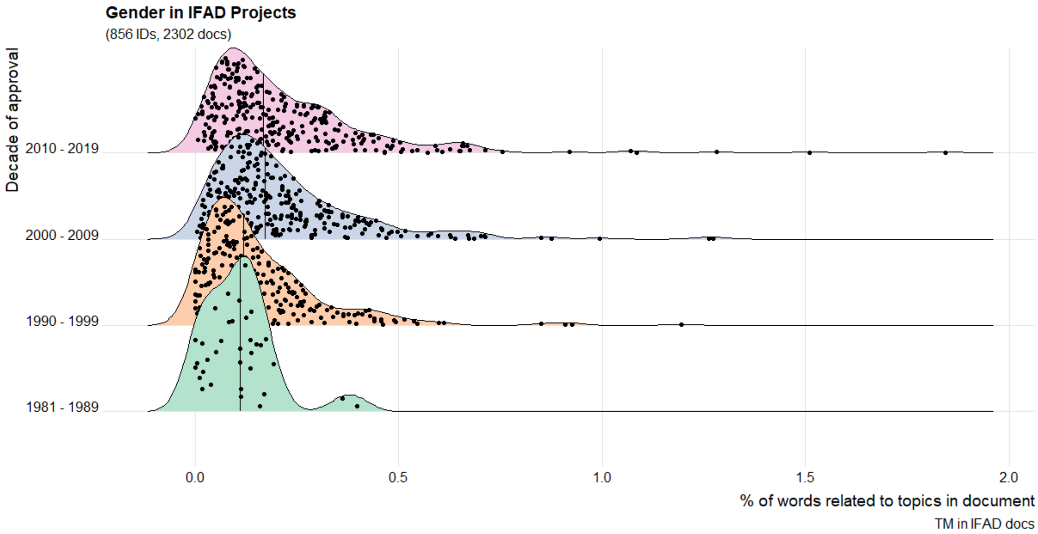
Share of words associated to IFAD's mainstreaming themes present in project documentation. Key terms collected manually and expanded by word2vec (2302 reports from 856 projects analysed).

Figure A3 – Distribution of Emerging Issues in project documentation per decade



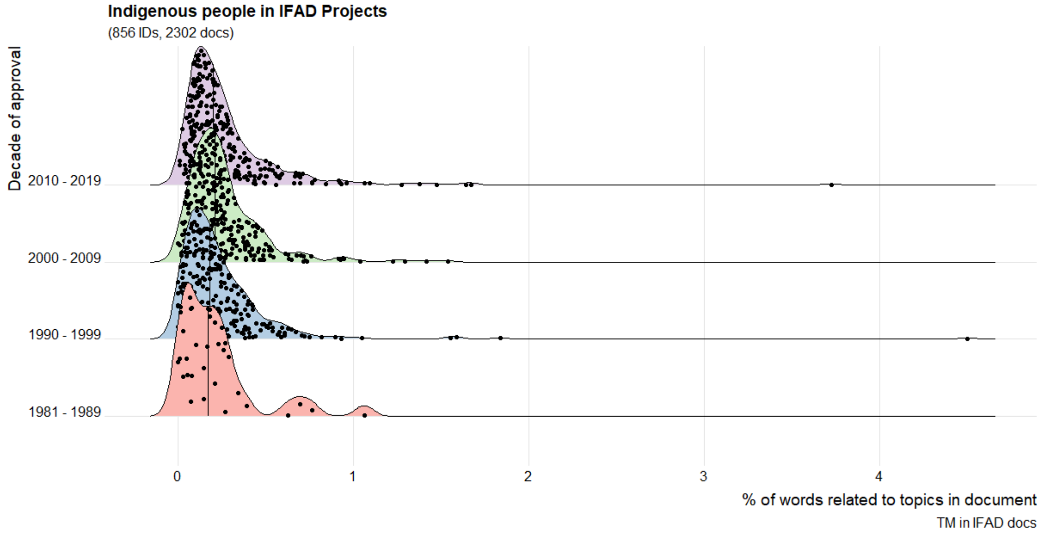
Share of words associated to IFAD's mainstreaming themes present in project documentation. Key terms collected manually and expanded by word2vec (2302 reports from 856 projects analysed).

Figure A4 – Distribution of Gender in project documentation per decade



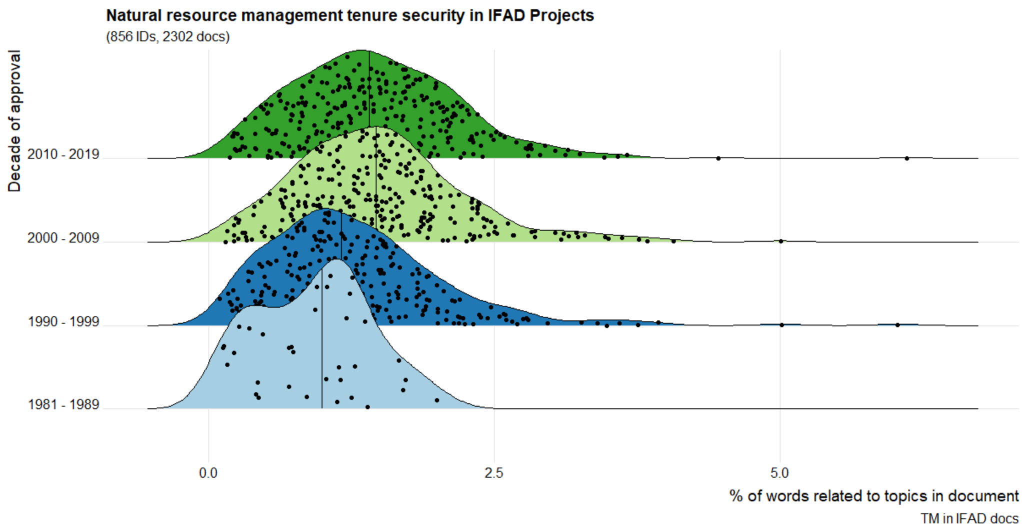
Share of words associated to IFAD's mainstreaming themes present in project documentation. Key terms collected manually and expanded by word2vec (2302 reports from 856 projects analysed).

Figure A5 – Distribution of Indigenous Peoples in project documentation per decade



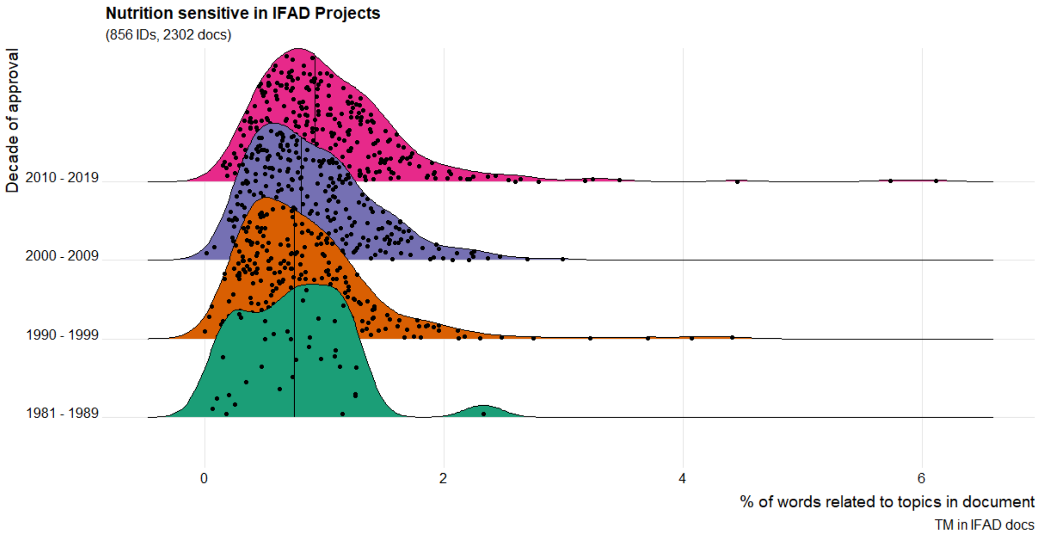
Share of words associated to IFAD's mainstreaming themes present in project documentation. Key terms collected manually and expanded by word2vec (2302 reports from 856 projects analysed).

Figure A6 – Distribution of Natural Resource Management and Land Tenure in project documentation per decade



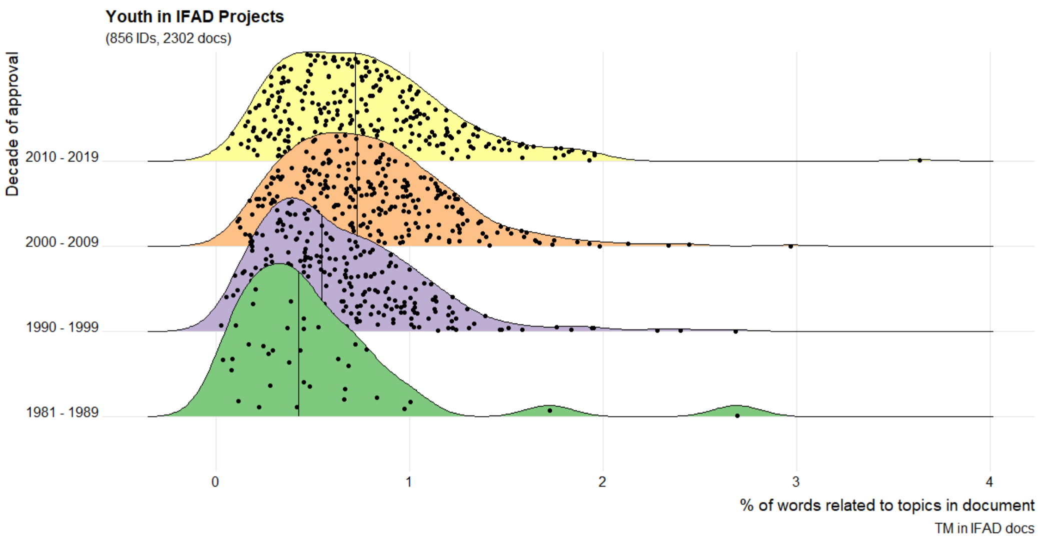
Share of words associated to IFAD's mainstreaming themes present in project documentation. Key terms collected manually and expanded by word2vec (2302 reports from 856 projects analysed).

Figure A7 – Distribution of Nutrition in project documentation per decade



Share of words associated to IFAD's mainstreaming themes present in project documentation. Key terms collected manually and expanded by word2vec (2302 reports from 856 projects analysed).

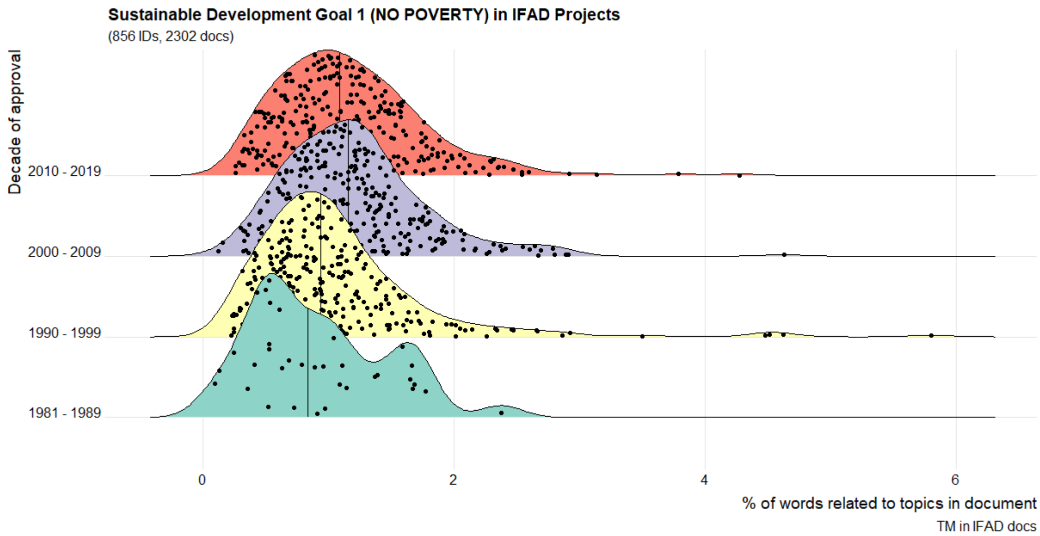
Figure A8 – Distribution of Youth in project documentation per decade



Share of words associated to IFAD's mainstreaming themes present in project documentation. Key terms collected manually and expanded by word2vec (2302 reports from 856 projects analysed).

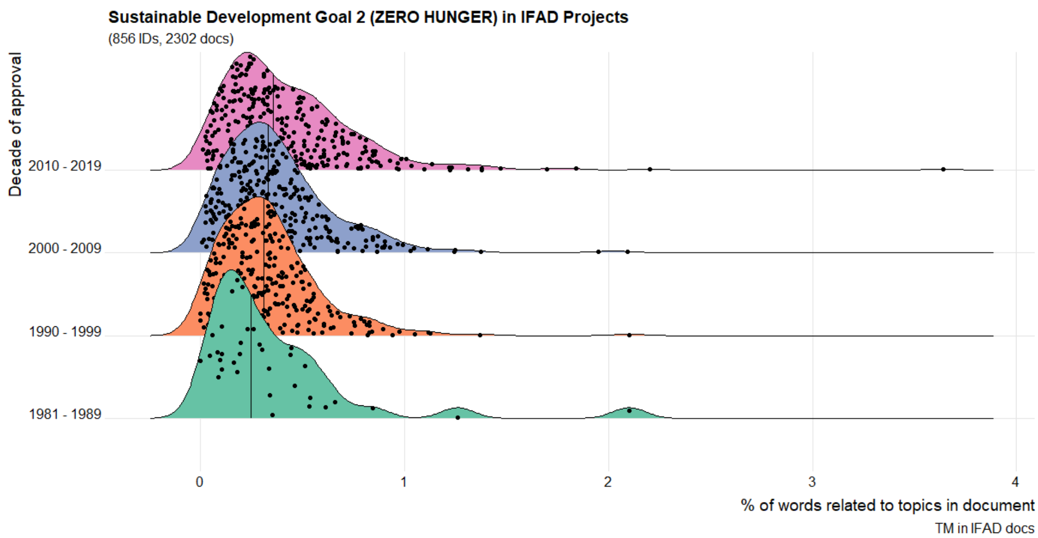
Appendix V – Presence of SDGs in IFAD documents over time

Figure A9 – Distribution of SDG 1 in project documentation per decade



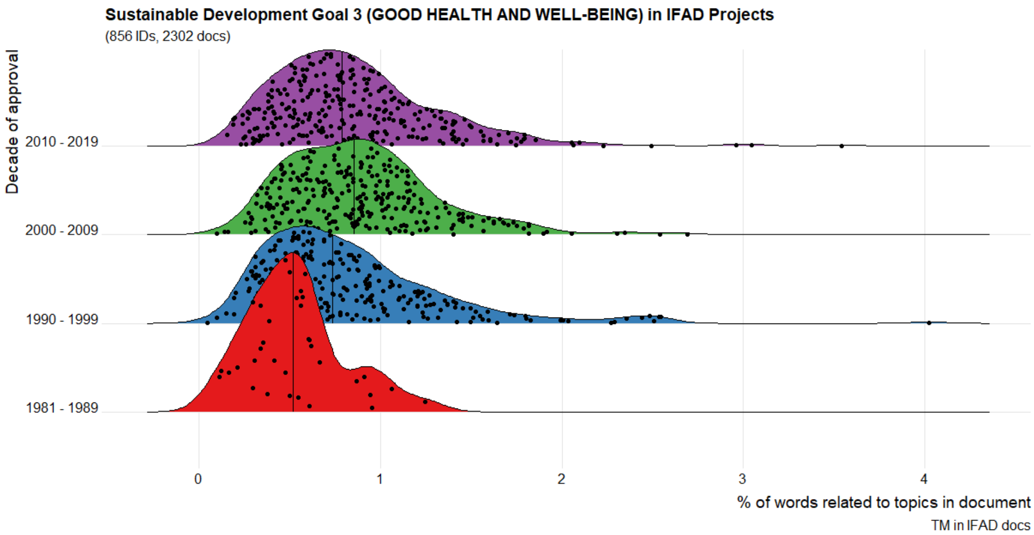
Share of words associated to SDGs present in project documentation. Key terms taken from public definition and expanded by word2vec (2302 reports from 856 projects analysed).

Figure A10 – Distribution of SDG 2 in project documentation per decade



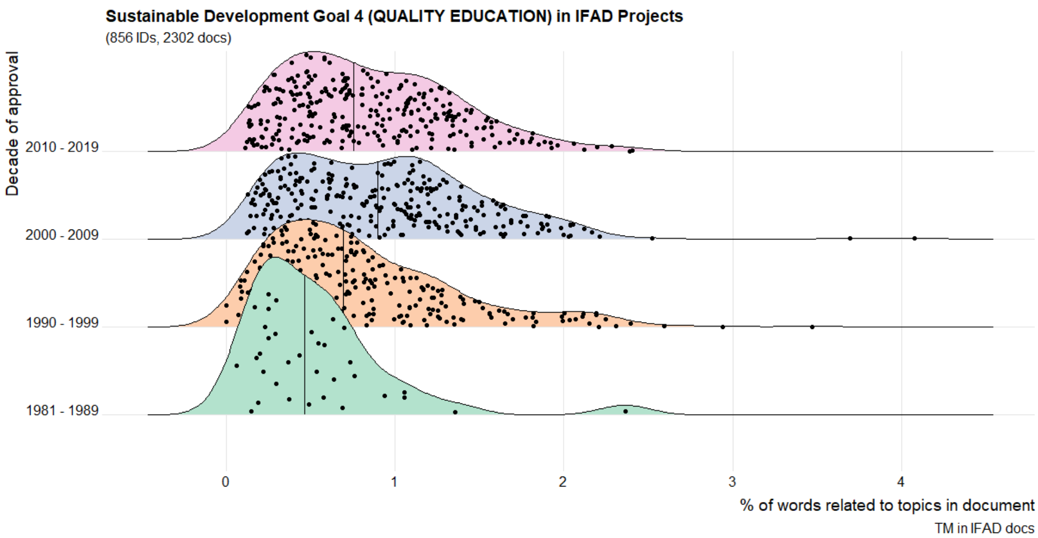
Share of words associated to SDGs present in project documentation. Key terms taken from public definition and expanded by word2vec (2302 reports from 856 projects analysed).

Figure A11 – Distribution of SDG 3 in project documentation per decade



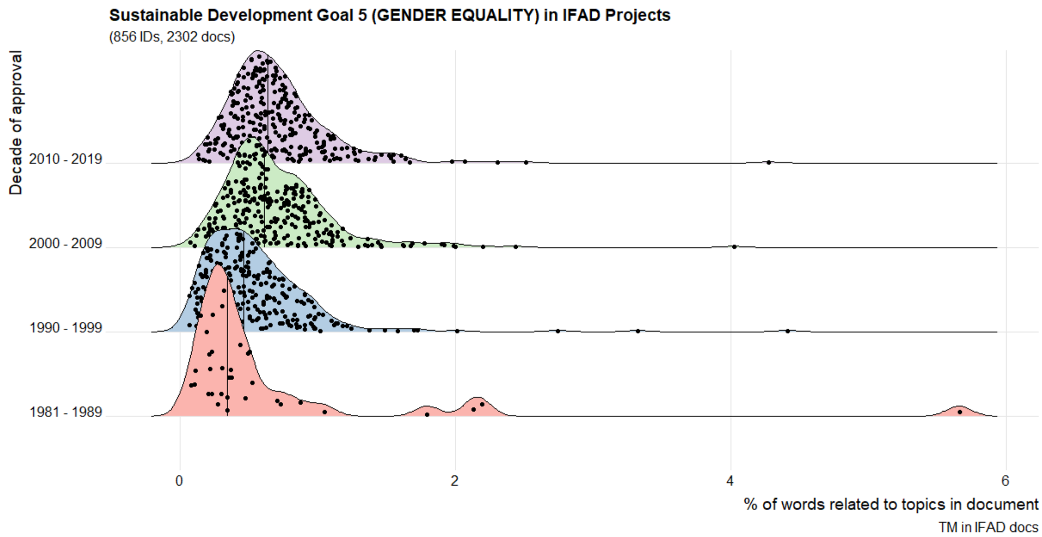
Share of words associated to SDGs present in project documentation. Key terms taken from public definition and expanded by word2vec (2302 reports from 856 projects analysed).

Figure A12 – Distribution of SDG 4 in project documentation per decade



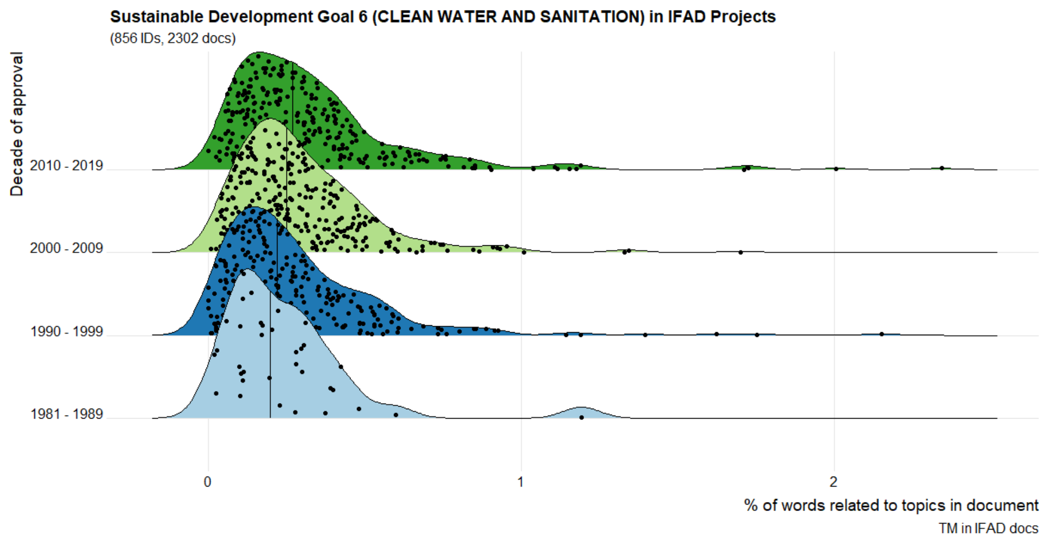
Share of words associated to SDGs present in project documentation. Key terms taken from public definition and expanded by word2vec (2302 reports from 856 projects analysed).

Figure A13 – Distribution of SDG 5 in project documentation per decade



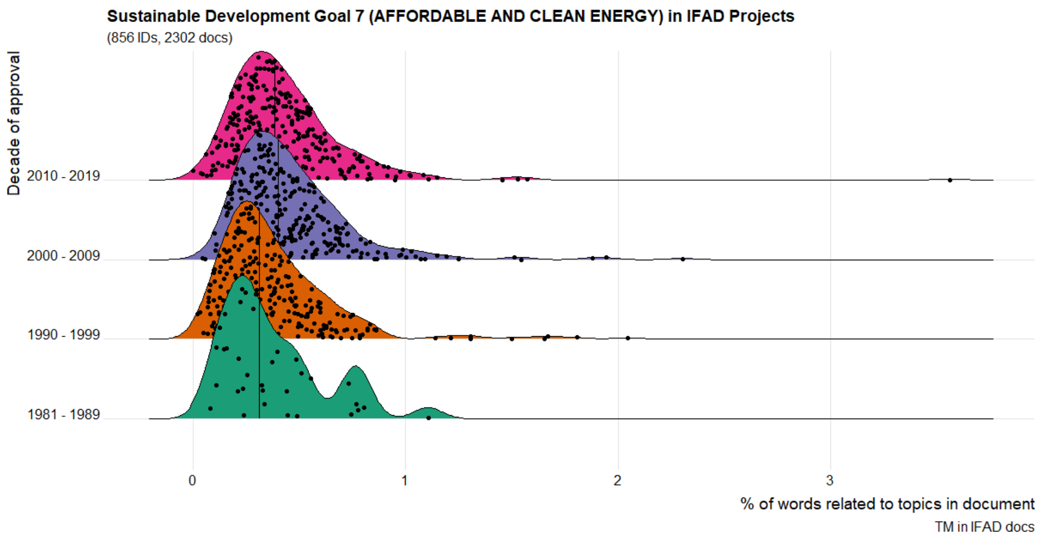
Share of words associated to SDGs present in project documentation. Key terms taken from public definition and expanded by word2vec (2302 reports from 856 projects analysed).

Figure A14 – Distribution of SDG 6 in project documentation per decade



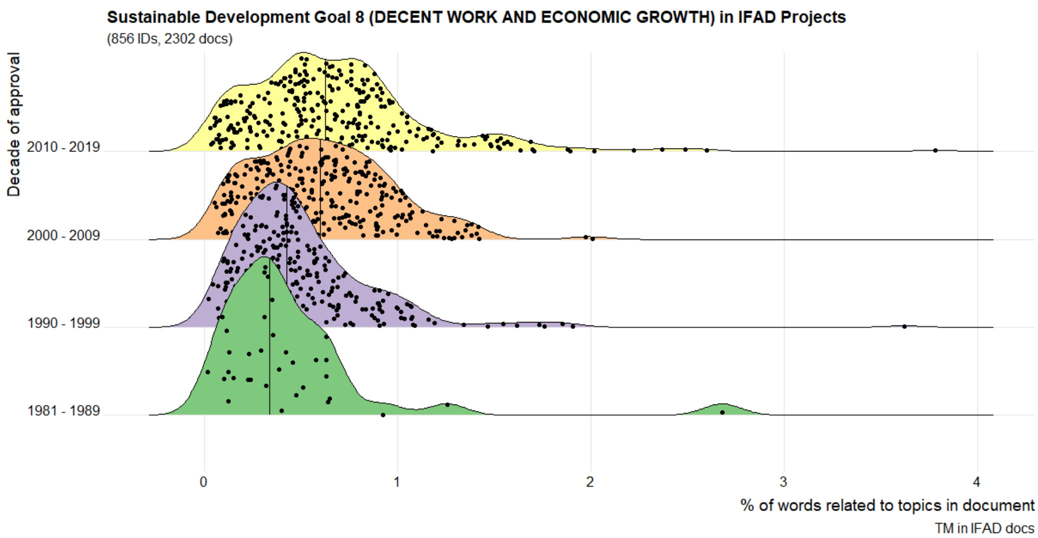
Share of words associated to SDGs present in project documentation. Key terms taken from public definition and expanded by word2vec (2302 reports from 856 projects analysed).

Figure A15 – Distribution of SDG 7 in project documentation per decade



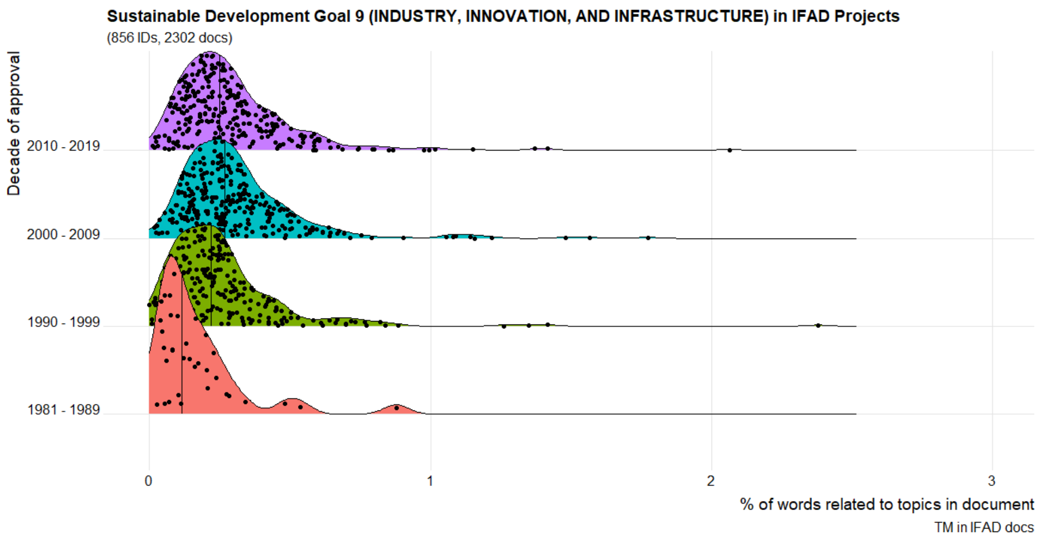
Share of words associated to SDGs present in project documentation. Key terms taken from public definition and expanded by word2vec (2302 reports from 856 projects analysed).

Figure A16 – Distribution of SDG 8 in project documentation per decade



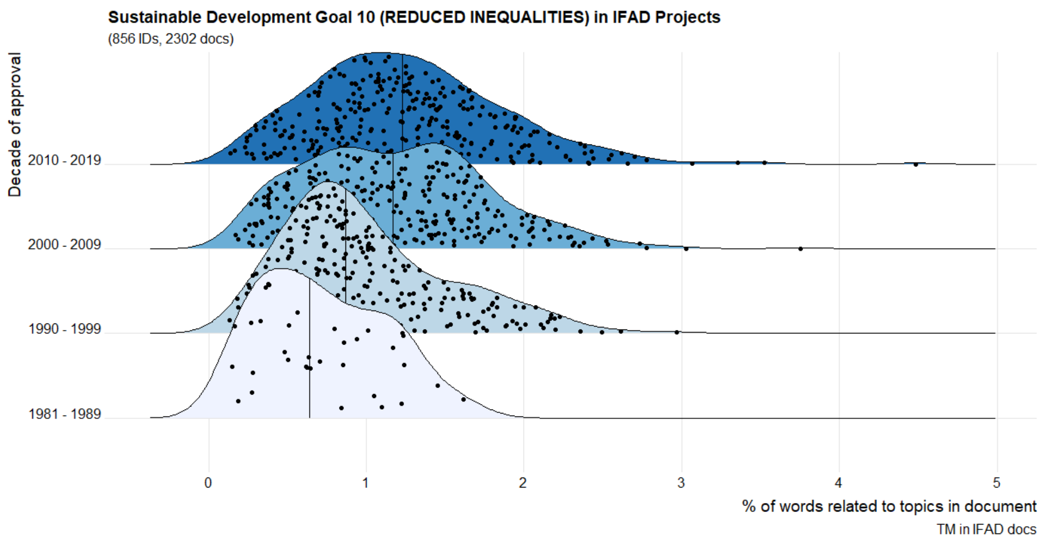
Share of words associated to SDGs present in project documentation. Key terms taken from public definition and expanded by word2vec (2302 reports from 856 projects analysed).

Figure A17 – Distribution of SDG 9 in project documentation per decade



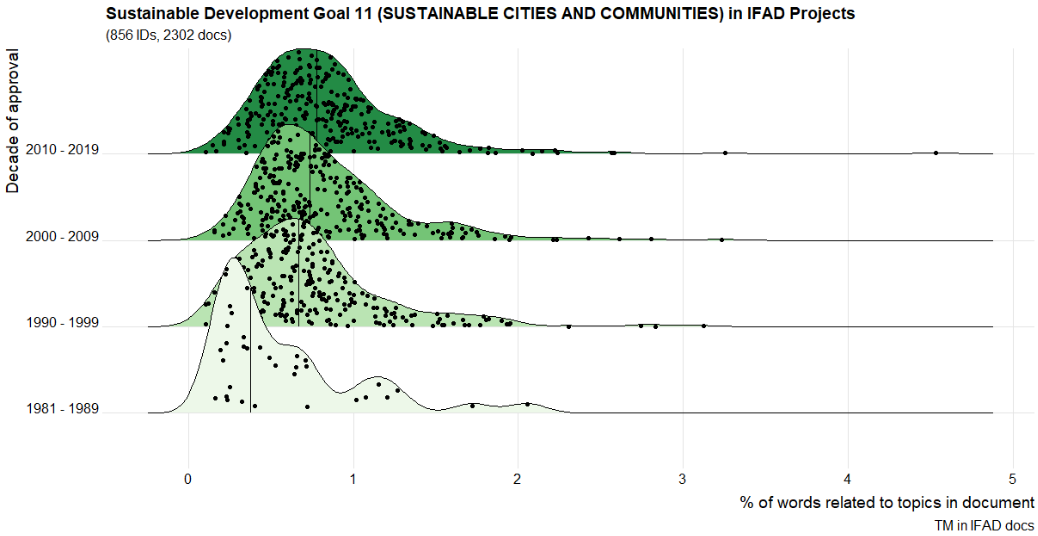
Share of words associated to SDGs present in project documentation. Key terms taken from public definition and expanded by word2vec (2302 reports from 856 projects analysed).

Figure A18 – Distribution of SDG 10 in project documentation per decade



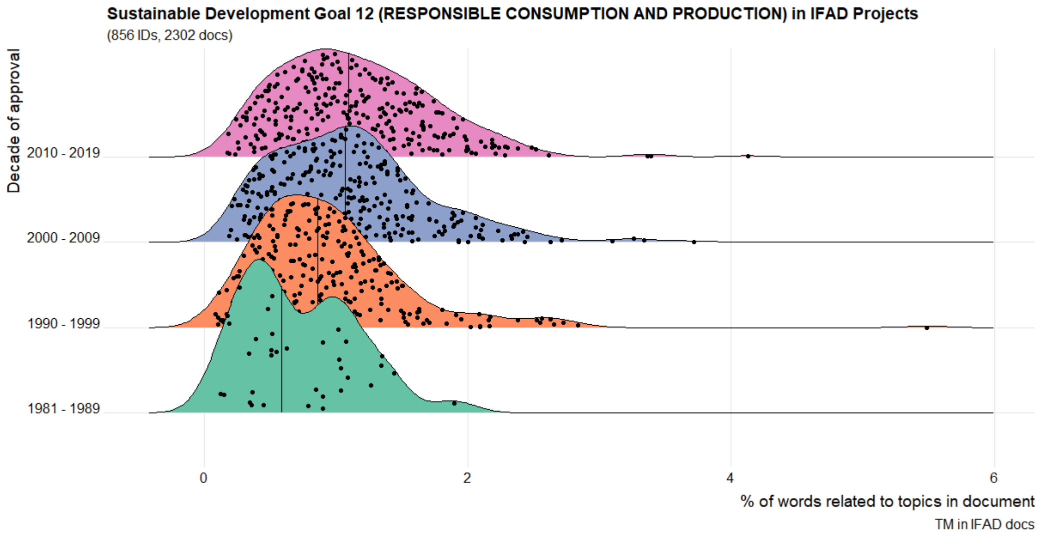
Share of words associated to SDGs present in project documentation. Key terms taken from public definition and expanded by word2vec (2302 reports from 856 projects analysed).

Figure A19 – Distribution of SDG 11 in project documentation per decade



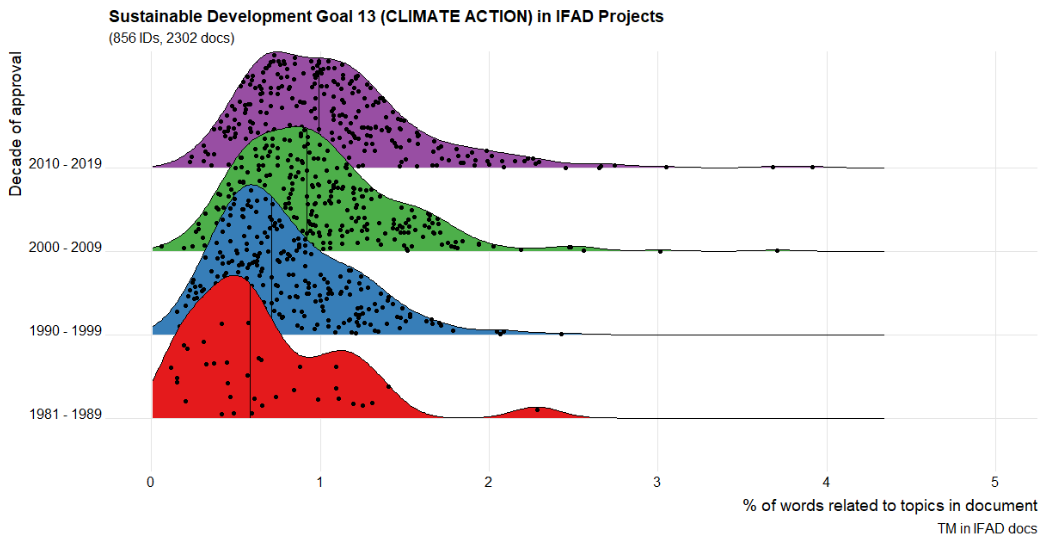
Share of words associated to SDGs present in project documentation. Key terms taken from public definition and expanded by word2vec (2302 reports from 856 projects analysed).

Figure A20 – Distribution of SDG 12 in project documentation per decade



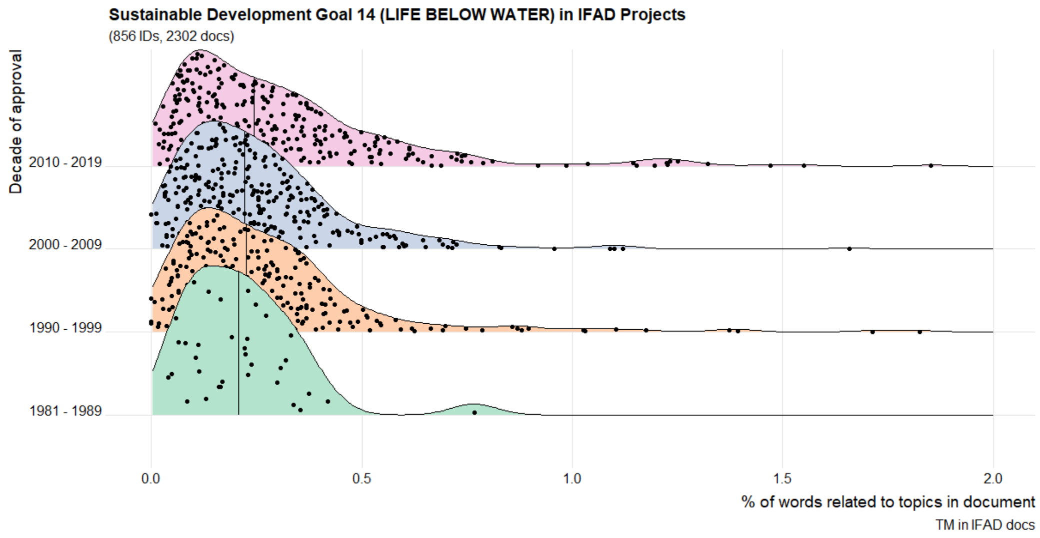
Share of words associated to SDGs present in project documentation. Key terms taken from public definition and expanded by word2vec (2302 reports from 856 projects analysed).

Figure A21 – Distribution of SDG 13 in project documentation per decade



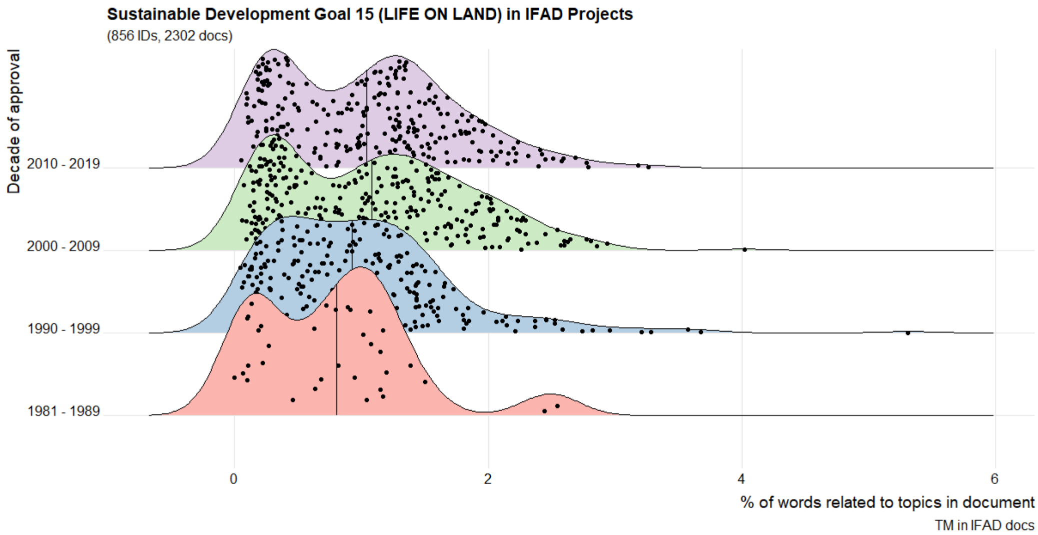
Share of words associated to SDGs present in project documentation. Key terms taken from public definition and expanded by word2vec (2302 reports from 856 projects analysed).

Figure A22 – Distribution of SDG 14 in project documentation per decade



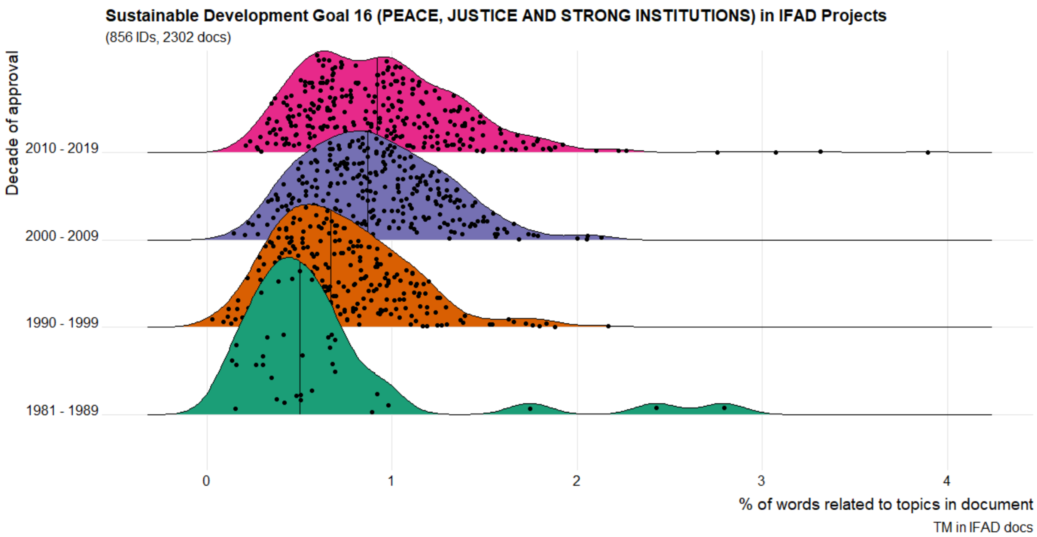
Share of words associated to SDGs present in project documentation. Key terms taken from public definition and expanded by word2vec (2302 reports from 856 projects analysed).

Figure A23 – Distribution of SDG 15 in project documentation per decade



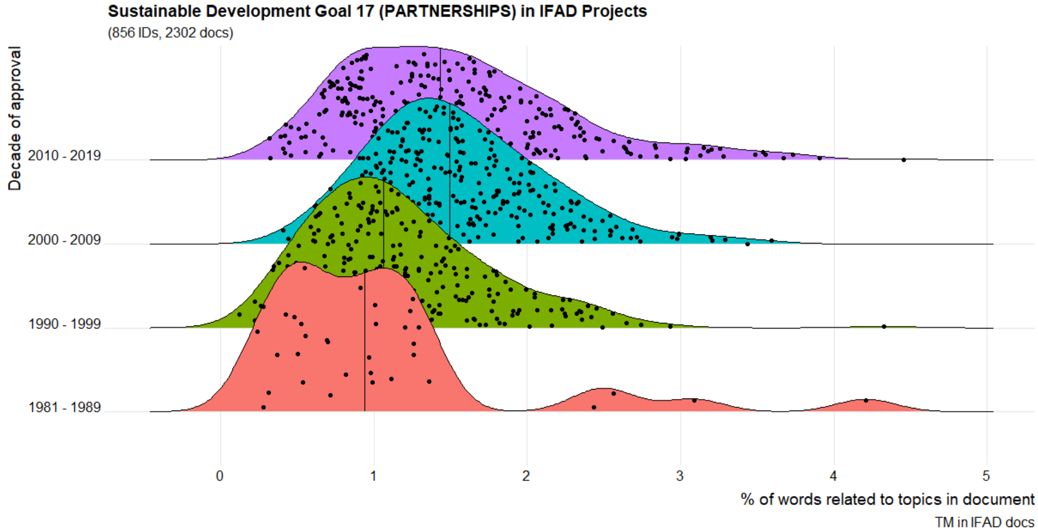
Share of words associated to SDGs present in project documentation. Key terms taken from public definition and expanded by word2vec (2302 reports from 856 projects analysed).

Figure A24 – Distribution of SDG 16 in project documentation per decade



Share of words associated to SDGs present in project documentation. Key terms taken from public definition and expanded by word2vec (2302 reports from 856 projects analysed).

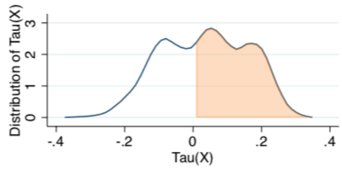
Figure A25 – Distribution of SDG 17 in project documentation per decade



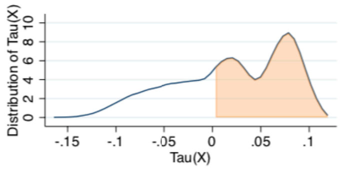
Share of words associated to SDGs present in project documentation. Key terms taken from public definition and expanded by word2vec (2302 reports from 856 projects analysed).

Appendix VI - Percentage of households obtaining a positive IFAD-10 treatment effect by country/project.

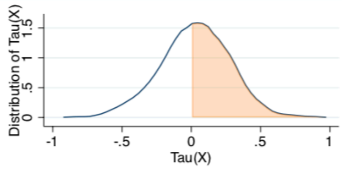
Bangladesh



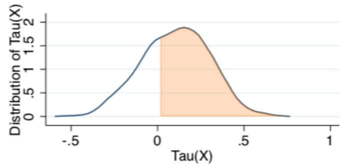
Country: Bangladesh
Outcome: Log of Total Gross Income from all Sources
N = 2743, N1 = 1191, N0 = 1552
Tau(X) > 0: 60%
Tau1(X) > 0: 58%



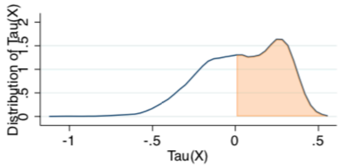
Country: Bangladesh
Outcome: Log of Total Net Income from all Sources
N = 2649, N1 = 1157, N0 = 1492
Tau(X) > 0: 65%
Tau1(X) > 0: 64%



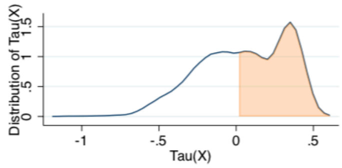
Country: Bangladesh
Outcome: Log of Gross Income from Crops and Byproducts
N = 1410, N1 = 659, N0 = 751
Tau(X) > 0: 56%
Tau1(X) > 0: 54%



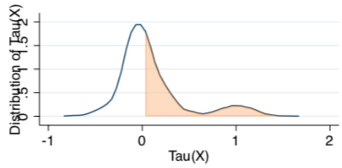
Country: Bangladesh
Outcome: Log of Net Income from Crops and Byproducts
N = 1200, N1 = 575, N0 = 625
Tau(X) > 0: 69%
Tau1(X) > 0: 69%



Country: Bangladesh
Outcome: Log of Gross Income from Livestock and Byproduct
N = 1112, N1 = 495, N0 = 617
Tau(X) > 0: 57%
Tau1(X) > 0: 57%

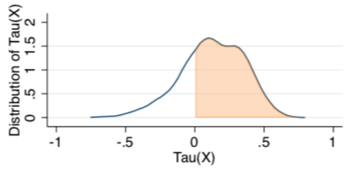


Country: Bangladesh
Outcome: Log of Net Income from Livestock and Byproducts
N = 934, N1 = 412, N0 = 522
Tau(X) > 0: 57%
Tau1(X) > 0: 56%

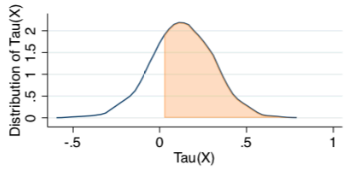


Country: Bangladesh
Outcome: Log of Bangladesh & Indonesia: Net Income from Fisheries
N = 204, N1 = 94, N0 = 110
Tau(X) > 0: 52%
Tau1(X) > 0: 50%

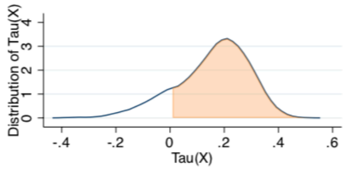
Bolivia



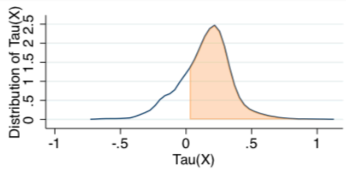
Country: Bolivia
Outcome: Log of Total Gross Income from all Sources
N = 2523, N1 = 855, N0 = 1668
Tau(X) > 0: 71%
Tau1(X) > 0: 71%



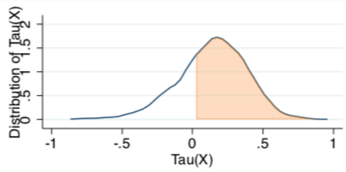
Country: Bolivia
Outcome: Log of Total Net Income from all Sources
N = 2415, N1 = 822, N0 = 1593
Tau(X) > 0: 76%
Tau1(X) > 0: 76%



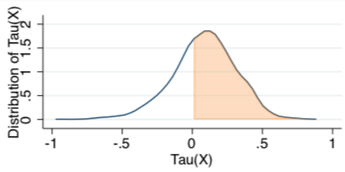
Country: Bolivia
Outcome: Log of Gross Income from Crops and Byproducts
N = 2364, N1 = 796, N0 = 1568
Tau(X) > 0: 85%
Tau1(X) > 0: 84%



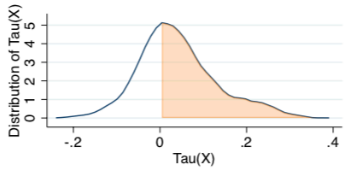
Country: Bolivia
Outcome: Log of Net Income from Crops and Byproducts
N = 2121, N1 = 715, N0 = 1406
Tau(X) > 0: 78%
Tau1(X) > 0: 78%



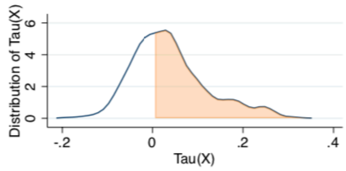
Country: Bolivia
Outcome: Log of Gross Income from Livestock and Byproduct
N = 1652, N1 = 595, N0 = 1057
Tau(X) > 0: 74%
Tau1(X) > 0: 71%



Country: Bolivia
Outcome: Log of Net Income from Livestock and Byproducts
N = 1586, N1 = 575, N0 = 1011
Tau(X) > 0: 67%
Tau1(X) > 0: 63%

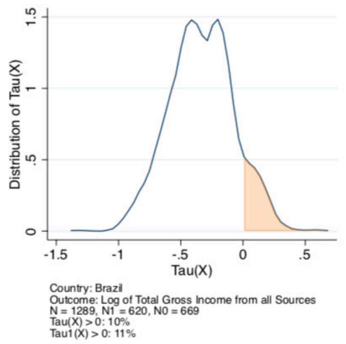


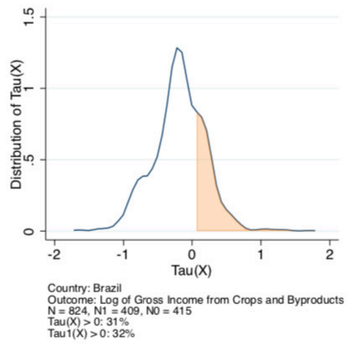
Country: Bolivia
Outcome: Log of Ability to recover from shocks
N = 2332, N1 = 790, N0 = 1542
Tau(X) > 0: 66%
Tau1(X) > 0: 67%

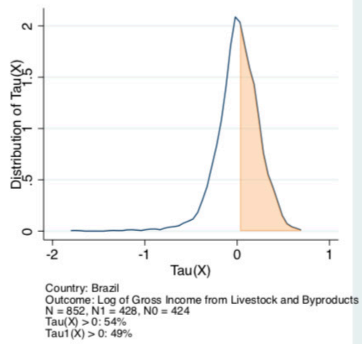


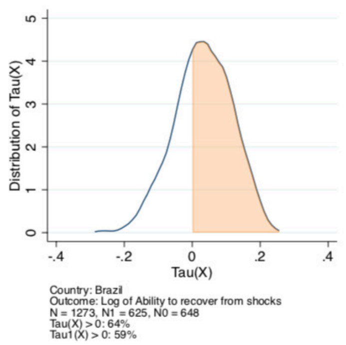
Country: Bolivia
Outcome: Log of Ability to recover from shocks - 2
N = 2326, N1 = 787, N0 = 1539
Tau(X) > 0: 63%
Tau1(X) > 0: 64%

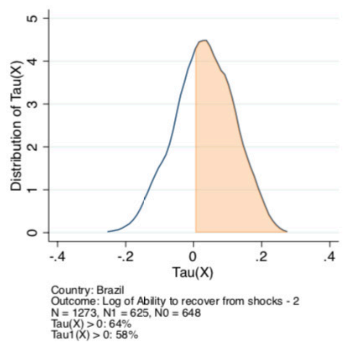
Brazil


Distribution of Tau(X)
Tau(X)
Country: Brazil
Outcome: Log of Total Gross Income from all Sources
N = 1289, N1 = 620, N0 = 669
Tau(X) > 0: 10%
Tau1(X) > 0: 11%

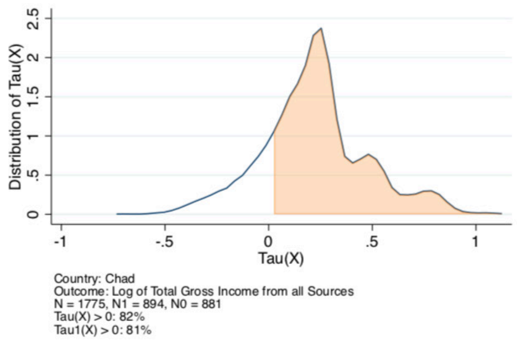

Distribution of Tau(X)
Tau(X)
Country: Brazil
Outcome: Log of Gross Income from Crops and Byproducts
N = 824, N1 = 409, N0 = 415
Tau(X) > 0: 31%
Tau1(X) > 0: 32%

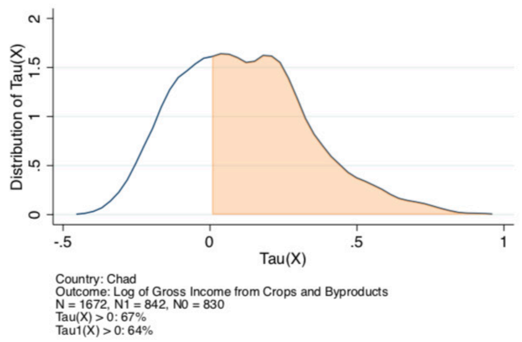

Distribution of Tau(X)
Tau(X)
Country: Brazil
Outcome: Log of Gross Income from Livestock and Byproducts
N = 852, N1 = 428, N0 = 424
Tau(X) > 0: 54%
Tau1(X) > 0: 49%

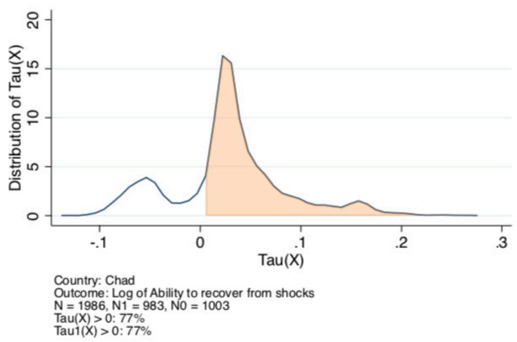

Distribution of Tau(X)
Tau(X)
Country: Brazil
Outcome: Log of Ability to recover from shocks
N = 1273, N1 = 625, N0 = 648
Tau(X) > 0: 64%
Tau1(X) > 0: 59%

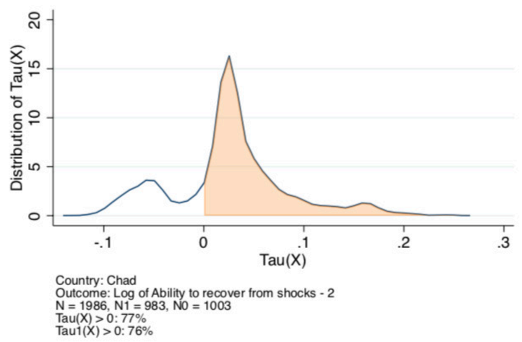

Distribution of Tau(X)
Tau(X)
Country: Brazil
Outcome: Log of Ability to recover from shocks - 2
N = 1273, N1 = 625, N0 = 648
Tau(X) > 0: 64%
Tau1(X) > 0: 58%

Chad

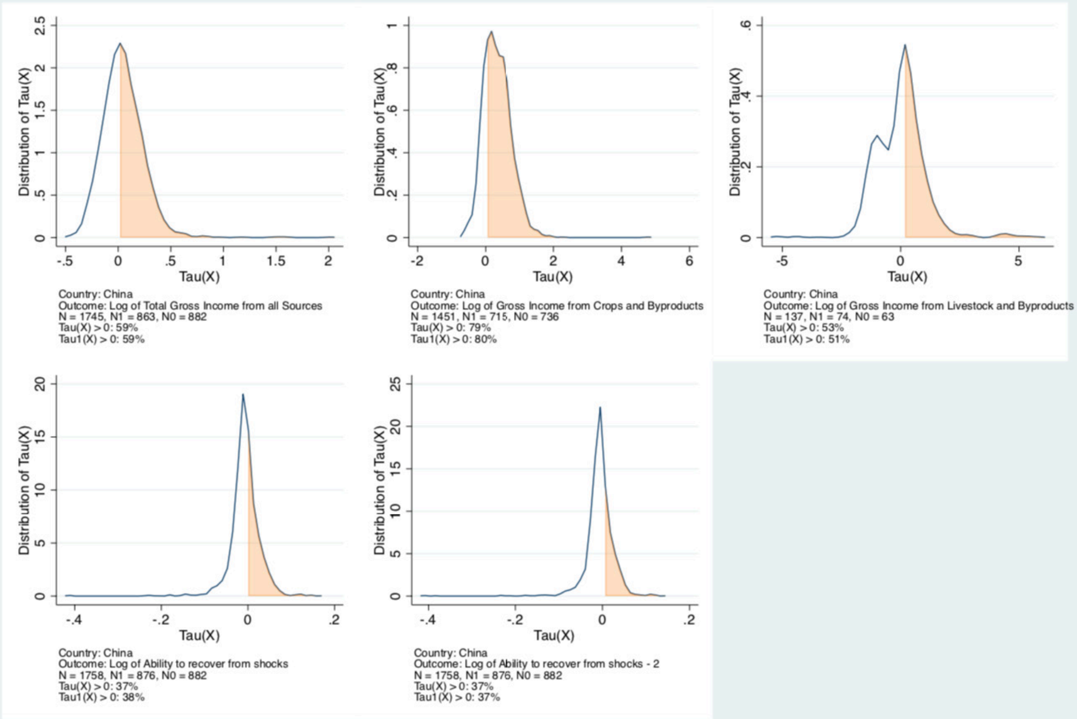

Distribution of Tau(X)
Tau(X)
Country: Chad
Outcome: Log of Total Gross Income from all Sources
N = 1775, N1 = 894, N0 = 881
Tau(X) > 0: 82%
Tau1(X) > 0: 81%


Distribution of Tau(X)
Tau(X)
Country: Chad
Outcome: Log of Gross Income from Crops and Byproducts
N = 1672, N1 = 842, N0 = 830
Tau(X) > 0: 67%
Tau1(X) > 0: 64%

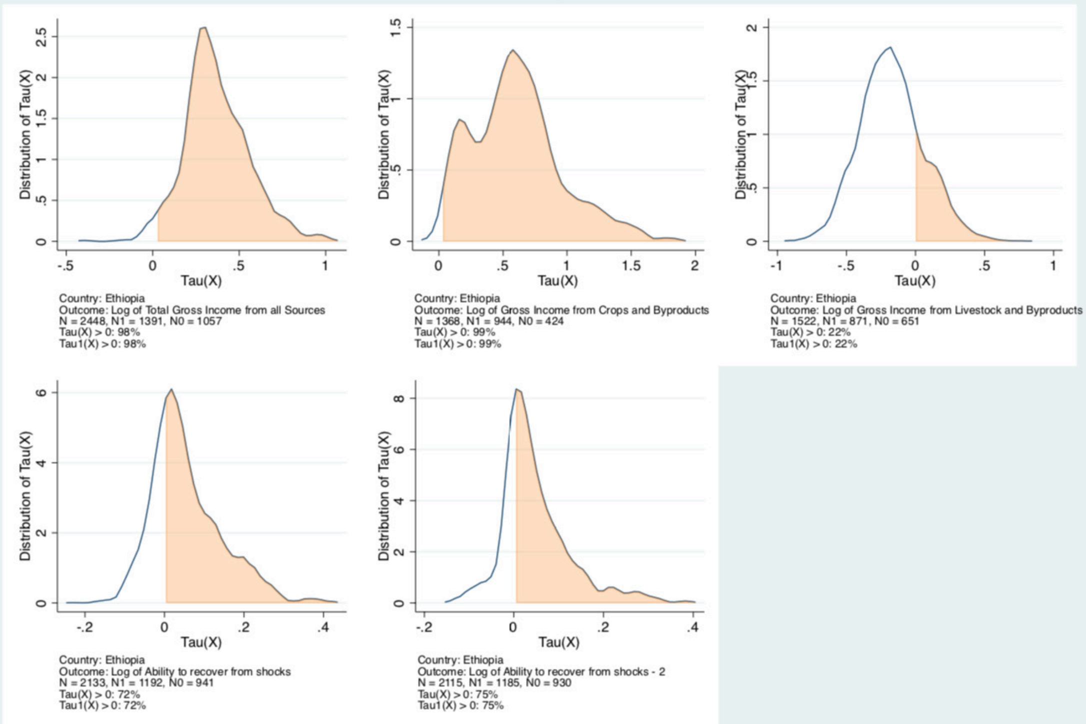

Distribution of Tau(X)
Tau(X)
Country: Chad
Outcome: Log of Ability to recover from shocks
N = 1986, N1 = 983, N0 = 1003
Tau(X) > 0: 77%
Tau1(X) > 0: 77%


Distribution of Tau(X)
Tau(X)
Country: Chad
Outcome: Log of Ability to recover from shocks - 2
N = 1986, N1 = 983, N0 = 1003
Tau(X) > 0: 77%
Tau1(X) > 0: 76%

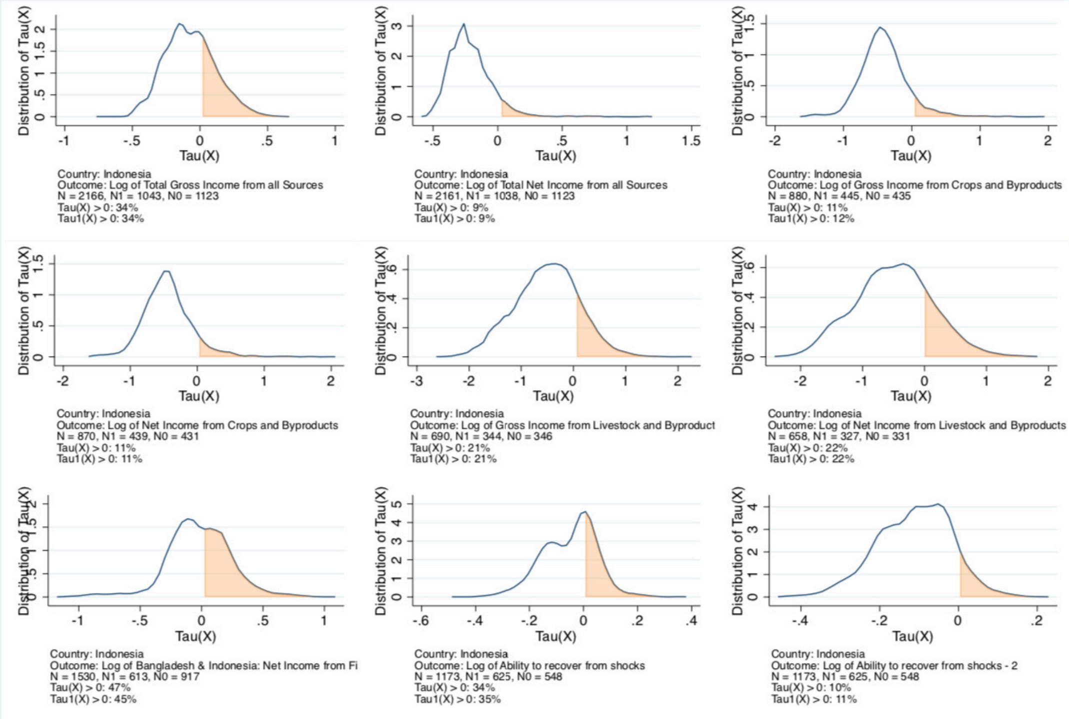
China



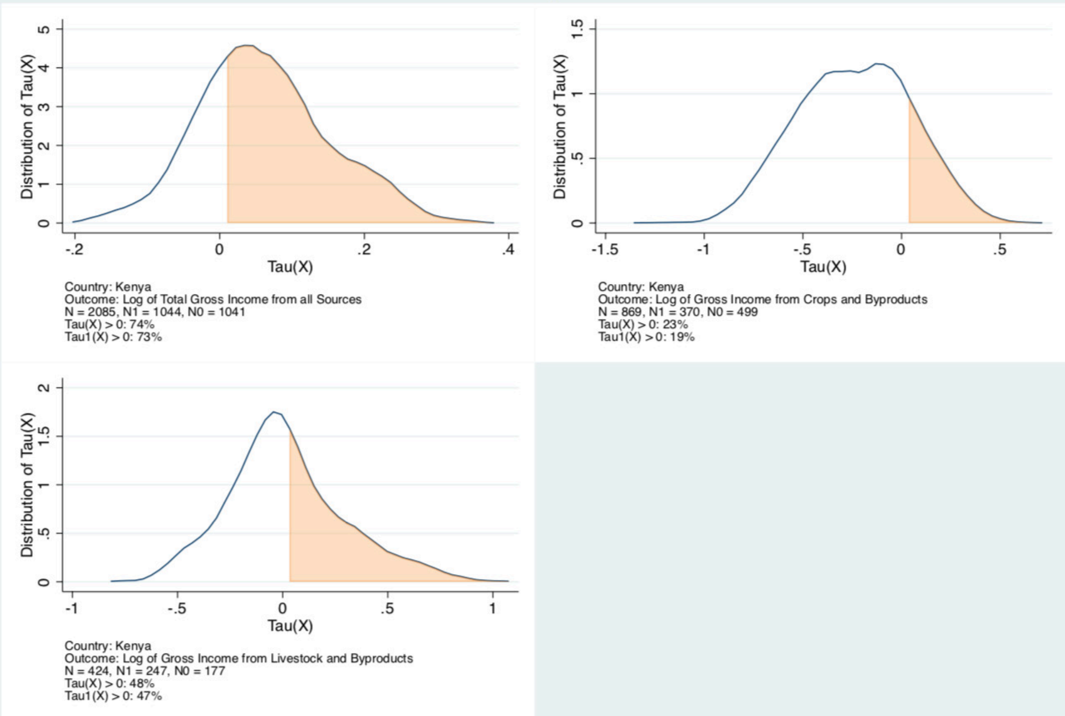
Ethiopia

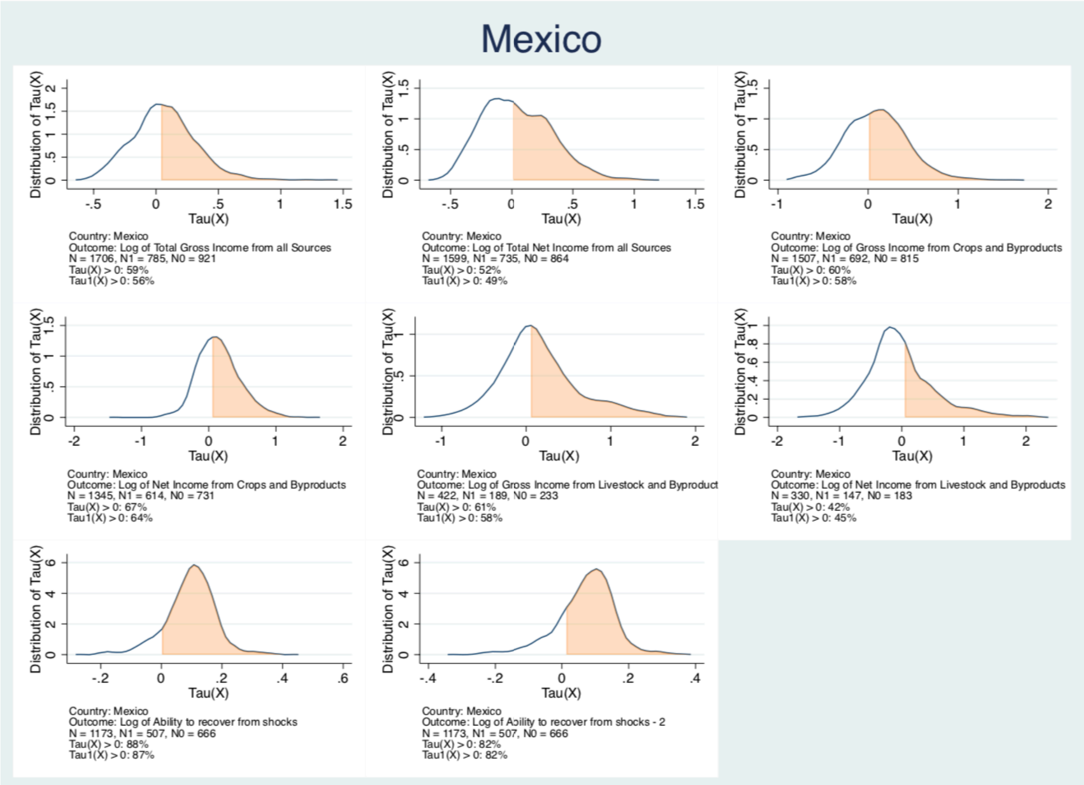


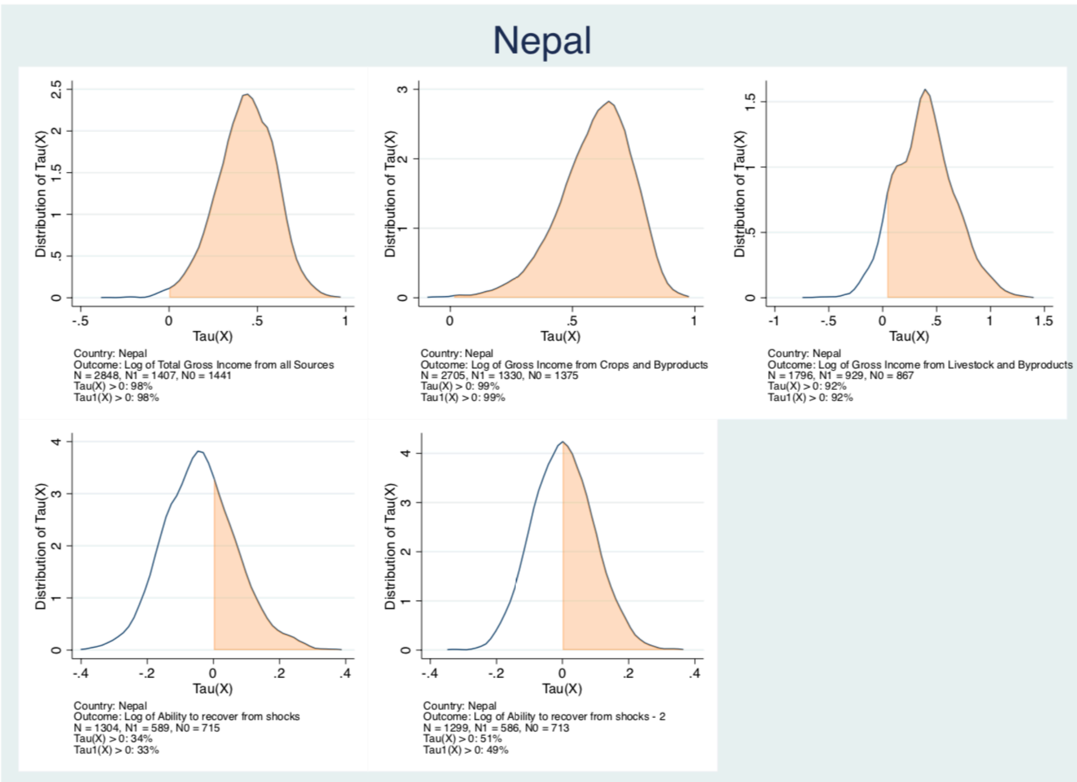
Indonesia



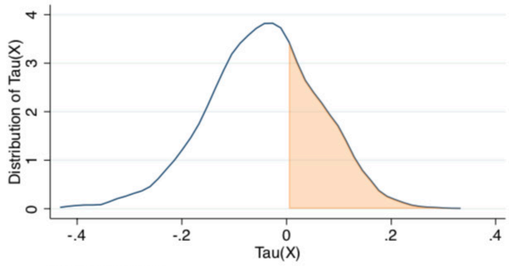
Kenya



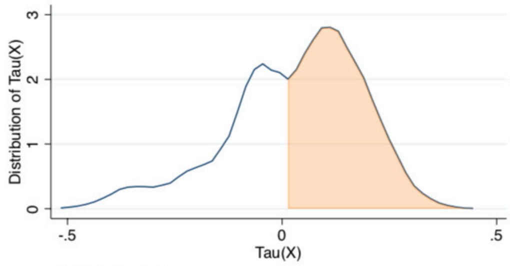

Mexico
Distribution of Tau(X)
Tau(X)
Country: Mexico
Outcome: Log of Total Gross Income from all Sources
N = 1706, N1 = 785, N0 = 921
Tau(X) > 0: 59%
Tau1(X) > 0: 56%
Country: Mexico
Outcome: Log of Total Net Income from all Sources
N = 1599, N1 = 735, N0 = 864
Tau(X) > 0: 52%
Tau1(X) > 0: 49%
Country: Mexico
Outcome: Log of Gross Income from Crops and Byproducts
N = 1507, N1 = 692, N0 = 815
Tau(X) > 0: 60%
Tau1(X) > 0: 58%
Country: Mexico
Outcome: Log of Net Income from Crops and Byproducts
N = 1345, N1 = 614, N0 = 731
Tau(X) > 0: 67%
Tau1(X) > 0: 64%
Country: Mexico
Outcome: Log of Gross Income from Livestock and Byproduct
N = 422, N1 = 189, N0 = 233
Tau(X) > 0: 61%
Tau1(X) > 0: 58%
Country: Mexico
Outcome: Log of Net Income from Livestock and Byproducts
N = 330, N1 = 147, N0 = 183
Tau(X) > 0: 42%
Tau1(X) > 0: 45%
Country: Mexico
Outcome: Log of Ability to recover from shocks
N = 1173, N1 = 507, N0 = 666
Tau(X) > 0: 88%
Tau1(X) > 0: 87%
Country: Mexico
Outcome: Log of Aility to recover from shocks - 2
N = 1173, N1 = 507, N0 = 666
Tau(X) > 0: 82%
Tau1(X) > 0: 82%


Nepal
Distribution of Tau(X)
Tau(X)
Country: Nepal
Outcome: Log of Total Gross Income from all Sources
N = 2848, N1 = 1407, N0 = 1441
Tau(X) > 0: 98%
Tau1(X) > 0: 98%
Country: Nepal
Outcome: Log of Gross Income from Crops and Byproducts
N = 2705, N1 = 1330, N0 = 1375
Tau(X) > 0: 99%
Tau1(X) > 0: 99%
Country: Nepal
Outcome: Log of Gross Income from Livestock and Byproducts
N = 1796, N1 = 929, N0 = 867
Tau(X) > 0: 92%
Tau1(X) > 0: 92%
Country: Nepal
Outcome: Log of Ability to recover from shocks
N = 1304, N1 = 589, N0 = 715
Tau(X) > 0: 34%
Tau1(X) > 0: 33%
Country: Nepal
Outcome: Log of Ability to recover from shocks - 2
N = 1299, N1 = 586, N0 = 713
Tau(X) > 0: 51%
Tau1(X) > 0: 49%

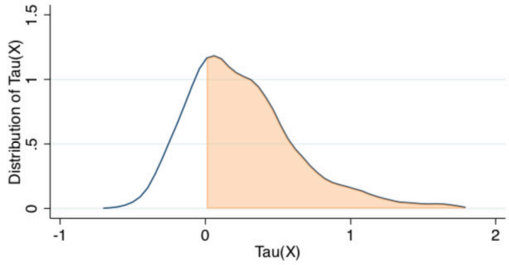
Philippines



Country: Philippines
Outcome: Log of Total Gross Income from all Sources
N = 1394, N1 = 561, N0 = 833
Tau(X) > 0: 33%
Tau1(X) > 0: 35%

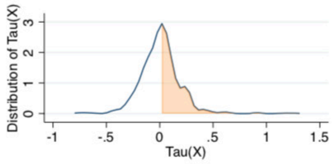


Country: Philippines
Outcome: Log of Gross Income from Crops and Byproducts
N = 973, N1 = 398, N0 = 575
Tau(X) > 0: 61%
Tau1(X) > 0: 64%

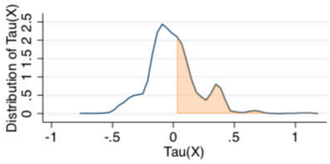


Country: Philippines
Outcome: Log of Gross Income from Livestock and Byproducts
N = 472, N1 = 187, N0 = 285
Tau(X) > 0: 72%
Tau1(X) > 0: 71%

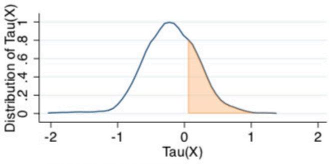
Rwanda_Horticolture



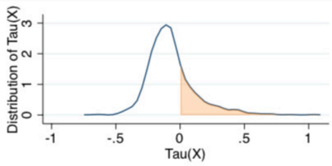
Country: Rwanda_Horticolture
Outcome: Log of Total Gross Income from all Sources
N = 1149, N1 = 250, N0 = 899
Tau(X) > 0: 51%
Tau1(X) > 0: 61%



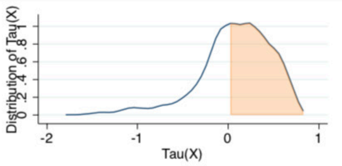
Country: Rwanda_Horticolture
Outcome: Log of Total Net Income from all Sources
N = 1094, N1 = 235, N0 = 859
Tau(X) > 0: 44%
Tau1(X) > 0: 52%



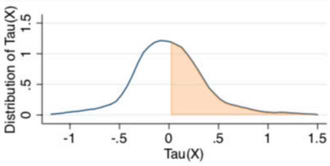
Country: Rwanda_Horticolture
Outcome: Log of Gross Income from Crops and Byproducts
N = 1149, N1 = 250, N0 = 899
Tau(X) > 0: 30%
Tau1(X) > 0: 31%



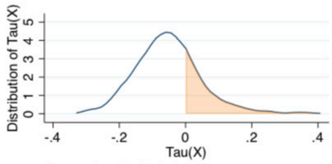
Country: Rwanda_Horticolture
Outcome: Log of Net Income from Crops and Byproducts
N = 1005, N1 = 206, N0 = 799
Tau(X) > 0: 26%
Tau1(X) > 0: 30%



Country: Rwanda_Horticolture
Outcome: Log of Gross Income from Livestock and Byproduct
N = 498, N1 = 103, N0 = 395
Tau(X) > 0: 62%
Tau1(X) > 0: 65%

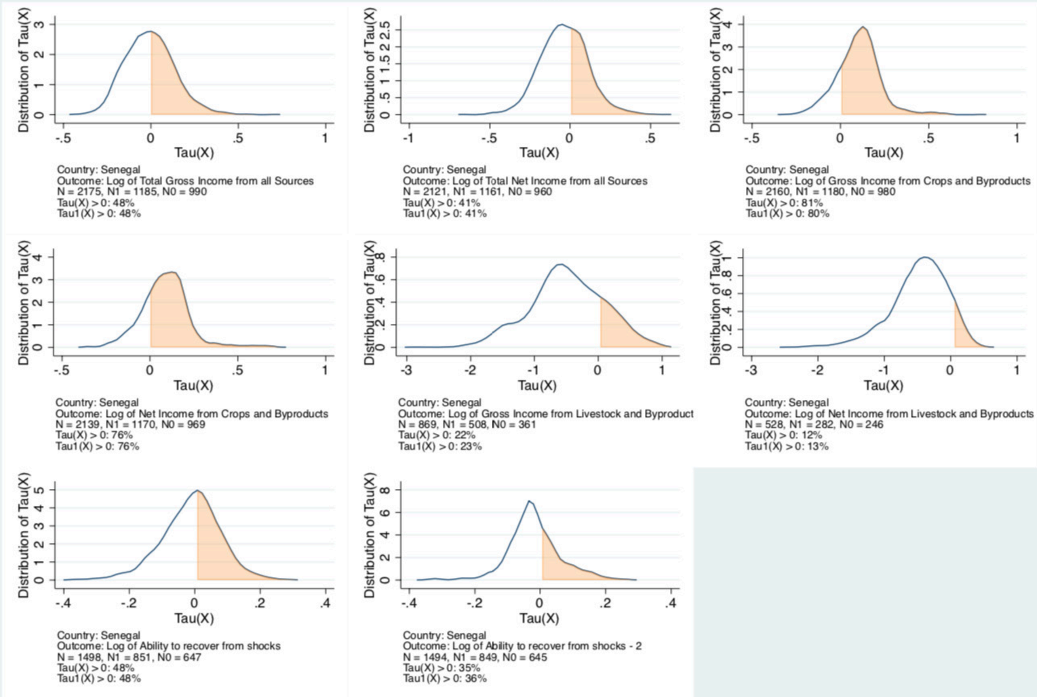


Country: Rwanda_Horticolture
Outcome: Log of Ability to recover from shocks
N = 1148, N1 = 249, N0 = 899
Tau(X) > 0: 49%
Tau1(X) > 0: 44%

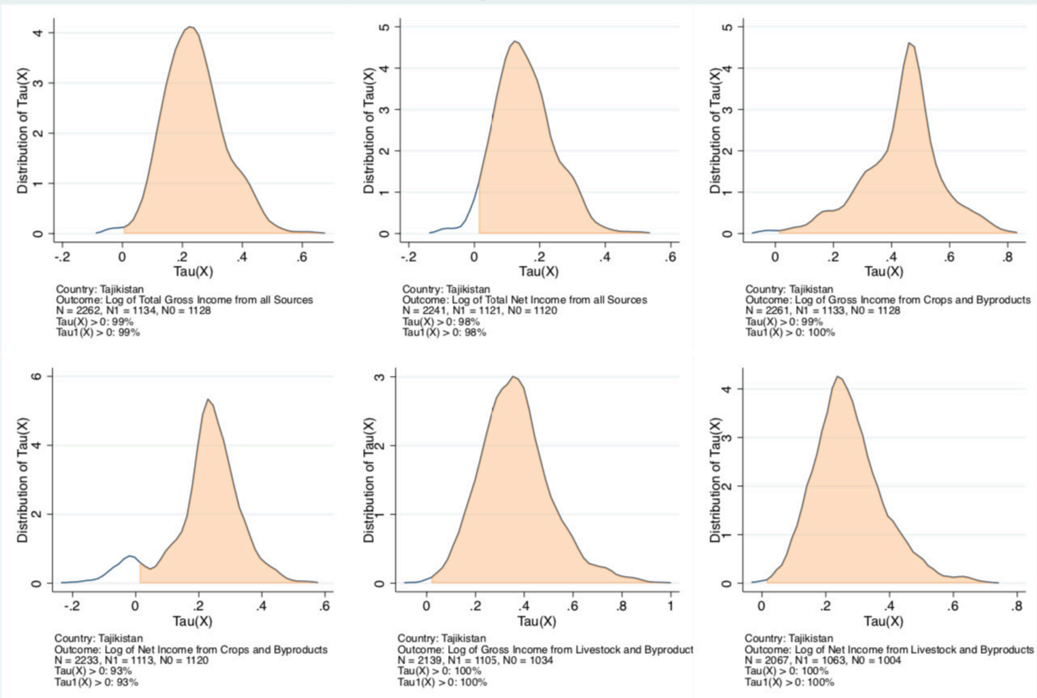


Country: Rwanda_Horticolture
Outcome: Log of Ability to recover from shocks - 2
N = 1117, N1 = 240, N0 = 877
Tau(X) > 0: 27%
Tau1(X) > 0: 21%

Senegal



Tajikistan



International Fund for Agricultural Development
Via Paolo di Dono, 44 - 00142 Rome, Italy
Tel: +39 06 54591 - Fax: +39 06 5043463
E-mail: ifad@ifad.org
www.ifad.org
www.ruralpovertyportal.org
ifad-un.blogspot.com
www.facebook.com/ifad
instagram.com/ifadnews
www.twitter.com/ifadnews
www.youtube.com/user/ifadTV