



The Global
SDG Synthesis
Coalition



Advancing Evidence.
Improving Lives.



Synthesis of Evidence on the Partnership Pillar of the Sustainable Development Goals

METHODOLOGICAL PROTOCOL

Thomas de Hoop, Andi Coombes, Melissa Rodgers, Averi Chakrabarti,
Ozen Guven, Marlous de Milliano, Philomena Panagoulas, Adria Molotsky,
Chinmaya Holla, Talla Cisse, Shalu Jain, Eve Namisango.

August, 2023

SYNTHESIS OF EVIDENCE ON THE PARTNERSHIP PILLAR OF THE SUSTAINABLE DEVELOPMENT GOALS

Copyright © The Global SDG Synthesis Coalition August 2023

This publication, produced by **American Institutes for Research** with inputs from **the Global SDG Synthesis Coalition** partners does not necessarily reflect the views of the individual organizations, entities and bodies partnering with the Global SDG Synthesis Coalition. Individual partners of the Global SDG synthesis Coalition (including staff and personnel) cannot ensure the accuracy, completeness or currency of the data included in this work and they do not assume responsibility for any errors, omissions, or discrepancies in the information, or liability with respect to the use of or failure to use the information, methods, processes or conclusions set forth.

Citation: de Hoop, Thomas, Andi Coombes, Melissa Rodgers, Averi Chakrabarti, Ozen Guven, Marlous de Milliano, Philomena Panagoulas, et al. "Synthesis of Evidence on the Partnership Pillar of the Sustainable Development Goals." New York: The Global SDG Synthesis Coalition, 2023. <https://www.sdgsynthesiscoalition.org/pillar/partnership-pillar>

ACKNOWLEDGEMENTS

The Global SDG Synthesis Coalition would like to thank all those who have contributed to this synthesis.

SDG SYNTHESIS TEAM

Synthesis Team (American Institutes for Research): Thomas de Hoop, Andi Coombes, Melissa Rodgers, Averi Chakrabarti, Ozen Guven, Marlous de Milliano, Philomena Panagoulas, Adria Molotsky, Chinmaya Holla, Talla Cisse, Shalu Jain, and Eve Namisango.

Directorate: Isabelle Mercier (Director) and Alan Fox (Deputy Director)

Synthesis & Lessons Section Chief: Ana Rosa Soares

Evaluation Specialists: Shivit Bakrania and Ben Murphy

Administrative Support: Taeyoung Kim and Kristina Pribilskaya

Publication and Dissemination: Kate Pond and Nthanda Manduwi

The Global SDG Synthesis Coalition could not have conducted this synthesis without the support of the following:

STAKEHOLDERS AND PARTNERS

Management Group: Deborah McWhinney (UNFPA), Jeremy Porzuczek (Global Affairs Canada), Molly Buckley, Katia Rooks), Judith Friedman, Francesca Bonino, Ramona Desole (WFP), Rachel Ingersoll (Department of Foreign Affairs, Ireland), Daniel Jacobo Orea (Ministry of Foreign Affairs, European Union and Cooperation, Spain), Eddie Yee Woo Guo (OIOS), Heewoong Kim (UN Joint SDG Fund), Erik Kinnhammar (UNV), Ada Ocampo (IDEAS International).

Donors: UNDP, UNFPA, WFP, IFAD, Ireland, Spain.

Contents

- Reasons for Conducting an Evidence Synthesis on SDG-17 8
- Conceptual Framework..... 11
- Research Protocol 14
 - Research Questions..... 15
 - Approach to Evidence Synthesis..... 19
 - Impact Evaluation Synthesis 36
 - Performance and Process Evaluation Synthesis 40
 - Synthesis of Systematic Reviews 45
 - Development of Evidence-Gap Maps..... 46
 - Positive Deviance Assessment 47
 - Limitations..... 51
- Work Plan..... 52
 - Quality Assurance 53
 - Coordination with Management Group and Technical Advisory Group 53
- References 55
- Annex A. Search Terms 65
- Annex B. Critical Appraisal Tools 79
- Annex C. Anchor Articles Reference List (organized by SDG-17 content area) 88
- Annex D. SDG-17 Interventions and Outcomes..... 96
- Annex E. Overview of SDG-17 Data 101

Exhibits

Exhibit 1. Methods Overview.....	8
Exhibit 2. Working Conceptual Framework	13
Exhibit 3. Evaluation Matrix.....	15
Exhibit 4. Summary of Final Search String by Sources.....	27
Exhibit 5. Inclusion Criteria.....	29
Exhibit 6: Finance Inclusion Criteria	31
Exhibit 7: Trade Inclusion Criteria.....	31
Exhibit 8: Technology Inclusion Criteria.....	31
Exhibit 9: Capacity Building Inclusion Criteria	32
Exhibit 10: Systemic Issue Inclusion Criteria	32
Exhibit 11. Coding Framework	34
Exhibit 12. Estimation of Odds Ratios	38
Exhibit 13. Draft Deductive Synthesis Framework.....	42
Exhibit 14. OECD DAC Evaluation Criteria and Definitions.....	44
Exhibit 13. Illustrative Example of Evidence-Gap Map	47
Exhibit 14. Phases of the Positive and Negative Deviance Analysis and VNR Assessment	48
Exhibit 15. Work Plan	53
Exhibit B-1: Risk of Bias Tool for Experimental and Quasi-Experimental Studies.....	79
Exhibit B-2: Qualitative Review Protocol.....	84

Introduction

1. The 17 Sustainable Development Goals (SDGs) are central to the 2030 Agenda for Sustainable Development, adopted by all United Nations (UN) Member States in 2015 (United Nations, 2022). The SDGs serve as an urgent call for action to stimulate all countries to work on ending poverty while improving health and education, reducing inequality, stimulating economic growth, and addressing climate change. With 2030 approaching, it is critical to monitor progress toward the SDGs and change course where SDG indicators are lagging in progress. In fact, the Sustainable Development Goals 2022 report (United Nations, 2022, pp. 3) indicates that “the 2030 Agenda for Sustainable Development is in grave jeopardy due to multiple, cascading and intersecting crises” such as the COVID-19 pandemic and climate change.

2. The limited progress toward the SDGs shows the importance of developing evidence-based and practical recommendations for accelerating progress. The SDGs were accompanied by a monitoring and evaluation framework with 248 indicators across each of the SDGs to measure progress. These indicators range from the eradication of extreme poverty for all people everywhere by 2030 (Goal 1.1) to the proportion of countries that have conducted at least one population and housing census in the last 10 years and have achieved 100 percent birth registration and 80 percent death registration (Goal 17.9.2). Measuring progress toward these indicators and assessing which programmes, policies, and interventions are most successful in achieving progress can contribute to the development of evidence-based and practical recommendations for accelerating progress on the SDGs.

3. To contribute to this objective, the Global Coalition for Evaluative Evidence for the SDGs (“the Coalition”) partnered with American Institutes for Research (AIR) to design and implement an evidence synthesis to understand what works, why, and in what context in improving SDG-17 or the Partnership Pillar of the SDGs, the first SDG for which the Coalition has commissioned an evidence synthesis (with others to follow on the Peace, Prosperity, and Planet Pillars of the SDGs). The intended users of the synthesis include UN agencies, UN Member States from high-income, middle-income, and low-income countries, researchers and evaluators, and other stakeholders focused on achieving the SDG-17 objectives.

This report presents the protocol for this evidence synthesis, which will cover the following broad research questions (we present more detailed research sub questions in the section describing the research protocol):

1. Which SDG-17 targets are currently on track and which are lagging?
2. Which countries (across contexts) have made the most progress on SDG-17 and why?
3. Which interventions are most effective in improving and accelerating SDG-17 indicators and targets?
4. How and why are some interventions more successful in achieving progress toward SDG-17-related outcomes?

4. To address these research questions, AIR designed a mixed-methods evidence synthesis that will include (a) a text analysis of voluntary national review (VNR) data combined with in-depth statistical analyses of selected case study countries to address the first two research questions; (b) an evidence synthesis of impact evaluations and specifically randomized controlled trials, quasi-experimental studies, and panel data analyses focused on finance, technology, trade, capacity building, and systemic issues to address the third research question; and (c) an evidence synthesis of process and performance evaluations¹ focused on finance, technology, trade, capacity building, and systemic issues that were independently conducted or commissioned by UN entities and development partners from multilateral or bilateral organizations, civil society organizations, or the private sector to address the fourth research question. Annex D provides details on how these topics relate to targets from the SDG performance measurement framework.

5. We will produce rigorous evidence to generate preliminary observations or lessons learned to accelerate progress toward SDG-17 in time for the High-Level Political Forum on Sustainable Development (HLPF). Since the HLPF will take place in July 2023, we aim

¹ We recognize that some UN agencies use different labels for evaluations with similar objectives as performance and process evaluations. We will also include these evaluations, which may include process tracing, implementation science, formative research, evaluations, or assessments, developmental evaluations, participatory evaluations, midterm evaluations, midterm reviews, or summative evaluations among other evaluation types.

to have preliminary observations or findings in June 2023. In this way, we can draft the PowerPoint presentation for the HLPF by mid-June 2023 and finalize this presentation at the end of June in 2023. Following the HLPF, we will conduct additional analyses to finalize a full evidence synthesis with lessons learned to accelerate progress on SDG-17 by September 2023.

6. Because of this ambitious timeline, we have organized the synthesis to primarily focus on the trade, finance, and technology aspects of SDG-17, as well as statistical capacity building and capacity building for the generation and use of quantitative and qualitative evidence.² Further, we will include evaluations that focus on the role of capacity building and systemic issues in achieving progress toward trade, finance, and technology objectives under SDG-17. Finally, we will generate lessons on the performance of activities related to support for national plans to implement all the SDGs, including through North-South, South-South, and triangular regional and international cooperation (target 17.9 on capacity building) and related to the mobilization and sharing of knowledge, expertise, technology, and financial resources through multi-stakeholder partnerships (target 17.16 on systemic issues).

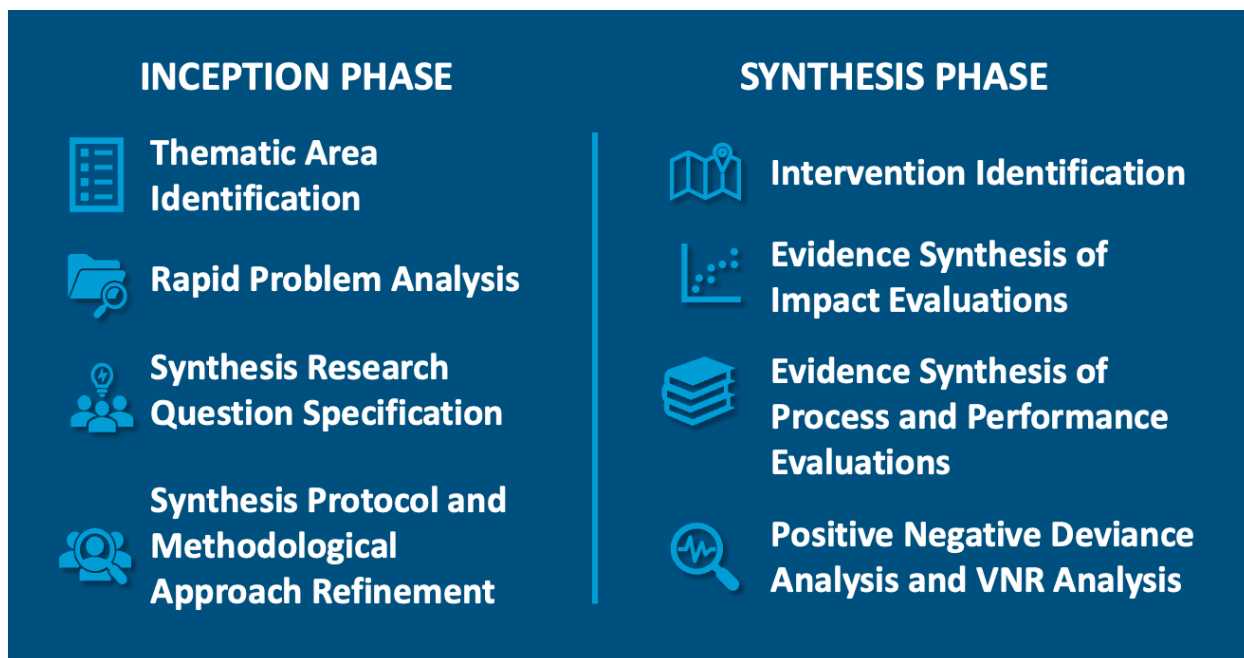
7. A comprehensive scoping of the available literature suggests that a large number of evaluations on capacity building or systemic issues focus on improving other SDGs (e.g., agricultural sustainability, quality education, or ensuring healthy lives and promoting well-being for all). Because of the volume of evaluations of capacity building and systemic issue programmes focused on other SDGs, we will exclusively focus on capacity building activities that target SDG-17 indicators and exclude evaluations of capacity building and systemic issue programmes that primarily focus on other SDG goals. This decision was based on feasibility after several meetings with a management group that includes representatives from UN agencies and Member States that are particularly interested in the evidence synthesis on the Partnership Pillar.

8. Our research design combines a systematic database search, critical appraisal of identified evaluations, and a synthesis of impact, performance, and process evaluations that meet the inclusion criteria. We will triangulate these findings with

² This decision is based on specific requests made by the custodian agencies of SDG-17 and the management group.

country case studies of countries that have made the most progress on SDG-17 using a positive deviance assessment. Exhibit 1 presents a summary of the research design during the inception phase and the synthesis phase.

Exhibit 1. Methods Overview



Note. Subsequent subsections include detailed descriptions of each activity highlighted in this figure.

9. This protocol summarizes the progress we made during the inception phase and the methodology we developed for the synthesis phase.

Reasons for Conducting an Evidence Synthesis on SDG-17

10. We consider it critical to use a rigorous evidence synthesis to assess what works to achieve progress on SDG-17. In contrast to literature reviews conducted to frame the debate or provide context surrounding a discreet theoretical argument, rigorous evidence syntheses aim to systematically synthesize the existing knowledge related to that topic. Evidence syntheses capture and consolidate evidence that is voluminous, diverse, and fragmented across disciplines through a synthesis of “all the existing high-quality evidence using transparent methods to give the best possible generalized statements about what is known” (Waddington et al., 2012, p. 360). Systematic evidence syntheses use “a clear protocol for systematically searching

defined databases, over a defined time period, with transparent criteria for the inclusion or exclusion of studies, as well as the analysis and reporting of results” (Waddington, p. 360). Rigorous evidence syntheses also provide a more comprehensive overview of the existing evidence across contexts than individual evaluations. As a result, evidence syntheses increase external validity relative to evaluations that only cover a limited number of contexts. Because of their increased external validity, evidence syntheses are a powerful tool to examine what works, why, and under what conditions (Waddington et al., 2012).

11. The primary objective of the evidence synthesis is to synthesize evidence for SDG-17. SDG-17, or the Partnership Pillar, emphasizes the importance of revitalizing the global partnership for sustainable development. In this way, SDG-17 can serve as an enabler for achieving progress on the other SDG objectives. For example, improvements in export promotion policies can result in regions specializing in exporting more sophisticated goods under SDG-17. Subsequent increases in the value of exports can in turn result in accelerated, sustained, inclusive, and sustainable economic growth under SDG-8 (i.e., Yao, 2006; Jarreau & Poncet, 2012), which can in turn reduce extreme poverty under SDG-1 (i.e., Dollar & Kraay, 2002). Similarly, improvements in tax collection policies can increase tax revenue (e.g., Khan, Khwaja, & Olken, 2016), which can in turn result in increases in education or health expenditures under SDG-4 and SDG-3. In the technology space, innovation subsidies can increase the adoption of solar energy or other clean technology (e.g., Popp, 2020), which can then result in improved environmental outcomes under SDG-15. By collating and aggregating findings on the effectiveness of SDG-17-related interventions, we will examine whether these hypotheses hold true for different interventions in different contexts.

12. As a secondary objective, the synthesis on SDG-17 will generate lessons learned that can inform the methods and approach for future SDG evidence syntheses commissioned by the Coalition. The Coalition will commission additional evidence syntheses on SDGs related to People, Planet, Prosperity, and Peace. The evidence synthesis on SDG-17 can serve as an input for these future evidence syntheses because SDG-17 serves as an enabler for achieving the other SDG goals. For this reason, we developed a conceptual framework and designed a methodology that can

create synergies with some of the future evidence syntheses (the next section presents more details on the conceptual framework). For example, we will include systematic reviews and other evidence syntheses on the relationship between SDG-17 indicators (e.g., tax revenue, export values, internet access) and other SDG indicators (e.g., economic growth, poverty reduction, education outcomes, health outcomes, women's empowerment). In addition, the coding framework we will use during the coding of papers will account for the role of capacity building and systemic issues in the papers we map and synthesize. In this way, we will not only be able to integrate the role of capacity building and systemic issues related to the evidence synthesis on the Partnership Pillar, but also provide a framework for incorporating the role of capacity building and systemic issues in the future evidence syntheses related to the other SDGs.

13. Finally, the synthesis will incorporate learning goals related to the role of artificial intelligence in evidence syntheses. Specifically, we will pilot the use of the Artificial Intelligence for Development Analytics (AIDA) database, which stores the evaluation reports of the United Nations Development Programme (UNDP).³ This database trains an algorithm to create a feedback loop supporting the machine's learning to optimize the information delivered following a specific search (Garcia, 2022). We aim to pilot AIDA for the screening and analysis of UNDP evaluations and will use the pilot to provide methodological recommendations related to the use of the database for future evidence syntheses.

14. The rest of this protocol is structured as follows. We start with a description of a basic conceptual framework related to SDG-17, which will guide the methodology and specifically the database search, evaluation mapping, and evidence synthesis. Next, we describe the research protocol, which includes methods for screening, coding, analysing, and triangulating evidence from impact, performance, and process evaluations as well as a description of the positive deviance assessment methods. We finalize the protocol with a description of the timeline including the deliverables.

³ In the future, the AIDA database will also store evaluation reports of other UN agencies.

Conceptual Framework

15. During the scoping of the evidence related to SDG-17, we created a working conceptual framework to guide the database search, evaluation mapping, and evidence synthesis. Here, we elaborate on the conceptual framework, which we summarize in Exhibit 2.

16. Inputs, programmes, and policies: SDG-17 has five components: finance, technology, trade, capacity building, and systemic issues. While finance, technology, and trade are discrete policies, programmes, or interventions, most efforts related to capacity building and systemic issues span areas, including those within SDG-17 and across other SDGs. Discrete inputs in finance include, for example, interventions to increase tax revenue, foreign direct investment, and remittances. Technology programmes and policies include interventions to increase internet access or stimulate renewable energy use. Trade interventions focus on increasing exports or enabling free trade by reducing tariffs and other non-trade barriers. Inputs related to capacity building and systemic issues, on the other hand, often support plans to achieve all SDG objectives. Discrete capacity building interventions that relate to specific SDGs could, for example, train government officials to efficiently deliver agriculture or education interventions. On the other hand, capacity building interventions that relate to SDG-17 focus on supporting national plans to implement all the SDGs, including through North-South, South-South, and triangular cooperation (target 17.9), as well as statistical capacity building to increase the availability of high-quality, timely, and reliable data (target 17.18) and capacity building related to the generation and use of quantitative and qualitative evidence. Policies related to systemic issues may include multi-stakeholder partnerships that have discrete goals related to other SDGs (e.g., public-private partnerships to improve health or education outcomes), while other policies related to systemic issues may support policy coherence (target 17.13), implementation of country-owned results frameworks (indicator 17.15.1), or increased involvement of private sector, civil society, and other stakeholders to mobilize and share knowledge, expertise, technology, and financial resources (target 17.16).⁴

⁴ Exhibits 6-10 describe the specific inclusion criteria for the evidence synthesis based on the description of these inputs, programmes, and policies.

17. Moderators and mediators: Our conceptual framework highlights three key moderators and mediators. Moderators are contextual, household-level, or individual-level factors that are hypothesized to determine the effectiveness of SDG-17 programmes, policies, and interventions (e.g., the income status of the country, gender norms in a specific setting, the age of the programme participants), while mediators are generally considered intermediate outcomes that reflect the pathways of the theory of change or the mechanisms through which the programmes, policies, and interventions can achieve their objectives. Anticipated moderators also include (1) programmes that support partnerships between governments, the private sector, and civil society, including multi-stakeholder partnerships, and North-South, South-South, and triangular regional and international cooperation; (2) other capacity building initiatives; and (3) other systemic issues. These factors could each influence positively or hinder the mechanisms through which trade, finance, technology, discrete capacity building, and systemic issue inputs could achieve their objectives. Partnership models that include North-South, South-South, and triangular cooperation, capacity building initiatives such as training governments, and systemic issue policies relating to policy coherence could each increase the effectiveness and impact of trade, technology, and finance interventions. For example, the development of national plans could improve the implementation and effectiveness of export promotion policies. Other moderators may include contextual and other implementation-related considerations, such as the gender, age, and social status of programme participants and contextual characteristics, such as geography (e.g., income and human development status and region). Programmes that support partnerships among governments, the private sector, and civil society could also serve as mediators or intermediate outcomes that contribute to achieving the objectives of other SDG-17 programmes, policies, and interventions. For example, partnerships could facilitate an increase in the commitment of ODA from high-income countries to meet the SDGs or stimulate cross-border trade with low- and middle-income countries (LMICs) after the reduction of tariffs or the introduction of trade agreements.

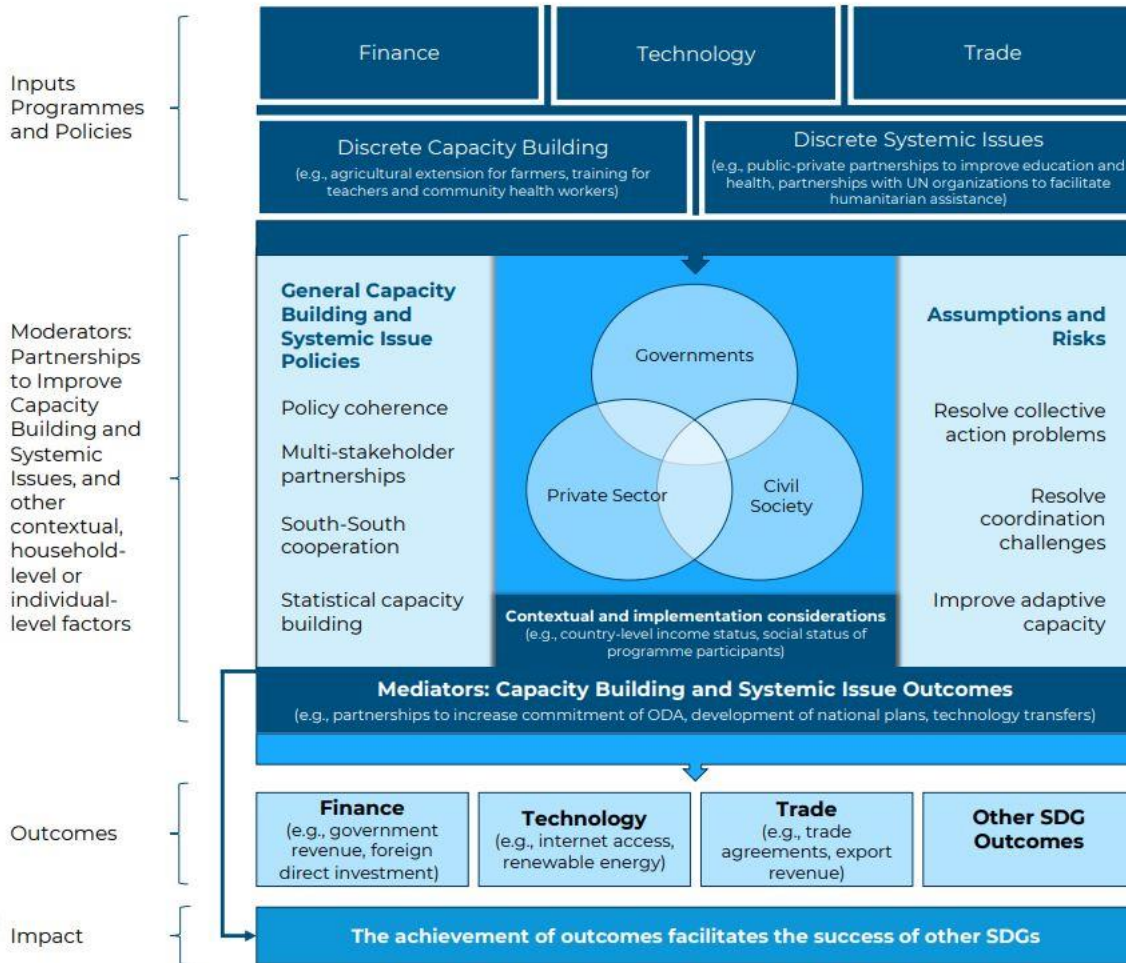
18. The SDG principles of Universality, Coherence, Integration, and Leaving No One Behind, which must consider equity, equality, and non-discrimination, play an important role in the conceptual framework. The effectiveness of finance, trade, and

technology interventions may, for example, depend on the income and human development status of the country, which relates to the Universality principle. It is also critical to consider environmental sustainability and equity when examining the effectiveness of trade, finance, and technology interventions, for example by assessing their potential impacts on environmental outcomes or their effects on inequality.

19. Capacity building and systemic issue interventions will also only increase the effectiveness of trade, finance, and technology interventions if they successfully address several key risks and assumptions. Specifically, it requires that the efforts focused on capacity building and systemic issues resolve collective action problems and coordination challenges and improve adaptive capacity. We highlight these key assumptions in our conceptual framework in Exhibit 2. We will adapt the conceptual framework based on the key findings of the evidence synthesis, which will help in capturing lessons and development of recommendations for policy and practice.

Exhibit 2. Working Conceptual Framework

20. Outcomes and Impacts: We interpret the SDG-17 outcomes as intermediate outcomes because they could enable impacts on other SDG indicators, including improvements in health, education, gender equality, environmental outcomes, economic growth, and poverty reduction. For example, the adoption of clean technology could lead to improvements in environmental outcomes. Similarly, increases in the value of exports could lead to economic growth, and increases in tax revenue or debt relief can enable governments to invest in education or health programmes. In this way, some SDG-17 outcomes can be interpreted as a means toward the end goal of achieving improvements in other SDG indicators. We will assess the achievement of intermediate outcomes both as objectives and as a means toward the end goal of achieving improvements in other SDG goals.



21. Overall, the conceptual framework suggests that SDG-17 may influence other SDGs through environmental, political, economic, and socio-cultural mechanisms. We will consider each of these mechanisms in the analysis and interpretation of the results.

Research Protocol

22. This section presents the research questions; our overall approach to the evidence synthesis, including refining search terms, pilot searching, inclusion criteria, and study screening; and methodological protocols specific to the synthesis of impact evaluations, the synthesis of process and performance evaluations, the analysis of VNR data, and the positive deviance analysis.

Research Questions

23. We will answer the research questions from the evaluation matrix depicted in Exhibit 3:

Exhibit 3. Evaluation Matrix

Research Questions	Methods ⁵
Voluntary national review data analysis	
<ul style="list-style-type: none">• What factors contribute to achieving SDG-17 objectives according to voluntary national review data?• What themes do the voluntary national review data uncover about progress toward the SDG-17 objectives?	<ul style="list-style-type: none">• Text analysis of VNR data, including natural language processing and sentiment analysis
Positive deviance analysis	
<ul style="list-style-type: none">• Which SDG-17 targets are currently on track and which are lagging? How does this differ by country?• How have the COVID-19 pandemic and other crises influenced progress towards SDG-17?• What countries made more progress in achieving progress towards SDG-17? How did their policies and contextual characteristics differ from countries that made less progress in achieving progress towards SDG-17?	<ul style="list-style-type: none">• Statistical analysis of country-level SDG-17 indicators⁶• Desk review of relationship between COVID-19 and SDG-17 indicators• In-depth statistical analyses that include comparisons between four countries (in four different regions) that performed well on SDG-17 indicators with other countries that performed less well (in the same regions)• Desk review of four case study countries that performed well on SDG-17 indicators, including an analysis of the influence of COVID-19 and other crises.
Evidence synthesis of impact evaluations	

⁵ We provide more details on the methods in the section on the approach to the evidence synthesis.

⁶ We will coordinate with the UN Department of Economic and Social Affairs to ensure complementarities between our analyses of these data and separate analyses by the UN Department of Economic and Social Affairs.

Research Questions	Methods ⁵
<ul style="list-style-type: none"> • What is the available evidence on the impact of trade, technology, finance, capacity building, and systemic issues interventions on SDG-17 indicators? • Which trade, technology, finance, capacity building, and systemic issue interventions are most effective in improving and accelerating SDG-17 indicators? • What is the impact of trade policies, programmes, and interventions on SDG-17 indicators (e.g., export values, export diversification, etc.)? • What is the impact of technology policies, programmes, and interventions on SDG-17 indicators (e.g., internet access, mobile banking access, etc.)? • What is the impact of finance policies, programmes, and interventions on SDG-17 indicators (e.g., tax revenue, foreign direct investment, etc.)? • What is the impact of capacity building policies, programmes, and interventions (e.g., support for national plans through North-South partnerships, South-South partnerships, triangular cooperation, statistical capacity building, and capacity building for evidence use) on SDG-17 indicators? • What is the impact of systemic issue policies, programmes, and interventions (e.g., multi-stakeholder partnerships) on SDG-17 indicators? • How do partnerships, capacity building, and systemic issues influence 	<ul style="list-style-type: none"> • Narrative synthesis of experimental and quasi-experimental studies • Meta-analysis of experimental and quasi-experimental studies when more than three studies are available that combine the same interventions and outcomes • Narrative synthesis to analyse potential heterogenous effects by gender, intersectionality, geography (e.g., income and human development status), and COVID-19 incidence if sufficient studies are available • Narrative synthesis to examine the role of partnerships, capacity building and systemic issues in achieving SDG-17 indicators • Risk of bias assessment of experimental and quasi-experimental studies • Triangulation of impact estimates with estimates of costs and cost-effectiveness • Triangulation of experimental and quasi-experimental studies with evidence from performance and process evaluations

Research Questions	Methods ⁵
<p>positively or hinder the impact of trade, finance, and technology interventions?</p> <ul style="list-style-type: none"> • How do gender and age considerations, in policies, programmes, and interventions related to SDG-17 (e.g., gender mainstreaming, the sex and age of programme participants, gender norms, women’s decision-making power, and other gender and age considerations) influence positively or hinder the effects of trade, technology, finance, capacity building, and systemic issue interventions? • How does geography (e.g., income status, human development status, region) influence positively or hinder the effects of trade, technology, finance, capacity building, and systemic issue interventions? • How do COVID-19 incidence and restrictions and other crises influence positively or hinder the effects of trade, technology, finance, capacity building, and systemic issue interventions? 	
<p>Evidence synthesis of performance and process evaluations</p>	
<ul style="list-style-type: none"> • Why and how are some interventions more or less successful in achieving progress towards SDG-17? • What does the evidence say about what does and what does not work to implement effective trade, finance, technology, capacity building, and systemic issue policies, programmes, and interventions under SDG-17? How 	<ul style="list-style-type: none"> • Qualitative thematic analysis combining deductive (top-down) and inductive (bottom-up) approaches • Analysis of UN capacity frameworks to assess the role of capacity building in achieving SDG-17 indicators

Research Questions	Methods ⁵
<p>equitable and environmentally sustainable are the trade, finance, technology, capacity building and systemic issue policies, programmes, and interventions that work?</p> <ul style="list-style-type: none"> • Under what conditions (e.g., income status, human development status, region, gender, age) were the interventions most effective? • What are the main obstacles that need to be removed to unleash the full potential of the Partnership Pillar of the SDGs? To what extent do partnerships encourage ownership, alignment, harmonization, results, and mutual accountability (Paris Declaration) in the areas of trade, technology, and finance? • How do partners approach capacity building and systemic issues in the areas of trade, technology, and finance? • What is the available evidence on how to achieve progress on statistical capacity building and capacity building on the generation and use of evidence? • How do partnerships help address systemic issues such as policy coherence, policy coordination, infrastructure, and country planning in the areas of trade, technology, and finance? 	<ul style="list-style-type: none"> • Quality appraisal of performance and process evaluations • Triangulation of evidence from performance and process evaluations with impact evaluation estimates
<p>Evidence synthesis of systematic reviews, evidence syntheses, and other comprehensive literature reviews</p>	

Research Questions	Methods ⁵
<ul style="list-style-type: none"> • What is the available evidence on the impact of trade, technology, finance, and capacity building interventions related to statistics, evidence generation and use on other SDG goals (e.g., economic growth, poverty reduction, food security, nutrition, education, health, environmental outcomes, gender equality)? 	<ul style="list-style-type: none"> • Synthesis of systematic reviews and evidence syntheses on the relationship among trade, technology, finance, and capacity building interventions related to statistics, evidence generation, and use and other SDG indicators • Triangulation of evidence from systematic reviews with data from the synthesis of impact evaluations

Note: For each research question, we will first assess the available evidence. If the available evidence is too thin to generate reliable conclusions about the research question, we will produce evidence-gap maps.

Approach to Evidence Synthesis

24. We will use transparent inclusion and exclusion criteria to increase internal and external validity and to ensure that claims about a range of studies are representative of the existing high-quality evidence on SDG-17. Typical literature reviews can provide biased views of the state of the evidence without transparent inclusion and exclusion criteria and are thus not sufficient to identify the state of the current evidence and the evidence gaps in research on SDG-17. Further, individual evaluations usually have a limited external validity because they are not able to provide a comprehensive overview of the literature. Their lack of external validity also may limit the ability of individual evaluations to assess how contextual considerations from different settings outside the evaluation context matter for the relevance, effectiveness, efficiency, sustainability, and impact of trade, technology, finance, capacity building, and systematic issue interventions at the global level.

Inclusion and exclusion criteria to guide the systematic database search and screening

25. We developed and refined inclusion and exclusion criteria and a search strategy, relying on population, indicator, comparison, and outcome (PICO) criteria, as well as consultations with the Coalition and Management Group in Washington DC and New York. During this process, we selected the most relevant websites and databases for

our review of impact, performance, and process evaluations. These websites and databases included the Web of Science, the 3ie Impact Evaluation Repository, the United Nations Evaluation Group (UNEG), and ERC databases of the United Nations, the DEREc and IEG databases with evaluations of bilateral partners and the World Bank, and the ALNAP database with evaluations of humanitarian programmes.

26. Next, we describe the inclusion and exclusion criteria based on the scoping during the inception phase.

Overview of inclusion criteria

27. Populations of Interest: While we will include some evidence related to interventions for populations of high-income countries, we will **primarily** select studies that cover interventions for populations of LMICs. Our search strings for impact evaluations will include keywords to identify studies that look at these populations (such as “low-income countr*” and “low-income econom*”) as well as the names of all the countries in these income categories and terms to capture their demonyms (such as “Azerbaijan” and “Azeri”).⁷

28. However, we will include evaluations of programmes, policies, and interventions in high-income countries that focus on SDG-17 indicators with an emphasis on relationships between high-income countries, and LMICs. Such evaluations may, for example, include a focus on official development assistance (ODA) and its impact on SDG-17 indicators, changes in the definition of ODA in high-income countries, or the number of countries that adopt and implement investment promotion regimes for least-developed countries. We will not include studies focused on the impact of ODA on indicators that do not focus on SDG-17 (e.g., economic growth, poverty reduction, education, and health outcomes).

29. Depending on the number of studies we encounter, we will also consider including performance and process evaluations focused on high-income countries from the DeREC database. Such evaluations may, for example, focus on trade between different countries in the European Union. However, at this moment we have not yet conducted

⁷ We will not include countries in the search string for the 3ie database because this database only includes evidence from LMICs.

a comprehensive assessment of the number of studies focused on high-income countries in the DeREC database. We will assess our ability to include evidence from high-income countries from the DeREC database after a comprehensive assessment. Depending on how many studies we find, we may decide to only include performance and process evaluation evidence from after 2018 (as opposed to 2015) to limit the number of evaluations to review and to enable the inclusion of some evidence from high-income countries.

30. While we recognize the universality principle of the SDGs, we do not consider it feasible, realistic, or methodologically appropriate to treat evidence from high-income countries the same as evidence from LMICs (unless the evidence from high-income countries includes LMICs as part of the scope). The main reason is that the volume of research from high-income countries is much larger than the volume of evidence from LMICs. Research output on a given country increases considerably with the country's wealth, as shown by Das et al. (2013) who use a database of 76,046 empirical economics papers to demonstrate that "over the 20-year span of the data, there were 4 empirical economics papers on Burundi, 9 on Cambodia and 27 on Mali. This compares to the 37,000 or so empirical economics papers published on the U.S. over the same time period." More recent evidence also shows that, of the research in health economics journals, only 2 percent covers LMICs (Hirvonen, 2020). Including all evidence from high-income countries would lead to skewed results with a disproportionate focus on evidence from high-income countries, such as the United States. For this reason, treating evidence from high-income countries the same as evidence from LMICs will likely result in misleading policy recommendations that primarily apply to high-income countries. However, we will include evidence on high-income countries that relates to LMICs and possibly some evidence specific to high-income countries from the DeREC database.

31. Given that SDG-17 indicators primarily emphasize country or higher-level processes (such as South-South cooperation, interventions to increase tax revenue, and tariff policies), we will primarily focus on macro-level interventions. For example, we will exclude micro-level interventions related to finance (e.g., microfinance, self-help groups, cash transfer programmes, vocational and business training) from the purview of the current evidence synthesis because they focus on other SDG goals (i.e.,

poverty reduction and gender equality). However, we will include micro-level interventions that aim to improve SDG-17 indicators (e.g., programmes that aim to increase the volume of remittances by providing information to migrants or increase the volume of exports through export promotion policies).

32. Comparison/Evaluation: We will include impact evaluations that have either an experimental or a quasi-experimental design. Studies with an experimental research design use random assignment to the intervention (as part of randomized controlled trials). Quasi-experimental studies eligible to be included in this evidence synthesis include regression discontinuity designs, difference-in-differences analyses, instrumental variable analyses, and matching based on propensity scores. Quasi-experimental studies should leverage longitudinal data (baseline and endline data) and at least one comparison group to account for counterfactual trends (other things occurring concurrently with the treatment) and selection issues (systematic differences between those receiving and not receiving an intervention) to be included. Performance and process evaluations will not require a control or comparison group. However, the performance or process evaluation needs to self-identify as an evaluation to be included (this may include evaluations using alternative terms such as mid-term, participatory, strategic, institutional, programme, policy, summative, or global evaluations). For example, we will not include ethnographic studies that do not identify as an evaluation.

33. For performance and process evaluations, we anticipate including three broad categories of studies, provided they meet high-quality criteria according to a qualitative appraisal tool adapted from quality assessment tools used across UN Agencies. The first category includes performance/process evaluations and self-identified cost-effectiveness analyses linked to the experimental and quasi-experimental studies identified for inclusion in this synthesis. We refer to these studies as sibling studies.

34. The second category includes performance/process evaluations and costing/cost-effectiveness analyses commissioned or conducted by the UN, funds and facilities working with the UN (i.e., the Joint SDG fund), and bilateral agencies that meet minimum quality criteria. For example, UNICEF and UNDP have already

ascertained the quality of evaluation reports through the Global Evaluation Report Oversight System (GEROS) (UNICEF, 2020) and we anticipate that we will only include evaluations that were rated Highly satisfactory or Satisfactory (i.e., reports that meet or exceed the UNICEF and UNDP standards for evaluation reports). Other UN agencies, such as UNFPA and WFP, use similar tools. We will use similar quality criteria for evaluations commissioned or conducted by other UN agencies, such as UNFPA, UN Women, the World Food Programme, etc. Regardless of the quality score determined by the different UN agencies, each of the performance and process evaluations will still require minimum quality criteria according to the quality appraisal tool for performance and process evaluations, which we adapted based on quality assessment tools across UN agencies. We present the adapted quality appraisal tool in Annex B.

35. Third, we will also include quality performance/process evaluations and cost-effectiveness analyses funded by other agencies, either from multilateral/bilateral agencies, civil society, or the private sector, even when these are not specifically linked to an impact evaluation. These evaluations will again have to meet minimum quality criteria. We will request more details on the methodology when evaluations from bilateral agencies do not include sufficient detail in their final report.⁸

36. Using a uniform search string for impact, performance, and process evaluations will enable us to conduct a single comprehensive search per SDG-17 area. As described below, once relevant studies have been identified, we will discern the different types of studies (impact from process/performance) when extracting information from them.

37. Interventions and Outcomes: Because the SDG-17 outcomes are closely aligned or sometimes synonymous with SDG-17 interventions, we will use a single block of keywords for interventions/outcomes under each SDG-17 area— trade, technology, finance, capacity building, and systemic issues. For example, while tariffs are usually considered a policy, the SDG-17 indicators include a reduction in tariffs as an outcome. Similarly, debt relief is considered an outcome under SDG-17, even though it is commonly considered a policy.

⁸ UN agencies and Member States can submit relevant reports to AIR if reports are not publicly available.

38. Outcomes: We will include impact, performance, and process evaluations focused on outcome related to SDG-17, but we will not include primary studies focused on the impact of trade, technology, finance, capacity building, and systemic issue interventions on outcomes related to other SDGs. For example, we will include evaluations that focus on the effects of trade, technology, finance, capacity building, or systemic issue interventions on export competitiveness, but we will not include evaluations that focus on the impacts of trade policies on poverty alleviation. Examples of SDG-17 outcomes in our final search string include tax revenue, debt relief, technology diffusion and duty free.

39. Inclusion of Systematic Reviews: While we will not include impact evaluations related to other SDG outcomes, we do intend to include systematic reviews and other evidence syntheses related to the impact of trade, technology, finance, and statistical capacity building interventions on other SDG outcomes. In this way, we can generate some hypotheses for the future evidence syntheses related to the SDGs. For example, we will include systematic reviews on the impact of trade liberalization on economic growth, poverty reduction, and environmental outcomes as well as evidence syntheses on the impact of foreign direct investment policies on poverty reduction.

40. Interventions: We will use separate intervention keywords for each of the SDG-17 areas included in the evidence synthesis. Examples of trade interventions we cover include export processing zones and export subsidies. Examples of finance programmes include environmental taxes, remittances, and overseas development flows. Examples of technology interventions include programmes that focus on enhancing internet access and clean energy. Examples of capacity building interventions include national plans to implement all the SDGs, including through North-South or South-South collaboration initiatives to stimulate knowledge transfer, while examples of systemic issues interventions include interventions to mobilize and share knowledge, expertise, technology, and financial resources through multi-stakeholder partnerships. We will exclude interventions such as technology-aided instruction programmes that align with SDG-17 topics (in this case, technology) but that are more aligned with other SDGs. For example, technology-aided instruction also falls under SDG-4 to ensure inclusive and equitable quality education. Similarly, MHealth programmes focus on health outcomes and agricultural technology usually

focuses on agricultural outcomes. We decided to exclude such interventions because we anticipate their inclusion in future evidence syntheses. We also decided to exclude microfinance, self-help group, and cash transfer interventions because these programmes are more closely aligned with SDG-1 (concerned with the elimination of poverty) and SDG-5 (related to gender equality and women's empowerment). Finally, we will also exclude capacity building and systemic issue interventions that primarily aim to influence other SDG goals. Exhibits 6 to 10 explicitly define our eligibility criteria for interventions.

Search strategy

41. While we had originally proposed a single string for an overall SDG-17 search, we split the keywords for interventions of interest into five groups, one for each SDG-17 area – finance, technology, trade, capacity building, and systemic issues. During our scoping of the evidence, we experimented with the searches to obtain a comprehensive but manageable number of studies.

42. Starting with the search string in Annex A, we piloted the search strategy in the following databases: EBSCO, EconLit, Web of Science, and 3ie Development Evidence Portal. We worked with a librarian to refine the search strings by adding qualifying terms to keywords (e.g., added evaluation to performance to search for “performance evaluation”); eliminating superfluous words (e.g., “implementation,” since phrases such as “implementation science” were already in the string); and introducing punctuation to make the string more targeted (e.g., including quotations around phrases such as “capacity strengthening” to return papers that had the full phrase instead of just “capacity” or just “strengthening”). We also dropped terms that the Coalition suggested were beyond the scope of the current evidence synthesis (e.g., “cash transfers” and “fair trade”).

43. We tailored the search strings to specific database requirements. For example, as the 3ie database only covers research conducted in LMICs, we omitted the country names from the search string for this database. The UNEG database is different from academic databases in that it allows for searching only a single keyword at a time and does not permit filtering studies by a range of years. Given these features, we conducted test searches using several of the intervention/outcome terms within each

of the five SDG-17 areas. We then exported the results of these searches and deduplicated reports identified for each of the five areas. UNEG search results include the title, authors, and direct links to reports, and omit the typical keyword and abstract information. As such, this information is not structurally compatible for the software programmes typically used to screen and code studies for evidence syntheses (such as EPPI-Reviewer). During the pilot, we thus had two screeners review all unduplicated UNEG reports using an alternative custom-made tool in the AirTable software to determine eligibility. We are now exploring options to include reports from the UNEG database in EPPI reviewer.

44. We identified anchor/pivotal papers for each of the five SDG-17 areas to ensure our search strings returned relevant documents in all databases. The anchor papers are either systematic reviews that provide critical evidence related to SDG-17, highly cited individual studies related to SDG-17, or reports suggested by the custodian agencies. During an initial scan of academic databases, our team identified primary impact studies, performance and process evaluations, and systematic reviews of relevant work within each SDG-17 area. We shared this initial list of papers with the Coalition for feedback and requested that they identify additional papers that they would expect to include a synthesis of the evidence on each SDG-17 area (including studies from UN agencies). Finally, our team extracted full-text articles for the final list of anchor studies (see Annex C: Anchor Papers Reference List). We identified at least nine anchor papers for each area, with a minimum of three published in academic, peer-reviewed journals, and the remaining from the custodian agencies of SDG-17 (e.g., World Bank, IMF, Paris21, OECD).

45. We also leveraged the anchor papers to increase the comprehensiveness of our search strings. To achieve this goal, we extracted the titles, keywords, and abstract or executive summary from the articles and reports in Zotero, a reference management tool. We then analysed the bibliographic information with an R package called litsearchr (Grames et al., 2019) that uses text-mining algorithms to analyse bibliographic information to produce a list of phrases that appear frequently in the papers. We then included additional terms that litsearchr identified via a scan of the title and abstracts of the anchor papers.

46. Subsequently, we re-ran our updated search in three key databases (i.e., Web of Science, 3ie, and UNEG) and screened the first few pages of results for each SDG-17 area to ensure the search strings identified relevant articles. We further refined inclusion/exclusion criteria after experimenting with keyword inclusions and exclusions. For example, we identified the search terms detecting irrelevant articles and excluded these terms (e.g., “assessment”). This iterative process resulted in the final search strings that we present in Annex A. Exhibit 4 summarizes the results of this final search string for each of the pilot databases by topic.

Exhibit 4. Summary of Final Search String by Sources

SDG-17 Component	Web of Science	3ie	UN Group	Evaluation
Finance	805	143	36	
Trade	3,461	229	37	
Technology	13,830	551	211	
Capacity building	1,038	87	118	
Systemic Issues	1,488	110	43	

Note 1. Web of Science terms contain intervention/outcome, evaluation/comparison, and population search terms. 3ie terms contain intervention/outcome and evaluation/comparison terms due to database limitations. UNEG terms include intervention/outcome terms due to database limitations.

Note 2. 3ie values contain duplicates due to database limitations. We initially found a higher number of studies in the 3ie database because of misunderstandings about the most efficient way to search in the 3ie database repository.

Scoping of the evidence

47. We selected trade, technology, and finance as the three focus areas under SDG-17 based on a comprehensive scoping of the evidence during the inception phase. These are the areas on which there appears to be a wealth of evidence from different types of studies that are directly related to SDG-17 and less to other SDGs. Accordingly, there is substantial potential to speak to what works, how, why, and under what circumstances for these SDG-17 components. The SDG-17 goals and indicators for capacity building and systemic issues are often applicable across the different SDGs, meaning that the studies returned for these areas often covered topics that are better suited under other SDGs (e.g., capacity building efforts in health or education). We will

include studies or evaluations related to capacity building and systemic issues that focus on SDG-17 indicators, such as support for national plans to implement all the SDGs, including through North-South, South-South, and triangular regional and international cooperation (target 17.9 on capacity building) and related to the mobilization and sharing of knowledge, expertise, technology, and financial resources through multi-stakeholder partnerships (target 17.16 on systemic issues), but not when those studies or evaluations focus on other SDG indicators.

48. We will also synthesize the evidence on capacity building and systemic issues as a mechanism for achieving impacts for trade, technology, and finance. Further, we will tag articles that are relevant for capacity building in other SDG areas to facilitate future syntheses. In addition, we decided to place special emphasis on capacity building for statistics or other types of evidence use, North-South partnerships, South-South partnerships, triangular cooperation, and multi-stakeholder partnerships based on conversations with the custodian agencies and the management group.

49. For capacity building and systemic issues, we decided with the management group to also pilot an approach to synthesize information that can be adapted and used in future syntheses related to the other SDGs. Doing so will not only facilitate the work of future syntheses, but also set up the potential for comparison of capacity building and systemic issues across SDGs.

50. In the next phase of this review, we will proceed with screening results from the databases listed above, while also considering the inclusion of additional databases (EconLit and ALNAP) and specific journals (e.g., Journal of International Economics, Journal of Developmental Economics, Journal of Developmental Studies, World Bank Economic Review). We will also use the DeREC and IEG databases to obtain evidence from bilateral and multilateral donors. Finally, we will use Google and Google Scholar for forward and backward citation searches and for obtaining evaluations commissioned or conducted by LMIC governments.

51. We will also pilot approaches for generating lessons about the use of artificial intelligence in systematic reviews and evidence syntheses by searching the AIDA database. This database currently houses UNDP evaluations and uses artificial intelligence to train an algorithm to create a feedback loop supporting the machine's

learning to optimize the information delivered following a specific search (Garcia, 2022).⁹ The information provided by AIDA summarizes information from full reports by providing paragraphs with findings, conclusions, and recommendations (based on a machine learning algorithm). While these paragraphs do not by themselves allow for an assessment of the quality of the available evidence, it is possible to link the paragraphs to quality assessments of full reports from which AIDA obtains the information. In addition, providing paragraphs could help the synthesis team review studies more efficiently, as it may limit the need to read full reports after the quality assessment. For this pilot, the use of AIDA will be restricted to UNDP evaluations.

52. Based on the search and screening process described above, we refined our proposed inclusion criteria and finalized the criteria for inclusion in critical appraisal as presented in Exhibit 5. During the screening phase of the evidence synthesis, we will screen all studies highlighted by the search and exclude those that do not meet the inclusion criteria. We anticipate that two members will work independently to screen a sub-sample of the abstracts. We will then discuss the results and continue with a single rater after we achieve sufficient interrater reliability (>0.95).

53. To create efficiencies, we will also use machine learning approaches in EPPI Reviewer to screen out studies that are clearly unrelated to SDG-17. Exhibits 6, 7, 8, 9, and 10 present inclusion criteria for the trade, finance, technology, capacity building, and systemic issue components. Annex D presents more details on the specific outcomes.

Exhibit 5. Inclusion Criteria

Topic	Inclusion Criteria
Publication dates	Published between 2015 and 2022 and possibly between 2018 and 2022 for the DeREC database
Publication accessibility	Published in English, Spanish, or French Publicly available or shared with the synthesis team

⁹ Other UN agencies may add their evaluation reports to AIDA in the near future.

Topic	Inclusion Criteria
Intervention focus	Targets interventions with clear objectives and strategies that are related to the Partnerships Pillar (i.e., SDG-17), components finance, trade, technology, and statistical capacity building.
Evaluation focus	Assesses the relevance, effectiveness, efficiency, sustainability, and impact of a policy, programme, or intervention in achieving SDG-17 related objectives, focusing on trade, technology, finance, and statistical capacity building outcomes
Level of focus	<p>Primary evaluations: focus interventions and outcomes that strictly pertain to SDG-17 components</p> <p>Secondary evaluations (e.g., systematic reviews, evidence syntheses, literature reviews): focus on other SDG outcomes (i.e., economic growth, poverty reduction, gender equality, education, health)</p> <p>Exclude: primary evaluations that focus on other SDG pillars and interventions such as microfinance, cash transfers, self-help groups, savings groups that seek to improve other SDG outcomes (e.g., education, health, poverty reduction)</p>
Population of interest	<p>Includes populations in low- and middle-income countries for finance, trade, technology, and statistical capacity building interventions</p> <p>Includes populations in high-income countries for SDG-17 indicators with an emphasis on relationships between high-income countries and LMICs, and potentially for evaluations from the DeREC database.</p>
Method	<p>For evaluations on the impact of SDG-17 interventions on SDG-17 outcomes: include impact evaluations (RCT or quasi-experimental study with a comparison group), regression analyses with a comparison group and panel data, cross-country regressions that use panel data.</p> <p>For evaluations on the link between trade, technology, and finance and capacity building and systemic issues: include impact evaluations, process evaluations, and performance evaluations using primary data</p>

Topic	Inclusion Criteria
	For evaluations on the impact of SDG-17 interventions on other SDG outcomes: include systematic reviews and other evidence syntheses

Note: These criteria are for all papers regardless of topic.

Exhibit 6: Finance Inclusion Criteria

Topic	Inclusion Criteria
Topics included	Taxes, government finance, debt, debt relief, remittances, ODA
Topics excluded	Microfinance, cash transfers, self-help groups, savings groups
Primary outcomes	Government revenue, debt, remittances, foreign direct investment
Capacity development	Any capacity building geared toward government personnel or institutions that strengthens finance-related SDG-17 outcomes
Systemic issues	Any systemic issues that facilitate or undermine progress on finance-related SDG-17 outcomes

Annex D presents more details on the specific outcomes

Exhibit 7: Trade Inclusion Criteria

Topic	Inclusion Criteria
Topics included	Trade barriers, export promotion, trade liberalization
Topics excluded	Fair trade interventions
Primary outcomes	Trade agreements, export value, export revenue, export diversification, tariff values, and foreign direct investment
Capacity development	Any capacity building geared toward government personnel or institutions that strengthens trade-related SDG-17 outcomes
Systemic issues	Any systemic issues that facilitate or undermine progress on trade-related SDG-17 outcomes

Annex D presents more details on the specific outcomes

Exhibit 8: Technology Inclusion Criteria

Topic	Inclusion Criteria
-------	--------------------

Topics included	Technology access, transfer, and innovation, internet access, mobile money, mobile phones
Topics excluded	Technology-aided education, mHealth interventions
Primary outcomes	Renewable energy technology, technology related to public infrastructure development, internet access, mobile phone access, access to mobile money
Capacity development	Any capacity building geared toward government personnel or institutions that strengthens technology-related SDG-17 outcomes
Systemic issues	Any systemic issues that facilitate or undermine progress on technology-related SDG-17 outcomes

Annex D presents more details on the specific outcomes

Exhibit 9: Capacity Building Inclusion Criteria

Topic	Inclusion Criteria
Topics included	Support for national plans to implement all the SDGs, including through North-South, South-South, and triangular regional and international cooperation, statistical capacity building, and capacity building related to the generation and use of quantitative and qualitative evidence
Topics excluded	Capacity building initiatives that primarily focus on other SDG goals, such as capacity building to improve agriculture, education, and health outcomes
Primary outcomes	Development of national plans, value of financial and technical assistance to LMICs, quality of implementation of national plans

Annex D presents more details on the specific outcomes

Exhibit 10: Systemic Issue Inclusion Criteria

Topic	Inclusion Criteria
Topics included	Public-private partnerships, other multi-stakeholder partnerships, policy coherence and coordination, the mobilization and sharing of knowledge, expertise, technology, and financial resources through multi-stakeholder partnerships, country-owned results frameworks

Topics excluded	Systemic issue initiatives that primarily focus on other SDG goals, such as multi-stakeholder partnerships to improve agriculture, education, and health outcome
Primary outcomes	Progress in multi-stakeholder development effectiveness frameworks, quality or fidelity of implementation of public-private partnerships, quality or fidelity of implementation of other multi-stakeholder partnerships, implementation of country-owned results frameworks, sharing of knowledge, expertise, technology and financial resources to support the achievement of the SDGs.

Annex D presents more details on the specific outcomes

54. During the title and abstract reviews using the above criteria, reviewers will select “yes,” or “no,” in EPPI-Reviewer. If a reviewer marks “yes” for any of the criteria, the reviewer will continue to the next criterion on the coding sheet. If the reviewer marks “yes”, the study qualifies for the review of the full text. If a reviewer marks “no”, the study does not meet the criteria for further review. If reviewers disagree, the study is tagged for reconciliation.

55. We plan to code each study based on the key indicators in Annex D using a coding sheet in EPPI-Reviewer web software for review management and coding. The indicators largely align with those used to tag papers in UNDP’s AIDA database, but we may include other indicators after screening. We will use the AIDA and other categories to manually tag papers that are not already tagged as part of the UNDP database and code each included study by beneficiary characteristics, geography, evaluation methodology, SDG indicator, and whether the study is a potential sibling study of an impact evaluation. We will code for the inclusion of vulnerable populations to enable the estimation of heterogeneous effects for these subgroups for all full-text studies. We will, for example, code for ethnicity, caste, indigenous group, poverty levels, and potentially intersectionality. Exhibit 11 presents the coding framework.

Exhibit II. Coding Framework

Thematic Category	Definition of Thematic Category
Evaluated intervention	
Topic Level 1 ("parent category")	E.g., financing, governance, SDG; categories aligned with those used in AIDA database
Topic Level 2 ("child category")	E.g., poverty alleviation, capacity development; categories aligned with those used in AIDA database
SDG	Tag most relevant to SDG
SDG targets	If SDG-17, tag relevant target
SDG indicator	If SDG-17, tag relevant indicator
Region	Region where the evaluated intervention was implemented
Country	Country (or countries) where the evaluated intervention was implemented
Income level	Low income, lower-middle income, upper-middle income
Focus on disadvantaged group(s) and younger populations	Indigenous, women, low-income, disabled, youth, caste, ethnicity, youth, other age groups, none
Capacity building efforts	Whether paper explicitly addresses capacity building efforts
Systemic issues addressed	Whether paper explicitly addresses systemic issues
Partnership efforts assessed	Whether paper explicitly addresses partnerships
Evaluation study	
Type of evaluation*	Process evaluation, performance evaluation, impact evaluation
Language of evaluation	English, French, Spanish
Evaluation year	Enter the evaluation year

Thematic Category	Definition of Thematic Category
Quality rating (<i>only for performance and process evaluations</i>)	High, medium, low, based on score ranges from qualitative quality appraisal tool
Risk of selection bias (<i>only for impact evaluations</i>)	High, medium, low, based on risk of bias assessment
Risk of performance bias (<i>only for impact evaluations</i>)	High, medium, low, based on risk of bias assessment
Included in evidence synthesis	Yes or No
Reason for exclusion	Explanation for why we excluded studies after closer examination

* We recognize evaluations may be tagged as implementation science, implementation fidelity, developmental, formative, and participatory evaluation. As we review the studies, we will make a note of the relevant evaluation typologies to the extent they are useful for improving programme implementation.

56. We will use two assessment tools to assess the methodological quality of included studies (see Annex B for the tools):

- A risk of bias assessment tool to appraise the quality of impact evaluations
- A qualitative review tool to assess the quality of performance and process evaluations

57. We will use different tools for impact evaluations and performance and process evaluations because, although there is overlap among the tools, each tool addresses unique methodological details not covered in the other tools. We anticipate that two members will work independently to apply the critical appraisal to a sub-sample of the full-text studies. We will then discuss the results and continue with a single rater after we achieve sufficient interrater reliability (>0.95).

58. At this moment, we anticipate that we can apply the critical appraisal tools for most of the included studies before the HLPF. However, we may have to postpone the

critical appraisal of some studies until after the HLPF depending on the number of included studies.

59. The following sections detail the approaches for impact and process and performance evaluation syntheses.

Impact Evaluation Synthesis

Data extraction

60. Two team members with expertise in impact evaluations will independently extract information from each experimental or quasi-experimental study included in the review. Both team members will use a data extraction form and fill the data in a table.

61. Where available, we will also code information related to the inclusion of other SDG outcomes. We will code this information to guide future evidence syntheses. However, we will not synthesize information related to other SDG outcomes.

Effect size calculations

62. To synthesize the quantitative studies, we will use the extracted information from each experimental or quasi-experimental study to estimate the standardized effect sizes (for continuous variables) or odds ratios (for binary variables) across studies. In addition, we will calculate standard errors and 95 percent confidence intervals, where possible. This section presents the process to calculate effect sizes and is heavily based on Brody et al. (2015).

63. We will estimate effect sizes as thoroughly as we can before the HLPF presentation but will prioritize effect-size calculations for interventions for which more than three impact evaluations are available. In these cases, we can conduct meta-analyses for the interventions before the HLPF presentation and finalize other effect-size calculations after the HLPF presentation to include in the consolidated report.

64. We will report two types of effect sizes. We will calculate the Hedges' *g* sample-size-corrected standardized mean differences (SMDs) for continuous outcome variables, which measure the effect size in units of standard deviation of the outcome variable. We will calculate odds ratios for binary outcome variables.

65. First, we will calculate SMD in Cohen's *d* effect sizes by dividing the mean difference with the pooled standard deviation by applying the formula in Equation 1:

$$(1) \text{ SMD} = \frac{Y_t - Y_c}{S_p}$$

66. Here SMD refers to the standardized mean differences, Y_t refers to the outcome for the treatment group, Y_c refers to the outcome for the comparison group, and S_p refers to the pooled standard deviation.

67. The pooled standard deviation S_p can be calculated by relying on the formulas in Equations 2 and 3:

$$(2) S_p = \sqrt{\frac{((SD_y)^2 * (nt + nc - 2)) - (\frac{\beta^2 * (nt * nc)}{nt + nc})}{nt + nc}}$$

$$(3) S_p = \sqrt{\frac{(nt - 1) * st^2 + (nc - 1) * sc^2}{nt + nc - 2}}$$

68. We will use Equation 2 for regression studies with a continuous dependent variable. In this equation, SD_y refers to the standard deviation for the point estimate from the regression, n_t refers to the sample size for the treatment group, n_c refers to the sample size for the control group, and β refers to the point estimate. We will use Equation 3 when there is information about the standard deviation for the treatment group and the control group separately.

69. We will correct the SMD for small sample size bias by relying on Equation 4, which transforms Cohen's *d* to Hedges' *g*:

$$(4) \text{ SMD}_{\text{corrected}} = \text{ SMD}_{\text{uncorrected}} * \left(1 - \frac{3}{4 * (nt + nc - 2) - 1}\right)$$

We will rely on Equation 5 to estimate the standard error of the SMD:

$$(5) \text{ SE} = \sqrt{\frac{nt + nc}{nc * nt} + \frac{\text{SMD}^2}{2 * (nc + nt)}}$$

70. Where possible, we will calculate odds ratios by relying on 2X2 contingency tables (Lipsey & Wilson, 2001; see Exhibit 12).

Exhibit 12. Estimation of Odds Ratios

Treatment or comparison group	Frequencies	
	Success	Failure
Treatment group	A	B
Comparison group	C	D

71. We will calculate the odds ratio using Equation 6, where \overline{ES} refers to the effect size:

$$(6) \quad \overline{ES} = \frac{ad}{bc}$$

72. In the cases in which we are not able to retrieve the missing data, we will extract or impute effect sizes and associated standard errors based on commonly reported statistics, such as the t or F statistic or p - or Z -values, using David B. Wilson's practical meta-analysis effect-size calculator. In studies that do not report sample sizes for the treatment and the control or comparison group, we will assume equal sample sizes across the groups.

73. While we will include cross-country regressions that use panel data, we will likely not be able to extract effect sizes from all of these studies. This is due to the non-binary nature of the treatment variable in these studies. For this reason, we will likely only synthesize cross-country regressions that use panel data in our narrative synthesis.

Risk of bias assessment

74. We will determine the rigor of the quantitative studies using an adaptation of a set of criteria, to assess risk of bias in experimental and quasi-experimental studies (Hombrados & Waddington, 2012). We will assess the risk of the following biases:

1. Selection bias and confounding, based on quality of identification strategy to determine causal effects and assessment of equivalence across the beneficiaries and nonbeneficiaries.
2. Performance bias, based on the extent of spillovers in comparison groups and contamination of the control or comparison group.

75. Assessing the risk of these biases will enable us to examine whether studies have a low, medium, or high risk of selection-bias or performance bias. Because of the

ambitious timeline, we decided to not include assessments of outcome and analysis reporting bias and other biases in our risk of bias assessment. We may also not be able to finalize the RoB assessment for all studies before the HLPF, but, at this moment, we still aim to do so.

Meta-analysis

76. We will pool the results of the quantitative studies that focus on the effects of trade, finance, technology, capacity building, and systemic issue interventions using meta-analysis for each combination of outcome measures and intervention type that includes three or more studies. We will conduct separate meta-analyses for the different outcome measures and separate meta-analyses by intervention. We will examine the heterogeneity of the effect sizes for each outcome across studies.

77. We will perform an extensive sensitivity analysis for two methodological effect size moderators after the presentations during the HLPF:

- Risk of bias status for each bias category.
- Study design (randomized controlled trials versus quasi-experimental studies).

78. We will start our analysis with separate meta-analyses of randomized controlled trials and quasi-experimental evaluations for determining the effects of interventions. Then we will use an iterative approach to determine the potential bias from pooling randomized controlled trials and quasi-experimental evaluations and studies with low, medium, and high risk of bias for each of the types of bias we assessed in our risk of bias assessment. We will use random-effects meta-analysis because the average effect of the interventions is likely to differ across contexts due to differences in programme design or contextual characteristics.

79. We will also investigate factors explaining heterogeneity by using inverse-variance weighted meta-regressions and stratified meta-analysis according to contextual and methodological moderator variables. We will use three contextual moderating variables: (1) type of intervention component, (2) geography, and (3) the inclusion of partnerships, capacity building, or systemic issues. We will, however, postpone analyses explaining heterogeneity until after the HLPF meeting.

Narrative synthesis

80. We will report the results following guidelines for systematic review without meta-analysis (SwIM) in cases where a combination of outcome measures and group type only results in two or fewer studies (Campbell et al., 2019). This will involve providing a rationale for grouping studies for the synthesis (based on group type and outcome measure as discussed above), describing the effect size, describing the synthesis methods, a description of the criteria used to prioritize results for the summary and synthesis, an investigation of the heterogeneity in the reporting of the effects, an analysis of the methods used to determine the certainty of evidence, and a description of the graphical and tabular methods to report the results.

Performance and Process Evaluation Synthesis

81. This section describes the approach to synthesizing performance and process evaluations, including quality review, mapping, coding, and thematic analysis.

Critical appraisal of performance and process evaluations

82. We will assess the methodological quality of all performance and process evaluations that meet the inclusion criteria using a qualitative review protocol. As a first assessment of quality and consistent with an evidence synthesis on SDG-6 (UNICEF, 2021), we will only include all the UN studies that scored satisfactory or higher on quality appraisal tools developed by UN agencies. Next, we will appraise the quality of all performance and process evaluations using a tool that includes indicators adapted from quality appraisals of evaluations across UN agencies, including UNICEF's GEROS (UNICEF, 2020), UNDP's evaluation quality assessment (UNDP IEO, 2021), UNEP's evaluation criteria and ratings (UNEP, n.d.), UNFPA's quality assurance and assessment tools (UNFPA, 2020), and WFP's evaluation quality assurance system (WFP, 2020).

83. Although we had initially proposed using the CASP protocol, the quality assessments used across UN agencies include most of the indicators from the CASP, but in a format that better aligns with UN evaluations. We will rate each item on a scale of High (mentioned and well explained), Medium (mentioned but missing at least one element), Low (alluded to but not described in full or explicitly), N/A, or Not mentioned to result in an overall methodological rating for each study. We will decide the cut-off score for inclusion of studies after the review of all studies is complete by calculating

the average assessment score to determine how well the study rated on the most critical items (i.e., design and methods) in the tool.

Qualitative evidence synthesis

84. We will conduct a rigorous evidence synthesis of the performance and process evaluations that meet the inclusion criteria and pass the critical appraisal. The purpose of this activity is to analyse and consolidate evidence from the included studies, documenting what works and what does not work for the successful implementation of interventions related to SDG-17.

85. **Data extraction.** The first step in the qualitative evidence synthesis is to extract data from the evaluations that passed the critical appraisal process. We will import all PDFs that meet the criteria for inclusion and pass the quality appraisal into NVivo. To extract data from the PDFs, we will focus on the sections on **findings, author's conclusions, and author's recommendations** (second-order data). Although we will focus on these sections of the studies, importing the full-text PDF studies will enable reviewers to understand the context of the full study as we code the process indicators and allow for identification of the characteristics that may have influenced the implementation of an intervention.

86. All members of the qualitative synthesis team will extract data from relevant evaluations. Initially, each team member will independently extract data on the same two evaluations. We will compare and address any inconsistencies in the types of data extracted for each category. Once consensus is achieved, each researcher will extract remaining data on a subset of studies. Quality assurance reviewers will conduct spot checks of the extracted data.

87. **Thematic analysis.** We will conduct a thematic analysis of the extracted data in NVivo to synthesize evidence from performance and process evaluations. The coding framework will build on a combination of deductive (top-down) and inductive (bottom-up) approaches. Using a deductive approach, we will develop several *a priori* themes informed by our conceptual framework, OECD-DAC criteria, and similar syntheses of process and performance evaluations (Johansson et al., 2022). An inductive approach will allow us to search for thematic patterns, emergent themes, and notable outliers in the data to identify the barriers and facilitators to successful

implementation. Exhibit 13 presents an indicative list of deductive codes followed by detail on the approach.

Exhibit 13. Draft Deductive Synthesis Framework

Thematic Category	Codes	Sub-codes
Design	Problem analysis Theory of change Relevance to population Planning Feasibility	Key outcomes of interest
Implementation of interventions related to: Finance; technology; trade; statistical capacity building; capacity building related to evidence generation and use; support for national plans to implement all the SDGs, including through North-South, South-South, and triangular regional and international cooperation and	Awareness Relevance to population Effectiveness Quality of delivery Sustainability Monitoring	Knowledge management, communication, public awareness Consideration and inclusion of local or disadvantaged groups Inclusion of gender considerations in design Stakeholder engagement Fidelity of implementation Coherence Efficiency Management Setting up conditions for sustainability and cultural transformation

<p>systemic issues policies and programmes related to policy coherence and coordination, the mobilization and sharing of knowledge, expertise, technology, and financial resources through multi-stakeholder partnerships, and country-owned results frameworks</p>		
<p>Moderators</p>	<p><i>Partnerships</i></p> <p>Institutional aspects</p> <p>External and contextual</p> <p>Sex of programme participants</p> <p>Gender mainstreaming approach</p>	<p>Ownership, alignment, harmonization, results, mutual accountability (Paris Declaration; Accra Agenda for Action); formal and informal</p> <p>Coordination</p> <p>Operational</p> <p>Economic</p> <p>Governance</p> <p>Socio political</p> <p>Culture</p> <p>Infrastructure</p>



	Inclusion of youth or other relevant age groups Other considerations related to age and gender	
--	---	--

OECD DAC Criteria: The deductive codes draw largely on the OECD–DAC criteria (Organisation for Economic Co-operation and Development, 2000) as defined in Exhibit 14.

Exhibit 14. OECD DAC Evaluation Criteria and Definitions

Evaluation Criteria	OECD Definition
Relevance	The extent to which the intervention objectives and design respond to beneficiaries’ global, country, and partner/institution needs, policies, and priorities, and continue to do so if circumstances change
Coherence	The compatibility of the intervention with other interventions in a country, sector, or institution
Effectiveness	The extent to which the intervention achieved, or is expected to achieve, its objectives, and its results, including any differential results across groups
Efficiency	The extent to which the intervention delivers, or is likely to deliver, results in an economic and timely way
Impact	The extent to which the intervention has generated, or is expected to generate, significant positive or negative, intended or unintended, higher-level effects
Sustainability	The extent to which the net benefits of the intervention continue or are likely to continue

Note. Criteria are defined by the OECD: <https://www.oecd.org/dac/evaluation/revised-evaluation-criteria-dec-2019.pdf>

88. OECD Partnerships Criteria: Where applicable, we will assess efforts to establish and build partnerships in trade, finance, and technology using inductive codes drawn from the Paris Declaration and revised under the Accra Agenda for Action: ownership,

alignment, harmonization, results, mutual accountability (OECD, 2008). In some cases, evaluations include a section on partnerships that specifies progress on these indicators. For those that do not specify efforts towards partnerships, we will code findings under categories, if possible. Finally, we will also inductively assess differences across types of partnerships where there is information available to do so.

89. Capacity Measurement Framework: We will assess the efforts related to systemic issues and capacity building as cross-cutting issues, as well as the discrete efforts related to technology, trade, and finance. To synthesize capacity building efforts, we will draw on indicators from the frameworks of various UN agencies, including, for example, the WFP country capacity strengthening glossary (WFP, 2019) and the UNDP capacity measurement framework (UNDP, 2010). As discussed during the meetings with the management group, we will refine our approach to synthesizing the evidence on capacity development with the intention of using it as a model for synthesizing capacity development evidence under the remaining SDGs.

90. The remaining codes incorporate other potential moderators of success of an intervention, including institutional coordination and operations, as well as external and contextual factors that may affect delivery.

Synthesis of Systematic Reviews

91. We will only include systematic reviews and other evidence synthesis that focus on the impact of SDG-17 programmes, policies, and interventions on other SDG outcomes (not on SDG-17 outcomes). We will synthesize relevant intervention and outcome pairings using a results matrix in which we include the intervention in a row and the outcome in a column. We will then examine the synthesis method of the systematic review or evidence synthesis and determine whether the intervention had statistically significant effects on the outcome and, if so, the magnitude of the effect.

92. The synthesis of systematic reviews will serve to generate lessons about the potential effects of SDG-17 interventions on other SDG outcomes. This information will help in generating recommendations on how to accelerate progress toward SDG-17 that could facilitate improvements in other SDG outcomes. It is likely that we will only finalize the evidence synthesis of systematic reviews after the HLPF meeting.

Development of Evidence-Gap Maps

93. Based on the results of the critical appraisal and the answers to the synthesis questions, we will develop evidence-gap maps to highlight areas and themes around SDG-17 with a limited evidence-base that require further strengthening in future evaluations. We will create the evidence-gap map using the same principles and methods as used for the four types of syntheses, meaning that we will align the scope and framework with each synthesis.

94. We will follow guidelines for 3ie evidence-gap maps adapted to this project (Snilstveit et al., 2017). First, we will create a matrix on each axis using the intervention and outcome concepts from the PICO criteria. We will base the intervention and outcome concepts on the results of the evidence syntheses. Second, using the information from the critical appraisal, we will map each of the included reports and evaluations on the intersections of intervention-outcome combinations. If a study has multiple outcomes, we will include the results in multiple intersections. We will use the findings of the critical appraisal and risk of bias assessment to assign a rating to each of the studies using a stoplight system (i.e., green for high-quality studies, orange for studies with some quality concern, red for low-quality studies; see Exhibit 13 for an example). The number and colours of the bubbles in the evidence-gap map will indicate the availability and quality of the existing evidence.

95. Based on the map, we will identify (1) absolute gaps, in which there are no studies on intersection between a given intervention and outcome, and (2) synthesis gaps for intersections with no or few high-quality studies.

96. We will link the evidence-gap map to the theory of change to identify clear evidence gaps on SDG-17. We will recommend that UN organizations and other stakeholders focus future research efforts on addressing these evidence gaps.

Exhibit 13. Illustrative Example of Evidence-Gap Map

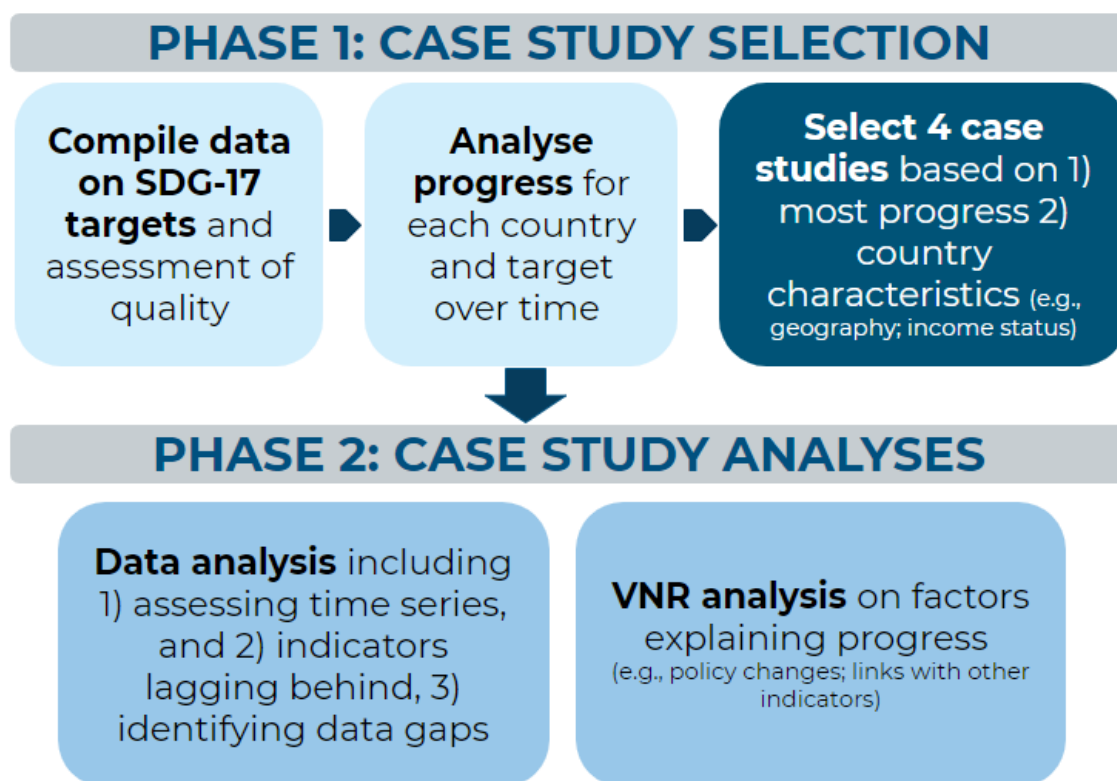
		Outcomes			
		Growth	Poverty	Employment	Labour conditions
Interventions	Tax reform	● ● ●		●	●
	Tax	● ●		● ●	● ●
	Export	●			● ● ●
	Mhealth		●	● ●	● ●
	Self-help group		● ●	●	

Positive Deviance Assessment

97. As a third objective of the synthesis study, we will conduct a positive deviance assessment of the current performance on SDG-17. The objective of this component is to identify countries from varying regions and variation in income and human development status with considerable progress towards achieving targets within SDG-17 and to highlight factors contributing to this progress using a case study approach that includes comparisons with other countries in the same region as the selected countries. Simultaneously, we will use this component to determine which targets of SDG-17 lag in terms of progress and/or data availability. The positive deviance analysis will contribute to the understanding of what circumstances and factors are associated with progress, as well as give an indication of potential barriers.

98. Methodology: The deviance assessment consists of two phases, namely, case study selection and case study analysis (see Exhibit 14). As part of the deviance analysis, we will incorporate the assessment of Voluntary national reviews (VNRs) to help explain factors associated with progress in SDG-17 indicators.

Exhibit 14. Phases of the Positive and Negative Deviance Analysis and VNR Assessment



99. Phase 1: Case Study Selection: for the selection of the case studies, we will assess the latest data available on the 19 targets (24 indicators) that are identified under SDG-17.

100. Data Analysis: We will use the data uploaded on the SDG tracker¹⁰ to determine the progress over time. The SDG tracker is frequently updated, and we retrieved the latest data from the original sources in June and July 2022. Each of the indicators comes from a reliable, original source that has data by country over multiple years. During the inception phase, we had consultations with various custodian agencies of SDG-17 indicators, in which we further discussed the availability and quality of the data, and any potential issues with definitions, measurements, etc. (see *Annex E for an overview of the data, timespan available, original database and source*).

¹⁰ <https://sdg-tracker.org/global-partnerships>

101. Based on the available data, we plan to create measures on the relative progress on each indicator (i.e., the percentage change in the indicator with regards to the baseline value). Depending on the timespan for which data is available, we will assess the progress over the last five or 10 years. In addition to progress measures, we will examine the overall quality of the available data. These quality assessments may include analyses of the number of years for which each country has data available, the latest and earliest year of the available data, and outlier values.

102. Using the available data, we will then identify (1) the countries that made the most progress on SDG-17 indicators potentially categorized by sub-topic (i.e., trade, finance, technology, capacity building, systemic issues); (2) the countries that made the most progress overall on SDG-17; and (3) the indicators that have issues with data availability or reliability.

103. VNR Assessment: VNRs are country-led and country-driven reviews of progress on the 2030 Agenda for Sustainable Development. The reports focus on the experience with the SDGs, including successes, challenges, and lessons learned, and are therefore an appropriate source to complement the data analysis in the deviance assessment. By 2021, 176 countries had submitted either one, two, or three VNRs. For countries with multiple reports, we will use the report content to analyse strategy or policy shifts, achievements at specific times, and changing relationships between SDG-17 and other goals. In 2020, the UN Committee for Development Policy (CDP) conducted an assessment using text analysis of 45 VNRs and identified 42 VNRs, which reported on issues addressed in SDG-17 (UN CDP, 2020). The assessment indicated that gaps remained in the coverage of issues central to SDG-17. While multi-stakeholder partnerships, fiscal policy ODA, and aid were mentioned in almost all of the reports, technology facilitation and measures of progress beyond GDP were only covered in two or three reports.

104. For the deviance analysis, we will conduct a new exploratory text analysis of the currently available VNRs with the aim of identifying the key themes in the most recent VNRs as well as any gaps in coverage. For this analysis, we will explore data-driven text analysis methods in the field of text mining, such as topic modelling and sentiment analysis, which can help us identify areas that countries consider important enough

to consistently touch upon in their reports as well as what the distribution of “sentiment” is in these reports (Silge and Robinson, 2017). The analyses will serve to identify potential trends in the case study countries that are associated with their “above-average” performance on certain SDG indicators or reasons for these indicators underperforming relative to others.

105. Selection Process: To select case study countries, we will use a stratified ranking of countries’ progress on key SDG-17 indicators based on the preliminary data analysis as well as the coverage of the indicator themes (i.e., the frequency with which key themes are mentioned) in the VNR reports. Because of the ambitious timeline, we will not select countries with sparse SDG indicator and VNR data. Instead, we will select a sample based on salient characteristics (i.e., population size, geography, income status) to ensure diversity and variability in the cases. In this way, we will ensure that we will include low-income, middle-income, and high-income countries from different regions. Ultimately, we will select four case study countries for which we will conduct a more in-depth analysis on “why” and “how” progress related to SDG-17 occurred. We will identify why these countries made more progress on SDG-17 by comparing the countries with other countries in the region. In this way, we will include additional variation, providing additional reassurance that we will include low-income, middle-income, and high-income countries from different regions in the deviance analysis.

106. Phase 2: Case Study Analysis: The case analysis phase will consist of three components: secondary data analysis, VNR assessment, and document review. Combining these elements will provide a comprehensive assessment of factors contributing to progress on SDG-17 indicators.

107. Data Analysis: The data analysis will build on the preliminary insights generated during phase 1 of the case study selection. Within the assessment, we will carry out a more in-depth analysis for each case study country by analysing time series data of case study countries and comparison countries in the region to detect systematic trends or patterns over time. We will attempt to link any emerging trends with events and/or circumstances that are described in the VNR. For instance, if the data indicates a large proportional change in the progress on a given indicator, we will assess

whether there have been major policy changes, investments, programmatic efforts, etc. reported within that same time period.

108. VNR Assessment: Within the case study analysis phase, we will make use of the results of the preliminary VNR analysis. We will use the text analysis results that have identified the coverage of key SDG-17 issues to guide further document review. For the selected case study countries and comparison countries in the region, we will analyse the sections in the report in which key issues are mentioned. At this stage, we will focus more on the sentiment attached to the key issues as well as mentions of policy mechanisms, barriers, and enabling environments that are associated with the key aspects of SDG-17.

109. Document Review: We will use document review to identify recent events, policy changes, or implementation that could have affected the progress on key issues reported in SDG-17. Specifically, we will review (1) relevant policy documents related to trade, finance, technology, capacity building, and systemic issues; and (2) existing programme reports or thematic studies on the SDG-17 sub-topics. We will use the document review to complement the VNR analysis and (if available) to address evidence gaps on explanatory factors for the progress on SDG-17.

Limitations

110. While the evidence synthesis will generate important lessons about how to accelerate progress toward SDG-17, the ambitious timeline creates some limitations. First, the evidence synthesis team will not be able to conduct a full systematic review in the time required to generate preliminary lessons before the HLPF. We will only conduct searches in a limited number of databases, which will limit the comprehensiveness of the review to some extent. We will address this limitation by limiting the scope as discussed in previous sections and spending additional time after the HLPF forum to conduct additional analyses related to SDG-17. Second, the scoping phase suggested that it is unlikely that we will find many rigorous impact evaluations related to SDG-17, and even fewer impact evaluations that include a cost-effectiveness analysis. As a result, any recommendations on how to accelerate progress towards SDG-17 will be based on only a limited evidence-base. To address this limitation, we will create an evidence-gap map to come with recommendations

for future research. Third, we will primarily focus on trade, technology, and finance, and only to a limited extent on capacity building and systemic issues. We will create a framework for analysing capacity building and systemic issues that future evidence syntheses can adapt to address this limitation.

Work Plan

111. Exhibit 15 presents our proposed timeline for conducting project activities and producing seven primary project deliverables: the methodological protocol, the report about the synthesis of the impact evaluation, the impact evaluation and systematic review synthesis report, the UN performance and process evaluation synthesis report, the VNR analysis and positive deviance analysis, the presentation for the HLPF, and the consolidated synthesis report. We will produce a preliminary report with progress and findings to date by May 2023. This report will likely not yet include findings of the synthesis of impact, performance, and process evaluations other than a general overview of the included studies and their characteristics.

112. During the inception phase, which occurred between December 2022 and March 2023, we devised and refined the methodological protocol and scope of the synthesis reports, incorporating guidance and feedback from UNDP, the Coalition, and the management group. With the draft methodological protocol complete, we will concurrently work on the synthesis of impact evaluations and systematic reviews, the synthesis of UN performance and process evaluations, and the analysis of VNR assessment data, including the positive deviance assessment, from March to June. These analyses will result in a PowerPoint presentation to the HLPF. Different team members will prioritize different reports with oversight from the team leader, Dr. Thomas de Hoop. We plan to finalize all deliverables by September 2023.

Exhibit 15. Work Plan

#	Deliverable	Expected date of completion
1	Draft Methodological Protocol	28 February 2023
2	Final Methodological Protocol	31 March 2023
3	Draft preliminary findings triangulating the evidence reviewed to date from the VNR Analysis, Positive and Negative Deviance Assessment, Synthesis of Impact Evaluations and Systematic Reviews, and Synthesis of UN Performance and Process Evaluations	15 May 2023
4	Draft PowerPoint Presentations to the HLPF with Consolidated Preliminary Findings and lessons of evidence triangulated to date	15 June 2023
5	Final PowerPoint Presentations to the HLPF with Consolidated Findings and Lessons of evidence triangulated to date	30 June 2023
6	Draft Consolidated Synthesis Report of All Evaluative Evidence and Gap Maps	21 August 2023
7	Final Consolidated Synthesis Report of All Evaluative Evidence and Gap Maps	4 September 2023
8	Draft PowerPoint presentation to the SDG Summit based on all Evaluative Evidence	5 September 2023
9	Final PowerPoint presentation to the SDG Summit based on all Evaluative Evidence	15 September 2023

Quality Assurance

113. We will submit each deliverable after quality assurance by Hannah Ring. She currently leads an evidence synthesis on process evaluations of food security interventions and is intimately familiar with the use of the OECD-DAC criteria for evidence synthesis. In addition, we will pilot each screening, coding, and quality appraisal tool with two reviewers to ensure a consistent approach.

Coordination with Management Group and Technical Advisory Group

114. We will also closely coordinate with a management group and a technical advisory group. The management group will serve to discuss the scope with the evidence synthesis team and will review each deliverable. It consists of representatives of

various UN organizations and Member States who each take a strong interest in the SDG evidence synthesis coalition. The technical advisory group will provide methodological and content expertise to the evidence synthesis team and conduct quality assurance of the deliverables. They will also support with the development of recommendations following the evidence synthesis. At this moment, the technical advisory group has not been formed yet. However, the evidence synthesis team will coordinate with several consultants with expertise in SDG-17 to ensure the relevance of the search string.

References

- Aiyar, P. S., Chen, J., Ebeke, C., Gudmundsson, T., Ilyina, A., Kangur, A., Kunaratskul, T., Rodriguez, S., Ruta, M., Schulze, T., Soderberg, G., & Trevino, J. P. (n.d.). Geoeconomic Fragmentation and the Future of Multilateralism. Staff Discussion Notes.
- Alfaro-Serrano, D., Balantrapu, T., Chaurey, R., Goicoechea, A., & Verhoogen, E. (2021). Interventions to promote technology adoption in firms: A systematic review. *Campbell Systematic Reviews*.
- Amare, A. T., Toni, A. T., Mekonnen, Z. A., Endehabtu, B. F., & Tilahun, B. C. (2021). Effectiveness and Feasibility of Using Local Medical Universities for Capacity Building to Improve the Immunization Program Ethiopia: Quasi-Experimental Study. *Journal of Multidisciplinary Healthcare*, Volume 14, 9–19.
<https://doi.org/10.2147/JMDH.S285280>
- Ashraf, N., Giné, X., & Karlan, D. (2009). Finding Missing Markets (and a disturbing epilogue): Evidence from an export crop adoption and marketing intervention in Kenya. *American Journal of Agricultural Economics*, 91(4): pp. 973–990.
- Better Evaluation. Evaluating Capacity Development.
<https://www.betterevaluation.org/methods-approaches/themes/evaluating-capacity-development>
- Basstel. (2022). Final evaluation of The Netherlands Trust Fund (NTF) Phase IV – Export Sector Competitiveness Programme. International Trade Centre.
- Brody, C., de Hoop, T., Vojtkova, M., Warnock, R., Dunbar, M., Murthy, P., & Dworkin, S.L. (2015). Economic self-help group programmes for improving women’s empowerment? A systematic review. *Campbell Systematic Reviews*, 11(1), 1–182.

- Campbell, M., McKenzie, J., Sowden, A., Vittal Katikireddi, S., Brennan, S. E., Ellis, S., Hartmann-Boyce, J., Ryan, R., Shepperd, S., Thomas, J., Welch, V., & Thomson, H. (2020). Synthesis without meta-analysis (SwIM) in systematic reviews: Reporting guideline. *BMJ*, (368).
- Cirera, X., & Lakshman, R. W. D. (2017). The impact of export processing zones on employment, wages and labour conditions in developing countries: Systematic review. *Journal of Development Effectiveness*, 9(3), 344–360.
- Cirera, X. W. & Dirk Lakshman, R. W. D. (2011). What is the evidence of the impact of tariff reductions on employment and fiscal revenue in developing countries?: A systematic review. Social Science Research Unit.
- Clarke, A., & Crane, A. (2018). Cross-Sector Partnerships for Systemic Change: Systematized Literature Review and Agenda for Further Research. *Journal of Business Ethics*, 150(2), 303–313. <https://doi.org/10.1007/s10551-018-3922-2>
- Climate Investment Funds. (2022). Enablers the role of enabling environment in scaling up climate finance. Climate Investment Funds.
- Das, J., Do, Q-T., Shaines, K., & Srikant, S. (2013). U.S. and them: The geography of academic research. *Journal of Development Economics* 105, 112–130.
- Dollar, D. & Kraay, A. (2002) Growth is Good for the Poor. *Journal of Economic Growth* 7, 195–225.
- Garcia, O.A. (2022). AIDA | Origins: How the journey into artificial intelligence for development began. <http://web.undp.org/evaluation/media-centre/blogs/aida.shtml>
- Grames, E. M., Stillman, A. N., Tingley, M. W., & Elphick, C. S. (2019). An automated approach to identifying search terms for systematic reviews using keyword co-occurrence networks. *Methods in Ecology and Evolution*, 10(10), 1645–1654.
- Guarini, E., Mori, E., & Zuffada, E. (2022). Localizing the Sustainable Development Goals: A managerial perspective. *Journal of Public Budgeting, Accounting & Financial Management*, 34(5), 583–601. <https://doi.org/10.1108/JPBAFM-02-2021-0031>

- Hirvonen, K. (2020). This is US: Geography of Evidence in Top Health Economics Journals. *Health Economics*, 29(10), 1316–1323.
- Hombrados, J., & Waddington, H. (2012). Internal validity in social experiments and quasi-experiments: An assessment tool for reviewers. Mimeo. 3ie.
- ICF. (2022). Evaluation of the Scaling up Renewable Energy Program in Low-income Countries: Evaluation Report.
- Independent Evaluation Group. (2020). The International Finance Corporation’s Blended Finance Operations: Findings from a Cluster of Project Performance Assessment Reports. World Bank, Washington, DC.
<https://doi.org/10.1596/33202>
- Independent Evaluation Group. (2020). The International Finance Corporation’s Blended Finance Operations: Findings from a Cluster of Project Performance Assessment Reports. World Bank, Washington, DC.
<https://doi.org/10.1596/33202>
- International Monetary Fund. Strategy, Policy, & Review Department. (2018). Guidance Note on the Bank-Fund Debt Sustainability Framework for Low Income Countries. *Policy Papers*, 18(009), 1.
<https://doi.org/10.5089/9781498307260.007>
- International Monetary Fund & World Bank. (2021). World Bank Group And International Monetary Fund Support For Debt Relief Under The Common Framework And Beyond. *Policy Papers*, 2021(022), 1.
<https://doi.org/10.5089/9781513576039.007>
- International Telecommunication Union. (2022a). Economic and fiscal incentives to accelerate digital transformation—9th ITU Economic Experts Roundtable: Outcome report.
- International Telecommunication Union. (2022b). Global Connectivity Report 2022.
- International Telecommunication Union. (2022c). ICT accessibility overview: Assessment for the Africa region.

International Trade Centre. (2022a). Evaluation of the ITC SheTrades Initiative. International Trade Centre.

International Trade Centre. (2022b). Midterm Evaluation of the ITC MENATEX Programme. International Trade Centre.

International Trade Centre. (2021a). Evaluation of ITC's Performance in Trade and Market Information.

International Trade Centre. (2021b). Sustainability review of the project 'Improving the competitiveness of Senegalese mangos' Results and lessons learned three years after project completion. International Trade Centre.

International Trade Centre. (2021c). Sustainability review of the project 'Integration of horticulture supply/value chains into tourism, Tanzania' Results and lessons learned three years after project completion. International Trade Centre.

International Trade Centre. (2019a). Evaluation of the ITC Participation and Performance in the UN Delivering as One System. International Trade Centre.

International Trade Centre. (2019b). Evaluation of the ITC Trade Facilitation Programme. International Trade Centre.

International Trade Center. (2018a). Evaluation of the Certified Trade Advisers Programme (CTAP). International Trade Centre.

International Trade Centre. (2018b). Evaluation of the Non-tariff Measures Programme. International Trade Centre.

International Trade Centre. (2017a). Final Evaluation Promoting Intra-Regional Trade in Eastern Africa Kenya Tanzania Zambia. International Trade Centre.

International Trade Centre. (2017b). Trade for Sustainable Development (T4SD) Project (2013 – 2016). International Trade Centre

International Trade Centre. (2016a). Evaluation of the Women and Trade Programme – Phase I. International Trade Centre.

International Trade Centre. (2016b). Final Evaluation of Sector Competitiveness & Export Diversification in The Gambia. International Trade Centre.

International Trade Centre. (2015). Evaluation of the Trade, Climate Change and Environment Programme. International Trade Centre.

Jarreau, J. & Poncet, S. (2012). Export sophistication and economic growth: Evidence from China, *Journal of Development Economics*, 97 (2), 281–292.

Keynejad, R., Semrau, M., Toynbee, M., Evans-Lacko, S., Lund, C., Gureje, O., Ndyabangi, S., Courtin, E., Abdulmalik, J. O., Alem, A., Fekadu, A., Thornicroft, G., & Hanlon, C. (2016). Building the capacity of policy makers and planners to strengthen mental health systems in low- and middle-income countries: A systematic review. *BMC*.

Khan, A. Q., & Khwaja, A. I. (2018). Tax farming redux: Experimental evidence on performance pay for tax collectors. *Quarterly Journal of Economics*, 131(1), 219–271.

KPMG & FACTS. (2016). Midterm Evaluation of Pashmina Enhancement and Trade Support (PETS) project, Nepal. KPMG & FACTS.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA, United States: SAGE.

Mbaye, S., Moreno Badia, M., & Chae, K. (2018). Global Debt Database: Methodology and Sources. *IMF Working Papers*, 18(111).
<https://doi.org/10.5089/9781484353592.001>

McCorrison, S., Hemming, D., Godwin, J., Osborn, J., Parr, M., & Roberts, A. (2013). What is the evidence of the impact of agricultural trade liberalisation on food security in developing countries? A systematic review.
<https://www.researchgate.net/publication/237973599>

- Nakhimovsky, S. S., Feigl, A. B., Avila, C., O’Sullivan, G., Macgregor-Skinner, E., & Spranca, M. (2016). Taxes on Sugar-Sweetened Beverages to Reduce Overweight and Obesity in Middle-Income Countries: A Systematic Review. PLOS ONE, 11(9), e0163358. <https://doi.org/10.1371/journal.pone.0163358>
- Nkala, D., & Kim, Y. (2022). The United Nations’ Role in the Evolution of Technical Cooperation among Developing Countries, South-South Cooperation, and Triangular Cooperation: Real or Cosmetic Approaches to Development Cooperation? Asian Journal of Peacebuilding, 10(1), 45–59. <https://doi.org/10.18588/202205.00a264>
- O’Brien, A., & Ulhoa, L. (2021). The World Bank Group’s Early Support to Addressing the COVID-19 Economic Response April 2020–June 2021 An Early-Stage Evaluation. The World Bank.
- O’Brien, A., Ulhoa, L., & Siegel, C. (2022). The International Finance Corporation’s and Multilateral Investment Guarantee Agency’s Support for Private Investment in Fragile and Conflict-Affected Situations, Fiscal Years 2010–21 An Independent Evaluation.
- OECD. (2022). Revenue Statistics 2022: The Impact of COVID-19 on OECD Tax Revenues. OECD Publishing. <https://doi.org/10.1787/8a691b03-en>
- OECD/UNCDF. (2020). Blended Finance in the Least Developed Countries 2020: Supporting a Resilient COVID-19 Recovery. OECD Publishing. <https://doi.org/10.1787/57620d04-en>
- OECD. (2019). Better Criteria for Better Evaluation: Revised Evaluation Criteria – Definitions and Principles for Use. OECD/DAC Network on Development Education. <https://www.oecd.org/dac/evaluation/revised-evaluation-criteria-dec-2019.pdf>
- OECD. (2008). The Accra Agenda for Action. OECD/DAC Network on Development Education. <https://www.oecd.org/dac/effectiveness/parisdeclarationandaccraagendaforaction.htm>

- OECD. (n.d.-b). OECD Toolkit for territorial approach to the SDGs.
- OECD. (n.d.-a). Achieving SDG Results in Development Co-operation: Summary for Policy Makers.
- Paris 21! (2022a). Digital Transformation of National Statistical Offices. OECD Publishing. <https://www.paris21.org/Digital-transformation-NSO>
- Paris 21! (2022b). The PARIS21 Partner Report on Support to Statistics 2022: A Wake-Up Call to Finance Better Data. OECD Publishing. https://www.oecd-ilibrary.org/development/the-paris21-partner-report-on-support-to-statistics-2022_c3cfb353-en
- Paris 21! (2021). 2021 Progress Report: Partnering for Better Data For Better Lives for All. Paris21!
- Popp, D. (2020). Promoting clean energy innovation at the state and local level. Agricultural and resource economics review
- Rodriguez-Crespo, E., Billon, M., & Marco, R. (2021). Impacts of Internet Use on Trade: New Evidence for Developed and Developing Countries. *Emerging Markets Finance and Trade*, 57(10), 3017–3032. <https://doi.org/10.1080/1540496X.2019.1676225>
- Schnatz, J., A. Guerrero-Ruiz and K. Sachin. (2022). Aligning development co-operation to the SDGs in lower middle-income countries: A case study of Bangladesh. OECD Development Co-operation Working Papers, No. 105, OECD Publishing, Paris, <https://doi.org/10.1787/7c83ab20-en>
- Silge, J., & Robinson, D. (2017). Text mining with R: A tidy approach. O'Reilly Media, Inc.
- Sobir, R. (2019). UNDESA – Review of capacity development gaps, needs, and priorities.
- Stacey, B. (2022). Citizen-Generated Data for SDG Reporting in Kenya. Paris21!

Umeh, C. A. (2016). Evaluating the effectiveness of online capacity building resources on capacity improvement of local Nigerian NGOs serving orphans and vulnerable children.

UN. (2022). The Sustainable Development Goals Report. United Nations.
<https://unstats.un.org/sdgs/report/2022/The-Sustainable-Development-Goals-Report-2022.pdf>

UN DESA. (n.d.-a). How should civil society stakeholders report their contribution to the implementation of the 2030 Agenda for Sustainable Development?

UN DESA. (n.d.-b). Micro-, Small and Medium-sized Enterprises (MSMEs) and their role in achieving the Sustainable Development Goals.

UNDP. (2021). High-level forum of directors general for development cooperation 2021: Trends, Opportunities and Challenges on the Path to Sustainable Development: Strengthening South-South Cooperation and Triangular Cooperation in the COVID-19 Context. UNDP.

UNDP. (2010). Measuring Capacity.
https://www.betterevaluation.org/sites/default/files/2022-08/UNDP__Measuring_Capacity___July_2010__pdf_.pdf

UNDP Independent Evaluation Office. (2021). Evaluation Quality Assessment.
<http://web.undp.org/evaluation/guideline/documents/PDF/section-6.pdf>

UNEP (n.d.) Evaluation Criteria and Ratings. <https://www.unep.org/evaluation-office/our-evaluation-approach/evaluation-criteria-and-ratings>

UNEP. (n.d.). Methodology for SDG-indicator 17.14.1: Mechanisms in place to enhance policy coherence for sustainable development.

UNFPA. (2020). Quality Assurance and Assessment Tools and Guidance.
<https://www.unfpa.org/admin-resource/evaluation-quality-assurance-and-assessment-tools-and-guidance>

UNICEF. (2021). Evaluation Synthesis of United Nations System and Development Bank Work Towards SDG6. <https://gdc.unicef.org/resource/evaluation-synthesis-united-nations-system-and-development-bank-work-towards-sdg-6>

UNICEF. (2020) Global Evaluation Report Oversight System Guidance Handbook. https://www.unicef.org/evaluation/media/2571/file/2020_GEROS_Handbook.pdf

UN Office for the South-South Cooperation. (2021). The Importance of South-South Cooperation in Strengthening Global South Trade, Investments and Regional Integration: A Contextual Overview. UNDP.

UNOSSC/UNDP. (2022). UN Fund for South-South Cooperation: Results report 2020-2021 results achieved with our partners.

UNOSSC/UNDP. (2021). South-South Ideas: Mapping South-South Cooperation in Science, Technology and Innovation for Theory and Practice. UNDP.

UN Water. (2017). Roles and responsibilities SDG monitoring and reporting. <https://www.unwater.org/news/roles-and-responsibilities-sdg-monitoring-and-reporting>

Waddington, H., White, H., Snilstveit, B., Hombrados, J.G., Vojtkova, M., Davies, P., Bhavsar, A., Evers, J., Perez Koehlmoos, T., Petticrew, M., Valentine, J.C., & Tugwell, P. (2012). How to do a good systematic review of effects in international development: A toolkit. *Journal of Development Effectiveness*, 4 (3): 359-387.

World Bank. (2022). Survey on the Implementation of the Cape Town Global Action Plan for Sustainable Development Data.

World Bank. (2021). World Bank Group Approaches to Mobilize Private Capital for Development: An Independent Evaluation. World Bank. <https://doi.org/10.1596/35040>

World Bank. (2019). Leveraging economic migration for development: A Briefing for the World Bank Board. World Bank.

World Bank. (2017). World Bank Group Joint Projects: A Review of Two Decades of Experience Lessons and Implications from Evaluation. World Bank.

World Bank. (2016). Findings from Evaluations of Policy-Based Guarantees. World Bank.

World Food Programme. (2020). Evaluation Quality Assurance System.
<https://www.wfp.org/publications/eqas-evaluation-quality-assurance-system-0>

World Food Programme. (2019). Country Capacity Strengthening Glossary

Yao, S. (2006) On economic growth, FDI and exports in China, Applied Economics, 38 (3), 339-351.

Annex A. Search Terms

115. Original PICO search strings to identify impact studies:

- **Population:** *low* income countr*, OR middle* income countr*, OR developing countr*, OR less* developed countr*, OR underdeveloped countr*, OR under developed countr*, OR underserved countr*, OR LMIC*, OR low GDP, OR low GNP, OR fragile state, OR third world, OR transitional countr*, OR high* burden countr*, OR Asia, OR South Asia*, OR Africa, OR Latin America, OR South America, OR Central America, OR Middle East, OR sub-Saharan Africa, OR sub Saharan Africa, OR Caribbean, OR West Indies, OR Afghanistan*, OR Albania, OR Algeria, OR American Samoa, OR Angola*, OR Argentina, OR Armenia*, Or Azerbaijan, OR Bangladesh*, OR Belarus, OR Belize, OR Benin, OR Bhutan*, OR Bolivia*, OR Bosnia, OR Botswana, OR Brazil, OR Bulgaria, OR Burkina Faso, OR Burundi, OR Cabo Verde, OR Cameroon, OR Cambodia*, OR Central African Republic*, OR Chad, OR China, OR Colombia, OR Comoros, OR Cote d'Ivoire, OR Ivory Coast, OR Congo*, OR Costa Rica, OR Cuba, OR Democratic Republic of Congo, OR Democratic People's Republic of Korea, OR Djibouti, OR Dominica, OR Dominican Republic, OR Ecuador, OR Egypt*, OR El Salvador, OR Eritrea, OR Eswatini, OR Ethiopia*, OR Equatorial Guinea, OR Fiji, OR Gabon, OR Gambia*, OR Gaza, OR Georgia, OR Ghana*, OR Grenada, OR Guatemala*, OR Guam, OR Guinea*,OR Guyana, OR Haiti*, OR Hondura*, OR India*, OR Indonesia*, OR Iran, OR Iraq, OR Jamaica, OR Jordan, OR Kazakhstan, OR Kenya*, OR Kiribati, OR Korea*, OR Kosovo, OR Kyrgyz*, OR Lao*, OR PDR, OR Lebanon, OR Lesotho, OR Liberia*, OR Libya, OR Madagascar, OR Malawi*, OR Malaysia, OR Maldives, OR Mali, OR Marshall Islands, OR Mauritius, OR Mauritania*, OR Mexico, OR Micronesia, OR Moldova*, OR Mongolia*, OR Montenegro, OR Morocc*, OR Mozambique, OR Burma, OR Myanmar, OR Myanma, OR Namibia, OR Nepal*, OR Nicaragua, OR Niger, OR Nigeria*, OR North Macedonia, OR Palau, OR Pakistan*, OR Paraguay, OR Peru, OR Philippines, OR Philipines, OR Phillipines, OR Phillipines, OR Papua New Guinea, OR Republic of Congo, OR Rwanda, OR Russian Federation, OR Samoa, OR Sao Tome and Principe, OR Senegal*, or Serbia, OR Sierra Leone, OR Sri Lanka, OR Solomon Islands, OR Somalia*, OR South Africa*, OR South Sudan, OR Sudan, OR St. Lucia, OR St. Vincent, OR Swaziland, OR Suriname, OR Syria*, OR Tajikistan, OR Tanzania*, OR Thailand, OR Timor-Leste, OR Tokelau, OR Togo, OR Tonga, OR Tunisia, OR Turkey, OR Turkmenistan, OR Tuvalu, OR Uganda*, OR Ukraine, OR Uzbekistan, OR Vanuatu, OR Vietnam*, OR Viet Nam, OR West Bank, OR Yemen, OR Zambia*, OR Zimbabwe*

- **Intervention:** *tax, OR trade, OR export, OR tariff, OR technolog*, OR digital technolog*, OR internet, OR phone, OR mHealth, OR mobile health, OR mobile money, OR mobile banking, OR finance, OR savings, OR self-help groups, OR market-led, OR market-based reforms, OR corporate social responsibility, OR certification schemes, OR organic, OR blockchain, OR laptop, OR business support services, OR public-private partnerships*
- **Comparison:** *evaluation, OR impact evaluation, OR random* controlled trial, OR experiment, OR quasi-experiment, OR regression discontinuity, OR difference-in-differences, OR assessment, OR propensity score, OR systematic review, OR rapid review, OR evidence synthesis*

116. Original PIE search strings to identify process and performance evaluations:

- **Population:** *low* income countr*, OR middle* income countr*, OR developing countr*, OR less* developed countr*, OR underdeveloped countr*, OR under developed countr*, OR underserved countr*, OR LMIC*, OR low GDP, OR low GNP, OR fragile state, OR third world, OR transitional countr*, OR high* burden countr*, OR Asia, OR South Asia*, OR Africa, OR Latin America, OR South America, OR Central America, OR Middle East, OR Sub-Saharan Africa, OR Sub Saharan Africa, OR Caribbean, OR West Indies, OR Afghanistan*, OR Albania, OR Algeria, OR American Samoa, OR Angola*, OR Argentina, OR Armenia*, Or Azerbaijan, OR Bangladesh*, OR Belarus, OR Belize, OR Benin, OR Bhutan*, OR Bolivia*, OR Bosnia, OR Botswana, OR Brazil, OR Bulgaria, OR Burkina Faso, OR Burundi, OR Cabo Verde, OR Cameroon, OR Cambodia*, OR Central African Republic*, OR Chad, OR China, OR Colombia, OR Comoros, OR Cote d'Ivoire, OR Ivory Coast, OR Congo*, OR Costa Rica, OR Cuba, OR Democratic Republic of Congo, OR Democratic People's Republic of Korea, OR Djibouti, OR Dominica, OR Dominican Republic, OR Ecuador, OR Egypt*, OR El Salvador, OR Eritrea, OR Eswatini, OR Ethiopia*, OR Equatorial Guinea, OR Fiji, OR Gabon, OR Gambia*, OR Gaza, OR Georgia, OR Ghana*, OR Grenada, OR Guatemala*, OR Guam, OR Guinea*,OR Guyana, OR Haiti*, OR Hondura*, OR India*, OR Indonesia*, OR Iran, OR Iraq, OR Jamaica, OR Jordan, OR Kazakhstan, OR Kenya*, OR Kiribati, OR Korea*, OR Kosovo, OR Kyrgyz*, OR Lao*, OR PDR, OR Lebanon, OR Lesotho, OR Liberia*, OR Libya, OR Madagascar, OR Malawi*, OR Malaysia, OR Maldives, OR Mali, OR Marshall Islands, OR Mauritius, OR Mauritania*, OR Mexico, OR Micronesia, OR Moldova*, OR Mongolia*, OR Montenegro, OR Morocc*, OR Mozambique, OR Burma, OR Myanmar, OR Myanma, OR Namibia, OR Nepal*, OR Nicaragua, OR Niger, OR Nigeria*, OR North Macedonia, OR Palau, OR Pakistan*, OR Paraguay, OR Peru, OR Philippines, OR Philipines, OR Phillipines, OR Phillipines, OR Papua New Guinea, OR Republic of Congo, OR Rwanda, OR Russian Federation, OR*

Samoa, OR Sao Tome and Principe, OR Senegal, or Serbia, OR Sierra Leone, OR Sri Lanka, OR Solomon Islands, OR Somalia*, OR South Africa*, OR South Sudan, OR Sudan, OR St. Lucia, OR St. Vincent, OR Swaziland, OR Suriname, OR Syria*, OR Tajikistan, OR Tanzania*, OR Thailand, OR Timor-Leste, OR Tokelau, OR Togo, OR Tonga, OR Tunisia, OR Turkey, OR Turkmenistan, OR Tuvalu, OR Uganda*, OR Ukraine, OR Uzbekistan, OR Vanuatu, OR Vietnam*, OR Viet Nam, OR West Bank, OR Yemen, OR Zambia*, OR Zimbabwe*

- **Intervention:** *tax, OR trade, OR export, OR tariff, OR technolog*, OR digital technolog*, OR internet, OR phone, OR mHealth, OR mobile health, OR mobile money, OR mobile banking, OR finance, OR savings, OR self-help groups, OR market-led, OR market-based reforms, OR corporate social responsibility, OR certification schemes, OR organic, OR blockchain, OR laptop, OR business support services, OR public-private partnerships*
- **Evaluation:** *evaluation, OR monitoring, OR process, OR performance, OR implementation, OR implementation science, OR fidelity of implementation, OR formative, OR assessment, OR developmental evaluation, OR participatory evaluation*

Final Search Strings

Web of Science

117. We conducted Web of Science searches on 18 February 2023 for all hits that returned from publication date of 1 January 2015, through 1 February 2023.

Population search terms in all searches:

"low income countr" OR "low-income countr*" OR "low-income econom*" OR "low income econom*" OR "lower-middle-income countr*" OR "lower middle income countr*" OR "lower-middle-income econom*" OR "lower middle income econom*" OR "middle income countr*" OR "middle-income countr*" OR "middle-income econom*" OR "middle income enconm*" OR "developing countr*" OR "less developed countr*" OR "less-developed countr*" OR "underdeveloped countr*" OR "under developed countr*" OR "under-developed countr*" OR "underserved countr*" OR "under served countr*" OR "under-served countr*" OR "LMIC*" OR "low GDP" OR "low-GDP" OR "low GNP" OR "low-GNP" OR "fragile state" OR "third world" OR "transitional countr" OR "high burden countr*" OR "high-burden countr*" OR Asia* OR "South Asia*" OR "Africa*" OR "Latin America*" OR "South America*" OR "Central America*" OR "Middle East*" OR "sub-*

Saharan Africa*" OR "sub Saharan Africa*" OR Caribbean OR "West Indies" OR Afghanistan* OR Afghan* OR Albania* OR Algeria* OR "American Samoa*" OR Angola* OR Argentin* OR Armenia* Or Azerbaijan* OR Azeri OR Bangladesh* OR Belarus* OR Belize* OR Benin* OR Bhutan* OR Bolivia* OR Bosnia* OR "Bosnia and Herzegovina" OR Botswana OR Mlotswana OR Brazil* OR Bulgaria* OR "Burkina Faso" OR Burkinabè OR Burkinabe OR Burundi* OR "Cabo Verde*" OR "Cape Verde*" OR Cameroon* OR Cambodia* OR "Central African Republic" OR "Central African" OR Chad* OR China OR Chinese OR Colombia* OR Comoros OR Comorian OR "Cote d'Ivoire" OR "Ivory Coast" OR Ivorian OR Congo* OR "Costa Rica*" OR Cuba* OR "Democratic Republic of Congo" OR "Republic of Congo" OR "Democratic People's Republic of Korea" OR "North Korea*" OR Korea* OR Djibouti* OR Dominica* OR "Dominican Republic" OR Ecuador* OR Egypt* OR "Arab Republic of Egypt" OR "El Salvador" OR Salvador* OR Eritrea* OR Eswatini OR Swazi OR Ethiopia* OR "Equatorial Guinea*" OR Equatoguinean OR Fiji* OR Gabon* OR Gambia* OR Gaza* OR Palestin* OR Georgia* OR Ghana* OR Grenada OR Granad* OR Guatemala* OR Guam* OR Guinea* OR "Guinea-Bissau" OR Guyan* OR Haiti* OR Hondura* OR India* OR Indonesia* OR Iran* OR "Islamic Republic of Iran" OR Iraq* OR Jamaica* OR Jordan* OR Kazakhstan* Or Kazakh* OR Kenya* OR Kiribati OR "I-Kiribati" OR Kosovo OR Kosova* OR Kyrgyz* OR Lao* OR Lao PDR OR Lao People's Democratic Republic OR Lebanon OR Leban* OR Lesotho OR Mosotho OR Basotho OR Liberia* OR Libya* OR Madagascar OR Malagasy OR Malawi* OR Malaysia* OR Maldives OR Maldivian OR Mali* OR "Marshall Islands" OR Marshallese OR Mauritius OR Mauritian OR Mauritania* OR Mexic* OR Micronesia* OR "Federated States of Micronesia" OR Moldova* OR Mongolia* OR Monteneg* OR Morocc* OR Mozambique OR Mozambican OR Burma OR Burmese OR Myanmar OR Myanma* OR Namibia* OR Nepal* OR Nicaragua* OR Niger* OR Nigeria* OR "North Macedonia" OR Macedonian OR Palau* OR Pakistan* OR Paraguay* OR Peru OR Philippines OR Philipines OR Phillipines OR Phillippines OR Filipino OR "Papua New Guinea*" OR "Republic of Congo" OR "Republic of Korea" OR "South Korea*" OR Rwanda OR Rwand* OR "Russian Federation" OR Russia* OR Samoa* OR "Sao Tome and Principe" OR "São Tomé*" OR "Sao Tome*" OR Santomean OR "SãoToméan" OR Senegal* or Serbia* OR "Sierra Leone*" OR "Sri Lanka*" OR "Solomon Island*" OR Somalia* OR "South Africa*" OR "South Sudan*" OR Sudan* OR "St. Lucia" OR "Saint Lucia*" OR "St. Vincent" OR "Saint Vincent and the Grenadines" OR "St. Vincent and the Grenadines" OR "Vincentian and Grenadinian" OR Vincy OR

Swaziland OR Emaswati OR Liswati OR Suriname* OR Syria* OR "Syrian Arab Republic"
OR Tajikistan* Or Tajik OR Tanzania* OR Thailand OR Thai OR "Timor-Leste" OR "Timor
Leste" OR "East Timor*" OR Timorese OR Maubere OR Tokelau* OR Togo* OR Tonga* OR
Tunisia* OR Turkey OR Turkish OR Turkiye OR Turk OR Turkmenistan* Or Turkmen* OR
Tuvalu* OR Uganda* OR Ukraine OR Ukrainian OR Uzbekistan OR Uzbek OR Vanuatu* OR
"Ni-vanuatu" OR Vietnam* OR "Viet Nam" OR "West Bank" OR Gaza* OR Yemen* OR
"Republic of Yemen*" OR Zambia* OR Zimbabwe* OR Zimbo

AND

Design/Comparison terms in all searches:

evaluation OR "impact evaluation" OR "impact analysis" OR "random* control* trial" OR
experiment* OR "quasi-experiment*" OR "regression discontinuity" OR "difference-in-
difference*" OR "difference in difference*" OR "propensity score" OR "systematic review"
OR "rapid review" OR "evidence synthesis" OR "quasi random" OR "quasi-random" OR
"independent evaluation group" OR "independent evaluation" OR "process tracing" OR
"process evaluation" OR "performance evaluation" OR "implementation science" OR
"fidelity of implementation" OR "formative research" OR "formative assessment" OR
"formative evaluation" OR "developmental evaluation" OR "participatory evaluation"
OR "costing analys*" OR "cost-effectiveness analys*" OR "instrumental variable*" OR
"interim evaluation" OR "midterm evaluation" OR "midterm review" OR "summative
evaluation"

AND

Intervention/Outcome terms specific to each SDG-17 area:

Trade

"anti-dumping*" OR "duty free" OR "duty-free" OR "export competi*" OR "export
market*" OR "export processing zone*" OR "export promotion" OR "export sector"
OR "export subsid*" OR "export value" OR "export*" OR "FDI" OR "foreign direct
investment*" OR "free trade agreement*" OR "free trade union" OR "green
procurement rule*" OR "import" OR "international trade" OR "intra-regional
trade" OR "intraregional trade" OR "non-tariff barrier" OR "preferential rules of
origin" OR "quantity restriction" OR "quota free" OR "quota-free" OR "regional
integration" OR "regional trade" OR "rules based trade" OR "rules-based trade"

OR "tariff average" OR "tariff reduction" OR "tariff*" OR "trade barrier*" OR "trade liberali*" OR "trade liberalization" OR "trade quota*" OR "trade reform*" OR "trade support" OR "weighted average tariff" OR "weighted average tariff" OR "weighted tariff average" OR "weighted tariff-average"

Finance

"aid dependency" OR "blended finance" OR "budget deficit" OR "budget support" OR "capital flow*" OR "carbon financ*" OR "clean energy financ*" OR "climate financ*" OR "concessional financ*" OR "debt as a percentage of GDP" OR "debt financ*" OR "debt relief" OR "debt restructur*" OR "debt servic*" OR "debt sustainability" OR "direct budget support" OR "domestic financ*" OR "domestic resource mobilization" OR "domestic revenue" OR "domestic tax*" OR " environmental financ*" OR "environmental tax*" OR "export value added tax*" OR "export value-added tax*" OR "external debt" OR "finance corporation*" OR "government deficit" OR "government revenue" OR "green financ*" OR "guarantee agenc*" OR "guarantee instrument*" OR "highly indebted poor countr*" OR "HIPC" OR "income tax*" OR "international monetary*" OR "investment guarantee agenc*" OR "investment promotion" OR "migrant income" OR "multi-lateral investment guarantee" OR "multilateral investment guarantee" OR "national budget support" OR "ODA" OR "official development assistance" OR "overseas development flows" OR "public debt" OR "remittanc*" OR "renewable energy financ*" OR "tax audit*" OR "tax collection" OR "tax restructur*" OR "tax revenue" OR "value added tax rebate*" OR "VAT rebate*" OR "Western Union"

Technology

blockchain OR broadband OR "broadband access" OR "digital divide" OR "digital technolog*" OR "digital transformation" OR "energy technolog*" OR hotspot OR "information and communication technology" OR "information technology" OR "communication technology" OR "internet access" OR "internet us*" OR "laptop*" OR "mobile internet access" OR "mobile phone*" OR phone* OR smartphone* OR "smart phone" OR "social media" OR smartphone* OR "smart phone" OR "social media" OR "solar panel*" OR "solar system*" OR "solar power" OR "solar device*" OR tablet* OR "technolog* innovation" OR

“technology access” OR “technology bank*” OR “technology diffusion” OR “WhatsApp” OR “ICT” OR “mobile money” OR “digital wallets” OR “digital bank account” OR MPESA

Systemic Issues

civil society organization” OR “United Nation* coordination” OR “UN* coordination” OR “country ownership” OR “CSO” OR “data monitoring accountability” OR “development co-operation” OR “development cooperation” OR “development impact bond*” OR “economic co-operation” OR “economic cooperation” OR “global coordination” OR “global partnership*” OR “impact invest*” OR “institutional coherence” OR “matching grant*” OR “multi stakeholder” OR “multi-stakeholder” OR “multiple stakeholders” OR “statistical legislation” OR “national statistical office*” OR “national statistical plan*” OR “national statistical system*” OR “north-south” OR “south-south” OR “performance based financing” OR “performance-based financing” OR “policy coherence” OR “public procurement*” OR “public-private partnership*” OR “SDG policy” OR “Sustainable Development Goal policy” OR “triangular cooperation” OR “triangular co-operation”

Capacity Building

“capacity building” OR “capacity development” OR “capacity strengthening” OR “capacity-building intervention” OR “implementation management” OR “national capacity” OR “national planning” OR “public sector training” OR “SDG planning” OR “technical assistance to government*” OR “training of government”

3ie Database

118. We conducted 3ie searches on 2/22/2023 for all hits that returned from publication date of January 1, 2015, through February 1, 2023. The following list of search terms should be searched for Title, Abstract, and Keyword fields only (using the search string below specific for the 3ie database).

Design/Comparison terms in all searches:

(title:(“evaluation OR “impact evaluation” OR “impact analysis” OR “random* control* trial” OR experiment* OR “quasi-experiment*” OR “regression discontinuity” OR “difference-in-difference*” OR “difference in difference*” OR “propensity score” OR “systematic review” OR “rapid review” OR “evidence synthesis” OR “quasi random” OR “quasi-random” OR “independent evaluation group” OR “independent evaluation” OR “process tracing” OR “process evaluation” OR “performance evaluation” OR “implementation science” OR “fidelity of implementation” OR “formative research” OR “formative assessment” OR “formative evaluation” OR “developmental evaluation” OR “participatory evaluation” OR “costing analys*” OR “cost-effectiveness analys*” OR “instrumental variable*” OR “interim evaluation” OR “midterm evaluation” OR “midterm review” OR “summative evaluation”) OR keywords: (“evaluation OR “impact evaluation” OR “impact analysis” OR “random* control* trial” OR experiment* OR “quasi-experiment*” OR “regression discontinuity” OR “difference-in-difference*” OR “difference in difference*” OR “propensity score” OR “systematic review” OR “rapid review” OR “evidence synthesis” OR “quasi random” OR “quasi-random” OR “independent evaluation group” OR “independent evaluation” OR “process tracing” OR “process evaluation” OR “performance evaluation” OR “implementation science” OR “fidelity of implementation” OR “formative research” OR “formative assessment” OR “formative evaluation” OR “developmental evaluation” OR “participatory evaluation” OR “costing analys*” OR “cost-effectiveness analys*” OR “instrumental variable*” OR “interim evaluation” OR “midterm evaluation” OR “midterm review” OR “summative evaluation”) OR abstract: (“evaluation OR “impact evaluation” OR “impact analysis” OR “random* control* trial” OR experiment* OR “quasi-experiment*” OR “regression discontinuity” OR “difference-in-difference*” OR “difference in difference*” OR “propensity score” OR “systematic review” OR “rapid review” OR “evidence synthesis” OR “quasi random” OR “quasi-random” OR “independent evaluation group” OR “independent evaluation” OR “process tracing” OR “process evaluation” OR “performance evaluation” OR “implementation science” OR “fidelity of implementation” OR “formative research” OR “formative assessment” OR “formative evaluation” OR “developmental evaluation” OR “participatory evaluation” OR “costing analys*” OR “cost-effectiveness analys*” OR “instrumental variable*” OR “interim

evaluation" OR "midterm evaluation" OR "midterm review" OR "summative evaluation"))

AND

Intervention/Outcome terms specific to each SDG-17 area:

Trade

(title:("anti-dumping*" OR "duty free" OR "duty-free" OR "export competi*" OR "export market*" OR "export processing zone*" OR "export promotion" OR "export sector" OR "export subsid*" OR "export value" OR "export*" OR "FDI" OR "foreign direct investment*" OR "free trade agreement*" OR "free trade union" OR "green procurement rule*" OR "import" OR "international trade" OR "intra-regional trade" OR "intraregional trade" OR "non-tariff barrier" OR "preferential rules of origin" OR "quantity restriction" OR "quota free" OR "quota-free" OR "regional integration" OR "regional trade" OR "rules based trade" OR "rules-based trade" OR "tariff average" OR "tariff reduction" OR "tariff*" OR "trade barrier*" OR "trade liberali*" OR "trade liberalization" OR "trade quota*" OR "trade reform*" OR "trade support" OR "weighted average tariff" OR "weighted average tariff" OR "weighted tariff average" OR "weighted tariff-average") OR abstract:("anti-dumping*" OR "duty free" OR "duty-free" OR "export competi*" OR "export market*" OR "export processing zone*" OR "export promotion" OR "export sector" OR "export subsid*" OR "export value" OR "export*" OR "FDI" OR "foreign direct investment*" OR "free trade agreement*" OR "free trade union" OR "green procurement rule*" OR "import" OR "international trade" OR "intra-regional trade" OR "intraregional trade" OR "non-tariff barrier" OR "preferential rules of origin" OR "quantity restriction" OR "quota free" OR "quota-free" OR "regional integration" OR "regional trade" OR "rules based trade" OR "rules-based trade" OR "tariff average" OR "tariff reduction" OR "tariff*" OR "trade barrier*" OR "trade liberali*" OR "trade liberalization" OR "trade quota*" OR "trade reform*" OR "trade support" OR "weighted average tariff" OR "weighted average tariff" OR "weighted tariff average" OR "weighted tariff-average") OR keywords:("anti-dumping*" OR "duty free" OR "duty-free" OR "export competi*" OR "export market*" OR "export processing zone*" OR "export promotion" OR "export sector" OR "export subsid*" OR "export value" OR "export*" OR "FDI" OR "foreign direct

investment*" OR "free trade agreement*" OR "free trade union" OR "green procurement rule*" OR "import" OR "international trade" OR "intra-regional trade" OR "intraregional trade" OR "non-tariff barrier" OR "preferential rules of origin" OR "quantity restriction" OR "quota free" OR "quota-free" OR "regional integration" OR "regional trade" OR "rules based trade" OR "rules-based trade" OR "tariff average" OR "tariff reduction" OR "tariff*" OR "trade barrier*" OR "trade liberali*" OR "trade liberalization" OR "trade quota*" OR "trade reform*" OR "trade support" OR "weighted average tariff" OR "weighted average tariff" OR "weighted tariff average" OR "weighted tariff-average"))

Finance

(title:(“aid dependency” OR “blended finance” OR “budget deficit” OR “budget support” OR “capital flow*” OR “carbon financ*” OR “clean energy financ*” OR “climate financ*” OR “concessional financ*” OR “debt as a percentage of GDP” OR “debt financ*” OR “debt relief” OR “debt restructur*” OR “debt servic*” OR “debt sustainability” OR “direct budget support” OR “domestic financ*” OR “domestic resource mobilization” OR “domestic revenue” OR “domestic tax*” OR “ environmental financ*” OR “environmental tax*” OR “export value added tax*” OR “export value-added tax*” OR “external debt” OR “finance corporation*” OR “government deficit” OR “government revenue” OR “green financ*” OR “guarantee agenc*” OR “guarantee instrument*” OR “highly indebted poor countr*” OR “HIPC” OR “income tax*” OR “international monetary*” OR “investment guarantee agenc*” OR “investment promotion” OR “migrant income” OR “multi-lateral investment guarantee” OR “multilateral investment guarantee” OR “national budget support” OR “ODA” OR “official development assistance” OR “overseas development flows” OR “public debt” OR “remittanc*” OR “renewable energy financ*” OR “tax audit*” OR “tax collection” OR “tax restructur*” OR “tax revenue” OR “value added tax rebate*” OR “VAT rebate*” OR “Western Union”) OR abstract:(“aid dependency” OR “blended finance” OR “budget deficit” OR “budget support” OR “capital flow*” OR “carbon financ*” OR “clean energy financ*” OR “climate financ*” OR “concessional financ*” OR “debt as a percentage of GDP” OR “debt financ*” OR “debt relief” OR “debt restructur*” OR “debt servic*” OR “debt sustainability” OR “direct budget support” OR “domestic financ*” OR “domestic resource mobilization” OR “domestic revenue”

OR "domestic tax*" OR " environmental financ*" OR "environmental tax*" OR "export value added tax*" OR "export value-added tax*" OR "external debt" OR "finance corporation*" OR "government deficit" OR "government revenue" OR "green financ*" OR "guarantee agenc*" OR "guarantee instrument*" OR "highly indebted poor countr*" OR "HIPC" OR "income tax*" OR "international monetary*" OR "investment guarantee agenc*" OR "investment promotion" OR "migrant income" OR "multi-lateral investment guarantee" OR "multilateral investment guarantee" OR "national budget support" OR "ODA" OR "official development assistance" OR "overseas development flows" OR "public debt" OR "remittanc*" OR "renewable energy financ*" OR "tax audit*" OR "tax collection" OR "tax restructur*" OR "tax revenue" OR "value added tax rebate*" OR "VAT rebate*" OR "Western Union") OR keywords:("aid dependency" OR "blended finance" OR "budget deficit" OR "budget support" OR "capital flow*" OR "carbon financ*" OR "clean energy financ*" OR "climate financ*" OR "concessional financ*" OR "debt as a percentage of GDP" OR "debt financ*" OR "debt relief" OR "debt restructur*" OR "debt servic*" OR "debt sustainability" OR "direct budget support" OR "domestic financ*" OR "domestic resource mobilization" OR "domestic revenue" OR "domestic tax*" OR " environmental financ*" OR "environmental tax*" OR "export value added tax*" OR "export value-added tax*" OR "external debt" OR "finance corporation*" OR "government deficit" OR "government revenue" OR "green financ*" OR "guarantee agenc*" OR "guarantee instrument*" OR "highly indebted poor countr*" OR "HIPC" OR "income tax*" OR "international monetary*" OR "investment guarantee agenc*" OR "investment promotion" OR "migrant income" OR "multi-lateral investment guarantee" OR "multilateral investment guarantee" OR "national budget support" OR "ODA" OR "official development assistance" OR "overseas development flows" OR "public debt" OR "remittanc*" OR "renewable energy financ*" OR "tax audit*" OR "tax collection" OR "tax restructur*" OR "tax revenue" OR "value added tax rebate*" OR "VAT rebate*" OR "Western Union"))

Technology

(title:(blockchain OR broadband OR "broadband access" OR "digital divide" OR " digital technolog*" OR "digital transformation" OR "energy technolog*" OR hotspot OR "information and communication technology" OR "information

technology" OR "communication technology" OR "internet access" OR "internet us*" OR "laptop*" OR "mobile internet access" OR "mobile phone*" OR phone* OR smartphone* OR "smart phone" OR "social media" OR smartphone* OR "smart phone" OR "social media" OR "solar panel*" OR "solar system*" OR "solar power" OR "solar device*" OR tablet* OR "technolog* innovation" OR "technology access" OR "technology bank*" OR "technology diffusion" OR "WhatsApp" OR "ICT" OR "mobile money" OR "digital wallets" OR "digital bank account" OR MPESA) OR abstract:(blockchain OR broadband OR "broadband access" OR "digital divide" OR " digital technolog*" OR "digital transformation" OR "energy technolog*" OR hotspot OR "information and communication technology" OR "information technology" OR "communication technology" OR "internet access" OR "internet us*" OR "laptop*" OR "mobile internet access" OR "mobile phone*" OR phone* OR smartphone* OR "smart phone" OR "social media" OR smartphone* OR "smart phone" OR "social media" OR "solar panel*" OR "solar system*" OR "solar power" OR "solar device*" OR tablet* OR "technolog* innovation" OR "technology access" OR "technology bank*" OR "technology diffusion" OR "WhatsApp" OR "ICT" OR "mobile money" OR "digital wallets" OR "digital bank account" OR MPESA) OR keywords:(blockchain OR broadband OR "broadband access" OR "digital divide" OR " digital technolog*" OR "digital transformation" OR "energy technolog*" OR hotspot OR "information and communication technology" OR "information technology" OR "communication technology" OR "internet access" OR "internet us*" OR "laptop*" OR "mobile internet access" OR "mobile phone*" OR phone* OR smartphone* OR "smart phone" OR "social media" OR smartphone* OR "smart phone" OR "social media" OR "solar panel*" OR "solar system*" OR "solar power" OR "solar device*" OR tablet* OR "technolog* innovation" OR "technology access" OR "technology bank*" OR "technology diffusion" OR "WhatsApp" OR "ICT" OR "mobile money" OR "digital wallets" OR "digital bank account" OR MPESA))

Systemic Issues

(title:("civil society organization" OR "United Nation* coordination" OR "UN* coordination" OR "country ownership" OR "CSO" OR "data monitoring accountability" OR "development co-operation" OR "development cooperation"

OR "development impact bond*" OR "economic co-operation" OR "economic cooperation" OR "global coordination" OR "global partnership*" OR "impact invest*" OR "institutional coherence" OR "matching grant*" OR "multi stakeholder" OR "multi-stakeholder" OR "multiple stakeholders" OR "statistical legislation" OR "national statistical office*" OR "national statistical plan*" OR "national statistical system*" OR "north-south" OR "south-south" OR "performance based financing" OR "performance-based financing" OR "policy coherence" OR "public procurement*" OR "public-private partnership*" OR "SDG policy" OR "Sustainable Development Goal policy" OR "triangular cooperation" OR "triangular co-operation") OR abstract:("civil society organization" OR "United Nation* coordination" OR "UN* coordination" OR "country ownership" OR "CSO" OR "data monitoring accountability" OR "development co-operation" OR "development cooperation" OR "development impact bond*" OR "economic co-operation" OR "economic cooperation" OR "global coordination" OR "global partnership*" OR "impact invest*" OR "institutional coherence" OR "matching grant*" OR "multi stakeholder" OR "multi-stakeholder" OR "multiple stakeholders" OR "statistical legislation" OR "national statistical office*" OR "national statistical plan*" OR "national statistical system*" OR "north-south" OR "south-south" OR "performance based financing" OR "performance-based financing" OR "policy coherence" OR "public procurement*" OR "public-private partnership*" OR "SDG policy" OR "Sustainable Development Goal policy" OR "triangular cooperation" OR "triangular co-operation") OR keywords:("civil society organization" OR "United Nation* coordination" OR "UN* coordination" OR "country ownership" OR "CSO" OR "data monitoring accountability" OR "development co-operation" OR "development cooperation" OR "development impact bond*" OR "economic co-operation" OR "economic cooperation" OR "global coordination" OR "global partnership*" OR "impact invest*" OR "institutional coherence" OR "matching grant*" OR "multi stakeholder" OR "multi-stakeholder" OR "multiple stakeholders" OR "statistical legislation" OR "national statistical office*" OR "national statistical plan*" OR "national statistical system*" OR "north-south" OR "south-south" OR "performance based financing" OR "performance-based financing" OR "policy coherence" OR "public procurement*" OR "public-private partnership*" OR "SDG policy" OR

“Sustainable Development Goal policy” OR “triangular cooperation” OR “triangular co-operation”))

Capacity Building

(title:(“capacity building” OR “capacity development” OR “capacity strengthening” OR “capacity-building intervention” OR “implementation management” OR “national capacity” OR “national planning” OR “public sector training” OR “SDG planning” OR “technical assistance to government*” OR “training of government”) OR abstract:(“capacity building” OR “capacity development” OR “capacity strengthening” OR “capacity-building intervention” OR “implementation management” OR “national capacity” OR “national planning” OR “public sector training” OR “SDG planning” OR “technical assistance to government*” OR “training of government”) OR keywords:(“capacity building” OR “capacity development” OR “capacity strengthening” OR “capacity-building intervention” OR “implementation management” OR “national capacity” OR “national planning” OR “public sector training” OR “SDG planning” OR “technical assistance to government*” OR “training of government”))

Annex B. Critical Appraisal Tools

Exhibit B-1: Risk of Bias Tool for Experimental and Quasi-Experimental Studies

Ask these questions for all quantitative studies

Does the study show baseline values of the outcomes of interest (as defined in the protocol) for beneficiaries and non-beneficiaries?

If baseline values of the outcome of interest are not available at baseline, does the study show baseline values of characteristics of beneficiaries and non-beneficiaries that are not likely to be affected by the intervention?

Are the mean values or the distributions of the covariates at baseline statistically different for beneficiaries and non-beneficiaries ($p < 0.05$)?

If there are statistically significant differences between beneficiaries and non-beneficiaries, are these differences controlled for using covariate analysis in the impact evaluation?

If baseline characteristics are not available, does the study qualitatively assess why beneficiaries are likely/unlikely to be a random draw of the population at baseline?

Confounding and selection bias (ask questions for all quantitative studies)

Does the study use a comparison/control group of women without access to the programme?

Does the study use a comparison/control group of women with access to the programme but who did not choose to participate in the programme?

Does the study include data on the outcomes of interest at baseline and endline (before and after the intervention)?

Are the data on covariates collected at the baseline?

Is difference-in-difference estimation used?

If the study is quasi-experimental and uses difference-in-difference estimation, do the authors assess the parallel trends assumption?

If the study does not use difference-in-difference, does the study control for baseline values of the outcome of interest (ANCOVA)?

If the study does not use difference-in-difference and does not control for baseline values of the outcome variable, does the study control for other covariates at baseline?

If the study does not use difference in differences estimation, is there any assessment of likely risk of bias from time invariant characteristics driving both participation and outcome?

If the study does not use difference-in-difference estimation but does assess likely risk of bias from time invariant characteristics, are these time invariant characteristics likely to bias the impact estimates?

Does the study report the table with the results of the outcome equation (including covariates)?

Where full results of the outcome equation are not reported, is it clear which covariates have been used?

Are all relevant observable covariates (confounding variables) included in the outcome equation that might explain outcomes, if estimation does not use a statistical technique to control for selection bias (RCT, PSM, RDD, or IV)?

Attrition (ask questions for all quantitative studies)

For studies including baseline data, does the study report attrition (drop-out) from the study?

Is the attrition rate from the study below 10%?

Does the study assess whether drop-outs from the study are random draws from the sample (e.g., by examining correlation with determinants of outcomes, in both treatment comparison group)?

Spillovers and contamination (ask questions for all quantitative studies)

Spillovers: are comparisons sufficiently isolated from the intervention (e.g., participants and non-participants are sufficiently geographically or socially separated) or are spillovers estimated by comparing non-beneficiaries with access to the intervention to non-beneficiaries without access to the intervention and/or through social network analysis?

Spillovers: if spillovers are not estimated, is the study likely to overestimate or underestimate the impact of the programme?

Contamination: does the study assess whether the control group receives the intervention?

Contamination: if the control group receives the intervention but for a shorter amount of time, does the study assess the likelihood that the control group has received equal benefits as the treatment group?

Contamination: if the control group receives the intervention, have they received the intervention sufficiently long to argue that they have benefited from the intervention?

Confidence Intervals (ask questions for all quantitative studies)

Does the study account for lack of independence between observations within assignment clusters if the outcome variables are clustered?

Is the sample size likely to be sufficient to find significant effects of the intervention?

Do the authors control for heteroskedasticity and/or use robust standard errors?

Ask questions below only for studies that apply randomization

Does the study apply randomized assignment?

Does the study use a unit of allocation with a sufficiently large sample size to ensure equivalence between the treatment and the control groups?

Ask questions below only for studies that apply regression discontinuity designs

Is the allocation of the programme based on a pre-determined continuity on a continuous variable and blinded to the beneficiaries or, if not blinded, individuals cannot reasonably affect the assignment variable in response to knowledge of the participation rule?

Is the sample size immediately at both sides of the cut-off point sufficiently large to equate groups on average?

Is the mean of the covariates of individuals immediately at both sides of the cut-off point statistically significantly different for beneficiaries and non-beneficiaries?

If there are statistically significant differences between beneficiaries and non-beneficiaries, are these differences controlled for using covariate analysis?

Ask questions below only for studies that apply matching

Quality of matching (PSM, covariate matching)

Are beneficiaries and non-beneficiaries matched on all relevant characteristics?

Does the study report the results of the matching function (e.g., for PSM the logit function)?

Does the study report the matching method?

Does the study exclude observations outside the common support?

Does the study use variables at follow-up that can be affected by the intervention in the matching equation?

Does the study report the mean or distribution for the covariates of the treatment and control groups after matching?

Are these characteristics similar, based on tests for statistically significant differences ($p > 0.05$)?

Ask questions below only for studies that apply instrumental variable estimation

Does the study describe clearly the instrumental variable(s)/identifier used?

Are the results of the participation equation reported?

Are the instruments jointly significant at the level of $F \geq 10$? If an F test is not reported, does the author report and assess whether the R-squared of the instrumenting equation is large enough for appropriate identification ($R\text{-sq} > 0.5$)?

Are the instruments individually significant ($p \leq 0.05$)?

For IV, if more than one instrument is used in the procedure, does the study include and report an overidentifying test ($p \leq 0.05$ is required to reject the null hypothesis)?

Does the study qualitatively assess the exogeneity of the instrument/identifier (both externality as well as why the variable should not enter by itself in the outcome equation)?

Ask questions below only for studies with censored outcome variables

Do the authors use appropriate methods (e.g., Heckman selection models, tobit models, duration models) to account for the censoring of the data?

For Heckman models, is there is a variable that is statistically significant in the first stage of the selection equation and excluded from the second stage?

Exhibit B-2: Qualitative Review Protocol

Number	Question
SECTION A	INTERVENTION, CONTEXT, AND KEY STAKEHOLDERS
Question 1.	Is the object of the evaluation clearly described?
i	Clear and relevant description of the intervention, including: location(s), timelines, cost/budget, and implementation status.
ii	Clear and relevant description of intended rightsholders (beneficiaries) and duty bearers (state and non-state actors with responsibilities regarding the intervention) by type (i.e., institutions/organizations, communities, individuals), by geographic location(s) (i.e., urban, rural, particular neighbourhoods, town/cities, subregions) and in terms of numbers reached, with disaggregation by gender, age, disability (as appropriate to the purpose of the evaluation).
Question 2.	Is the context of the intervention clearly described?
i	Clear and relevant description of the context of the intervention (i.e., relevant policy, socio-economic, political, cultural, power/privilege, institutional, international factors) and how context relates to the implementation of the intervention.
ii	Linkages drawn to the SDGs and relevant targets and indicators for the area being evaluated.
iii	Clear and relevant description of the status and needs of the rightsholders/beneficiaries of the intervention.
Question 3.	Are key stakeholders, their relationships, and contributions clearly identified?
i	Identification of implementing agency(ies), development partners, right holders, and additional duty bearers and other stakeholders; and of linkages between them (e.g., stakeholder map) (if relevant).
ii	Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UN agencies.
SECTION B:	EVALUATION PURPOSE, OBJECTIVES, AND SCOPE
Question 4.	Is the purpose of the evaluation clearly described?
i	Purpose of evaluation is clearly defined, including why it was needed at that point in time, its intended use, and key intended users.
ii	Clear and relevant description of the scope of the evaluation: what will and will not be covered (thematically, chronologically, geographically with key terms defined), as well as, if applicable, the reasons for this scope (e.g., specifications by the Terms of Reference, lack of access to particular geographic areas)

	for political or safety reasons at the time of the evaluation, lack of data/evidence on particular elements of the intervention).
Question 5.	Is the theory of change, results chain or logic model well articulated?
i	Clear description of the intervention's intended results, or of the parts of implementation that are applicable to, or are being assessed by, the evaluation.
ii	Relationship between implementation components, including pathways from input to activities to outputs, is presented in narrative and/or graphic form (e.g., logic model, theory of change, evaluation matrix).
iii	For theory-based evaluations, the theory of change or results framework is assessed.
SECTION C:	EVALUATION DESIGN AND METHODOLOGY
Question 6.	Does the evaluation use questions and the relevant evaluation criteria that are explicitly justified as appropriate for the purpose of the evaluation?
i	Evaluation questions and sub-questions are appropriate for meeting the objectives and purpose of the evaluation. The relevant criteria are specified and are aligned with the questions.
ii	In addition to the questions and sub-questions, the evaluation matrix includes indicators, benchmarks, assumptions, and/or other processes from which the analysis can be based and conclusions drawn.
Question 7.	Does the report specify adequate methods for data collection, analysis, and sampling?
i	Evaluation design and set of methods are relevant and adequately robust for the evaluation's purpose, objectives, and scope and are fully and clearly described.
ii	Qualitative and quantitative data sources are appropriate and are clearly described.
iii	Sampling strategy is provided - it should include a description of how diverse perspectives are captured (or, if not, provide reasons for this), with articulated consideration and/or inclusion of vulnerable/marginalized groups, equity, and intersectionality
iv	Clear and complete description of the methods of data analysis.
v	Clear and complete description of limitations and constraints faced by the evaluation, including gaps in the evidence that was generated and mitigation of bias and how these were addressed by the evaluators (as feasible).
Question 8.	Are ethical issues and considerations described?
i	Explicit and contextualized reference to the obligations of evaluators (independence, impartiality, credibility, conflicts of interest, accountability).

ii	Description of ethical safeguards for participants appropriate for the issues relevant to methodology and how they are applied (respect for dignity and diversity, right to self-determination, fair representation, compliance with codes for vulnerable groups, confidentiality, and avoidance of harm).
SECTION D:	EVALUATION FINDINGS
Question 9.	Do the findings clearly address all evaluation objectives and scope?
i	Findings marshal sufficient levels of evidence to systematically address all of the evaluation's questions, sub-questions and criteria.
ii	Explicit use of the intervention's results framework/ToC/logic model in the formulation of the findings.
Question 11.	Are evaluation findings derived from the conscientious, explicit, and judicious use of the best available, objective, reliable, and valid data and by accurate quantitative and qualitative analysis of evidence.
i	Evaluation uses credible forms of qualitative and quantitative data. It presents both output and outcome-level data as relevant to the evaluation framework. Triangulation is evident through the use of multiple data sources.
ii	Findings are clearly supported by, and respond to, the evidence presented, including both positive and negative. Findings are based on clear performance indicators, standards, benchmarks, or other means of comparison as relevant for each question.
iii	Factors (contextual, organizational, managerial, etc.) related to successful or unsuccessful implementation are clearly identified. For theory-based evaluations, findings analyse the logical chain of implementation processes expected to produce targeted results.
Question 12.	Does the evaluation assess and use the intervention's Results-Based Management elements?
i	Assessment of the adequacy of the intervention's monitoring system (including completeness and appropriateness of results/performance framework - including vertical and horizontal logic, M&E tools, and their usage) to support decision-making.
SECTION E:	EVALUATION CONCLUSIONS & LESSONS LEARNED
Question 13.	Do the conclusions clearly present an objective overall assessment of the intervention?
i	Conclusions are clearly formulated and reflect the purpose and objectives of the evaluation. They are sufficiently forward-looking (if a formative evaluation or if the implementation is expected to continue or have additional phase).

ii	Conclusions are derived appropriately from findings and present a picture of the strengths and limitations of the intervention that adds insight and analysis beyond the findings.
Question 14.	Are logical and informative lessons learned identified? [N/A if lessons are not presented and not requested in ToR]
i	Identified lessons stem logically from the findings, have wider applicability and relevance beyond the object of the evaluation.
ii	Lessons are clearly and concisely presented, yet have sufficient detail to be useful for intended audience.
SECTION F:	RECOMMENDATIONS
Question 15.	Are recommendations well grounded in the evaluation?
i	Recommendations align with the evaluation purpose, are clearly formulated and logically derived from the findings and/or conclusions.
ii	Recommendations are useful and actionable for primary intended users and uses (relevant to the intervention); guidance is given for implementation, as appropriate.
iii	Process for developing the recommendations is described, and includes the involvement of duty-bearers, as well as rights holders when feasible (or explanation given for why they were not involved).
Question 16.	Are recommendations clearly presented?
i	Clear identification of groups or duty-bearers responsible for action for each recommendation (or clearly clustered group of recommendations). Clear prioritization and/or classification of recommendations to support use.

Annex C. Anchor Articles Reference List (organized by SDG-17 content area)

Capacity Building

- Amare, A. T., Toni, A. T., Mekonnen, Z. A., Endehabtu, B. F., & Tilahun, B. C. (2021). Effectiveness and Feasibility of Using Local Medical Universities for Capacity Building to Improve the Immunization Program in Ethiopia: Quasi-Experimental Study. *Journal of Multidisciplinary Healthcare*, Volume 14, 9–19. <https://doi.org/10.2147/JMDH.S285280>
- Guarini, E., Mori, E., & Zuffada, E. (2022). Localizing the Sustainable Development Goals: A managerial perspective. *Journal of Public Budgeting, Accounting & Financial Management*, 34(5), 583–601. <https://doi.org/10.1108/JPBAFM-02-2021-0031>
- Keynejad, R., Semrau, M., Toynbee, M., Evans-Lacko, S., Lund, C., Gureje, O., Ndyanabangi, S., Courtin, E., Abdulmalik, J. O., Alem, A., Fekadu, A., Thornicroft, G., & Hanlon, C. (2016). Building the capacity of policy-makers and planners to strengthen mental health systems in low- and middle-income countries: A systematic review. *BMC Health Services Research*, 16(1), 601. <https://doi.org/10.1186/s12913-016-1853-0>
- Khan, A. Q., & Khwaja, A. I. (2018). Tax farming redux: Experimental evidence on performance pay for tax collectors. *Quarterly Journal of Economics*, 131(1), 219–271.
- OECD. (n.d.-a). *Achieving SDG Results in Development Co-operation: Summary for Policy Makers*.
- OECD. (n.d.-b). *OECD Toolkit for territorial approach to the SDGs*.
- Schnatz, J., A. Guerrero-Ruiz and Sachin, K. (2022). *Aligning development co-operation to the SDGs in lower middle-income countries: A case study of Bangladesh*. OECD Development Co-operation Working Papers, No. 105, OECD Publishing, Paris. <https://doi.org/10.1787/7c83ab20-en>

Sobir, R. (2019). UNDESA – Review of capacity development gaps, needs, and priorities.

Umeh, C. A. (2016). Evaluating the effectiveness of online capacity building resources on capacity improvement of local Nigerian NGOs serving orphans and vulnerable children.

Finance (18; 3 academic papers; 15 reports)

Aiyar, P. S., Chen, J., Ebeke, C., Gudmundsson, T., Ilyina, A., Kangur, A., Kunaratskul, T., Rodriguez, S., Ruta, M., Schulze, T., Soderberg, G., & Trevino, J. P. (n.d.). Geoeconomic Fragmentation and the Future of Multilateralism. STAFF DISCUSSION NOTES.

Climate Investment Funds. (2022). Enablers the role of enabling environment in scaling up climate finance. Climate Investment Funds.

Independent Evaluation Group. (2020). The International Finance Corporation's Blended Finance Operations: Findings from a Cluster of Project Performance Assessment Reports. World Bank, Washington, DC.
<https://doi.org/10.1596/33202>

International Monetary Fund. Strategy, Policy, & Review Department. (2018). Guidance Note on the Bank-Fund Debt Sustainability Framework for Low Income Countries. Policy Papers, 18(009), 1.
<https://doi.org/10.5089/9781498307260.007>

International Monetary Fund & World Bank. (2021). World Bank Group And International Monetary Fund Support For Debt Relief Under The Common Framework And Beyond. Policy Papers, 2021(022), 1.
<https://doi.org/10.5089/9781513576039.007>

Mbaye, S., Moreno Badia, M., & Chae, K. (2018). Global Debt Database: Methodology and Sources. IMF Working Papers, 18(111).
<https://doi.org/10.5089/9781484353592.001>

Nakhimovsky, S. S., Feigl, A. B., Avila, C., O'Sullivan, G., Macgregor-Skinner, E., & Spranca, M. (2016). Taxes on Sugar-Sweetened Beverages to Reduce Overweight and Obesity in Middle-Income Countries: A Systematic Review. PLOS ONE, 11(9), e0163358. <https://doi.org/10.1371/journal.pone.0163358>

- O'Brien, A., & Ulhoa, L. (2021). The World Bank Group's Early Support to Addressing the COVID-19 Economic Response April 2020–June 2021 An Early-Stage Evaluation. The World Bank.
- O'Brien, A., Ulhoa, L., & Siegel, C. (2022). The International Finance Corporation's and Multilateral Investment Guarantee Agency's Support for Private Investment in Fragile and Conflict-Affected Situations, Fiscal Years 2010–21 An Independent Evaluation.
- OECD. (2022). Revenue Statistics 2022: The Impact of COVID-19 on OECD Tax Revenues. OECD Publishing. <https://doi.org/10.1787/8a691b03-en>
- OECD/UNCDF. (2020). Blended Finance in the Least Developed Countries 2020: Supporting a Resilient COVID-19 Recovery. OECD Publishing. <https://doi.org/10.1787/57620d04-en>
- World Bank. (2016). Findings from Evaluations of Policy-Based Guarantees. World Bank.
- World Bank. (2017). World Bank Group Joint Projects: A Review of Two Decades of Experience Lessons and Implications from Evaluation. World Bank.
- World Bank. (2019). Leveraging economic migration for development: A Briefing for the World Bank Board. World Bank.
- World Bank. (2021). World Bank Group Approaches to Mobilize Private Capital for Development: An Independent Evaluation. World Bank. <https://doi.org/10.1596/35040>
- World Bank. (2022). The Multilateral Investment Guarantee Agency's Experience with Non-Honoring of Sovereign, SubSovereign, and StateOwned Enterprise Financial Obligation Guarantees Meso-Evaluation. World Bank.
- World Bank. (2023). 17 Partnerships for the goals: Remittances: A lifeline for many economies (pp. 1–16). <https://datatopics.worldbank.org/sdgtlas/goal-17-partnerships-for-the-goals/>

Systematic Issues (20; 4 academic papers; 16 reports)

- Aseete, P., Barkley, A., Katungi, E., Ugen, M. A., & Birachi, E. (2023). Public–private partnership generates economic benefits to smallholder bean growers in Uganda. *Food Security*, 15(1), 201–218. <https://doi.org/10.1007/s12571-022-01309-5>
- Beck, D., & Ferasso, M. (2023). How can Stakeholder Capitalism contribute to achieving the Sustainable Development Goals? A Cross-network Literature Analysis. *Ecological Economics*, 204, 107673. <https://doi.org/10.1016/j.ecolecon.2022.107673>
- Chong, A., & Valdivia, M. (2018). Health Impacts of a Public–Private Partnership Rural Roads Maintenance Programme: Health Impacts of a Public–Private Partnership. *Journal of International Development*, 30(7), 1248–1251. <https://doi.org/10.1002/jid.3393>
- Clarke, A., & Crane, A. (2018). Cross-Sector Partnerships for Systemic Change: Systematized Literature Review and Agenda for Further Research. *Journal of Business Ethics*, 150(2), 303–313. <https://doi.org/10.1007/s10551-018-3922-2>
- Independent Evaluation Group. (2019). Creating Markets to Leverage the Private Sector for Sustainable Development and Growth. World Bank. <https://doi.org/10.1596/32078>
- Nilsson, A., Ngugi, C., Barclay, E. G., Fernandez, I., Noren, J., & Dabelstein, N. (2019). Evaluation of IBON International and the CSO Partnership for Development Effectiveness (CPDE) Project. Sida Decentralised Evaluation.
- OECD. (n.d.). Making Development Co-Operation More Effective: 2019 Progress Report.
- OECD. (2003). Coordination Capacity in National Statistical Systems: Background Report. In OECD, *Reviews of National Policies for Education: Tertiary Education in Switzerland 2003* (pp. 16–125). OECD. <https://doi.org/10.1787/9789264103092-3-en>
- OECD. (2012). Recommendation of the Council on Principles for Public Governance of Public–Private Partnerships. OECD Publishing.

OECD. (2021). Understanding the Spillovers and Transboundary Impacts of Public Policies: Implementing the 2030 Agenda for More Resilient Societies. OECD Publishing. <https://doi.org/10.1787/862c0db7-en>

OECD. (2022). The Protection and Promotion of Civic Space: Strengthening Alignment with International Standards and Guidance. OECD. <https://doi.org/10.1787/d234e975-en>

Paris 21! (2021). 2021 Progress Report: Partnering for Better Data For Better Lives for All. Paris21!

Paris 21! (2022a). Digital Transformation of National Statistical Offices. OECD Publishing. <https://www.paris21.org/Digital-transformation-NSO>

Paris 21! (2022b). The PARIS21 Partner Report on Support to Statistics 2022: A Wake-Up Call to Finance Better Data. OECD Publishing. https://www.oecd-ilibrary.org/development/the-paris21-partner-report-on-support-to-statistics-2022_c3cfb353-en

Stacey, B. (2022). Citizen-Generated Data for SDG Reporting in Kenya. Paris21!

UN DESA. (n.d.-a). How should civil society stakeholders report their contribution to the implementation of the 2030 Agenda for Sustainable Development?

UN DESA. (n.d.-b). Micro-, Small and Medium-sized Enterprises (MSMEs) and their role in achieving the Sustainable Development Goals.

UNEP. (n.d.). Methodology for SDG-indicator 17.14.1: Mechanisms in place to enhance policy coherence for sustainable development.

World Bank. (2021). The World Bank Group's Experience with the IDA Private Sector Window An Early-Stage Assessment.

World Bank. (2022). Survey on the Implementation of the Cape Town Global Action Plan for Sustainable Development Data.

Technology (13; 3 academic papers; 10 reports)

Alfaro-Serrano, D., Balantrapu, T., Chaurey, R., Goicoechea, A., & Verhoogen, E. (2021). Interventions to promote technology adoption in firms: A systematic review. Campbell Systematic Reviews.

- ICF. (2022). Evaluation of the Scaling up Renewable Energy Program in Low-income Countries: Evaluation Report.
- International Telecommunication Union. (2022a). Economic and fiscal incentives to accelerate digital transformation—9th ITU Economic Experts Roundtable: Outcome report.
- International Telecommunication Union. (2022b). Global Connectivity Report 2022.
- International Telecommunication Union. (2022c). ICT accessibility overview: Assessment for the Africa region.
- Nkala, D., & Kim, Y. (2022). The United Nations' Role in the Evolution of Technical Cooperation among Developing Countries, South-South Cooperation, and Triangular Cooperation: Real or Cosmetic Approaches to Development Cooperation? *Asian Journal of Peacebuilding*, 10(1), 45–59.
<https://doi.org/10.18588/202205.00a264>
- OECD. (2021). Development Co-operation Report 2021: Shaping a Just Digital Transformation. OECD Publishing.
- OECD. (2022). Main Science and Technology Indicators. OECD Publishing.
- Rodriguez-Crespo, E., Billon, M., & Marco, R. (2021). Impacts of Internet Use on Trade: New Evidence for Developed and Developing Countries. *Emerging Markets Finance and Trade*, 57(10), 3017–3032.
<https://doi.org/10.1080/1540496X.2019.1676225>
- Taxonomy of Environmentally Sound Technologies. (n.d.).
- UNDP. (2021). High-level forum of directors general for development cooperation 2021: Trends, Opportunities and Challenges on the Path to Sustainable Development: Strengthening South-South Cooperation and Triangular Cooperation in the COVID-19 Context. UNDP.
- UNOSSC/UNDP. (2021). South-South Ideas: Mapping South-South Cooperation in Science, Technology and Innovation for Theory and Practice. UNDP.
- UNOSSC/UNDP. (2022). UN Fund for South-South Cooperation: Results report 2020–2021 results achieved with our partners.

Trade (21; 3 academic papers; 18 reports)

Ashraf, N., Giné, X., & Karlan, D. (2009). Finding Missing Markets (and a disturbing epilogue): Evidence from an export crop adoption and marketing intervention in Kenya. *American Journal of Agricultural Economics*, 91(4): pp. 973–990.

Basstel. (2022). Final evaluation of The Netherlands Trust Fund (NTF) Phase IV – Export Sector Competitiveness Programme. International Trade Centre.

Cirera, X., & Lakshman, R. (2017). The impact of export processing zones on employment, wages and labour conditions in developing countries: Systematic review. *Journal of Development Effectiveness*, 9(3), 344–360.

Cirera, X., Willenbockel, D., Lakshman, R. (2011). What is the evidence of the impact of tariff reductions on employment and fiscal revenue in developing countries?: A systematic review. Social Science Research Unit.

International Trade Centre. (2015). Evaluation of the Trade, Climate Change and Environment Programme. International Trade Centre.

International Trade Centre. (2016a). Evaluation of the Women and Trade Programme – Phase I. International Trade Centre.

International Trade Centre. (2016b). Final Evaluation of Sector Competitiveness & Export Diversification in The Gambia. International Trade Centre.

International Trade Centre. (2017a). Final Evaluation Promoting Intra-Regional Trade in Eastern Africa Kenya Tanzania Zambia. International Trade Centre.

International Trade Centre. (2017b). Trade for Sustainable Development (T4SD) Project (2013 – 2016). International Trade Centre.

International Trade Center. (2018a). Evaluation of the Certified Trade Advisers Programme (CTAP). International Trade Centre.

International Trade Centre. (2018b). Evaluation of the Non-tariff Measures Programme. International Trade Centre.

International Trade Centre. (2019a). Evaluation of the ITC Participation and Performance in the UN Delivering as One System. International Trade Centre.

International Trade Centre. (2019b). Evaluation of the ITC Trade Facilitation Programme. International Trade Centre.

International Trade Centre. (2021a). Evaluation of ITC's Performance in Trade and Market Information. International Trade Centre.

International Trade Centre. (2021b). Sustainability review of the project 'Improving the competitiveness of Senegalese mangos' Results and lessons learned three years after project completion. International Trade Centre.

International Trade Centre. (2021c). Sustainability review of the project 'Integration of horticulture supply/value chains into tourism, Tanzania' Results and lessons learned three years after project completion. International Trade Centre.

International Trade Centre. (2022a). Evaluation of the ITC SheTrades Initiative. International Trade Centre.

International Trade Centre. (2022b). Midterm Evaluation of the ITC MENATEX Programme. International Trade Centre.

KPMG & FACTS. (2016). Midterm Evaluation of Pashmina Enhancement and Trade Support (PETS) project, Nepal. KPMG & FACTS.

McCorriston, S., Hemming, D., Godwin, J., Osborn, J., Parr, M., & Roberts, A. (2013) What is the evidence of the impact of agricultural trade liberalisation on food security in developing countries? A systematic review. (2013). https://www.researchgate.net/publication/237973599_What_is_the_evidence_of_the_impact_of_agricultural_trade_liberalisation_on_food_security_in_developing_countries#fullTextFileContent

UN Office for the South-South Cooperation. (2021). The Importance of South-South Cooperation in Strengthening Global South Trade, Investments and Regional Integration: A Contextual Overview. UNDP.

Annex D. SDG-17 Interventions and Outcomes

Interventions and Outcomes
Finance
Aid dependency
Blended finance
Budget deficit
Budget support
Capital flow
Carbon finance
Clean energy finance
Climate finance
Concessional finance
Debt as a percentage of GDP
Debt finance
Debt relief
Debt restructure
Debt service
Debt sustainability
Direct budget support
Domestic finance
Domestic resource mobilization
Domestic revenue
Domestic tax
Environmental finance

Environmental tax
Export value added tax
External debt
Finance corporation
Government deficit
Government revenue
Green finance
Guarantee agency
Guarantee instrument
Highly indebted poor country
HIPC
Income tax
International monetary
Investment guarantee agency
Investment promotion
Migrant income
Multi-lateral investment guarantee
Multilateral investment guarantee
National budget support

Technology

Blockchain
Broadband
Broadband access
Communication technology
Digital bank account
Digital divide
Digital technology
Digital transformation
Digital wallets
Energy technology
Hotspot
ICT
Information and communication technology
Information technology
Internet access
Internet use
Laptop
Mobile internet access
Mobile money

Mobile phone

MPESA

Phone

Smart phone

Smartphone

Social media

Solar device

Solar panel

Solar power

Solar system

Tablet

Technology access

Technology bank

Technology diffusion

Technology innovation

WhatsApp

Capacity Building

Capacity building

Capacity development

Capacity strengthening

Capacity-building intervention

Implementation management

National capacity

National planning

Public sector training

SDG planning

Technical assistance to government

Training of government

Trade

Duty-free

Export competitiveness

Export market

Export processing zone

Export promotion

Export sector

Export subsidies

Export subsidy

Export value

Foreign Direct Investment

Free Trade Agreement

Free trade union

Green procurement rules

Import

International trade

Intraregional trade

Non-tariff barrier

Preferential rules of origin

Quantity restriction

Quota free

Regional integration

Regional trade

Rules based trade

Tariff

Tariff average

Tariff reduction

Trade barrier

Trade liberalization

Trade quota

Trade reform

Trade support

Weighted average tariff

Weighted tariff average

Systemic Issues

Civil society organization

Country ownership

Data monitoring accountability

Development co-operation

Development cooperation

Development impact bond

Economic co-operation

Economic cooperation

Global coordination

Global partnership

Impact investment

Institutional coherence

Matching grant

Multi stakeholder

Multi-stakeholder

Multiple stakeholders
National statistical office
National statistical plan
National statistical system
North-south
Performance based financing
Performance-based financing
Policy coherence
Public procurement
Public-private partnership
SDG policy
South-South
Statistical legislation
Sustainable Development Goal policy
Triangular cooperation

Annex E. Overview of SDG-17 Data

Target/ Indicator	Time span	Published by	Original source	Link to source	Category
Target 17.1: Mobilize resources to improve domestic revenue collection					
Indicator 17.1.1 is total government revenue as a proportion of GDP.	1972 – 2020	World Development Indicators - World Bank (2022.05.26)	Demographic and Health Surveys, Multiple Indicator Cluster Surveys, Household surveys, UN Population Division	https://datacatalog.worldbank.org/search/dataset/0037712/World-Development-Indicators	Finance
Indicator 17.1.2 is the proportion of domestic budget funded by domestic taxes.	1972 – 2020	World Development Indicators - World Bank (2022.05.26)	Demographic and Health Surveys, Multiple Indicator Cluster Surveys, Household surveys, UN Population Division	https://datacatalog.worldbank.org/search/dataset/0037712/World-Development-Indicators	Finance
Target 17.2: Implement all development assistance commitments					
Indicator 17.2.1 is net official development assistance, as a proportion of the Organization for Economic Cooperation and Development (OECD) Development Assistance	2000 – 2017	UN Sustainable Development Goals - United Nations (2022-07-07)	Organisation for Economic Co-operation and Development	https://sdgs.un.org/goals	Finance

Committee donors' gross national income (GNI).

Target 17.3: Mobilize financial resources for developing countries

Indicator 17.3.1 is foreign direct investment (FDI), official development assistance and South-South cooperation as a proportion of total domestic budget.	1970 – 2020	World Development Indicators - World Bank (2022.05.26)	Demographic and Health Surveys, Multiple Indicator Cluster Surveys, Household surveys, UN Population Division	https://datacatalog.worldbank.org/search/dataset/0037712/World-Development-Indicators	Finance
Indicator 17.3.2 is the volume of remittances (in United States dollars) as a proportion of total GDP.	1972 – 2018	World Bank based on the International Monetary Fund, World Bank and OECD		http://data.worldbank.org/data-catalog/world-development-indicators	Finance

Target 17.4: Assist developing countries in attaining debt sustainability

Indicator 17.4.1 is debt service as a proportion of exports of goods and services.	1970 – 2020	UN Sustainable Development Goals - United Nations (2022-07-07)	World Bank	https://sdgs.un.org/goals	Finance
--	-------------	--	------------	---	---------

Target 17.5: Invest in least-developed countries

Indicator 17.5.1 is the number of countries that adopt and implement investment promotion regimes for least-developed countries.	1959 – 2021	UN Sustainable Development Goals - United Nations (2022-07-07)	UN Conference on Trade and Development	https://sdgs.un.org/goals	Finance
--	-------------	--	--	---	---------

Target 17.6: Knowledge sharing and cooperation for access to science, technology and innovation

Indicator 17.6.1 is fixed Internet broadband subscriptions per 100 inhabitants.	1998 – 2020	World Development Indicators - World Bank (2022.05.26)	World Telecommunication / ICT Indicators Database - International Telecommunication Union	https://datacatalog.worldbank.org/search/dataset/0037712/world-development-indicators	Technology
---	-------------	--	---	---	------------

Target 17.7: Promote sustainable technologies to developing countries

Indicator 17.7.1 is the total amount of approved funding for developing countries to promote the development, transfer, dissemination, and diffusion of environmentally sound technologies.	2010 – 2020	United Nations Sustainable Development Goals - United Nations (2022-07-07)	UN Statistics Division	https://sdgs.un.org/goals	Technology
---	-------------	--	------------------------	---	------------

Target 17.8: Strengthen the science, technology and innovation capacity for least-developed countries

Indicator 17.8.1 is the proportion of individuals using the Internet.	1960 – 2020	World Development Indicators - World Bank (2022.05.26)	World Telecommunication / ICT Indicators Database - International Telecommunication Union	https://datacatalog.worldbank.org/search/dataset/0037712/world-development-indicators	Technology
---	-------------	--	---	---	------------

Target 17.9: Enhanced SDG capacity in developing countries

Indicator 17.9.1 is the dollar value of financial and technical assistance committed to developing countries.	1960 – 2019	World Development Indicators - World Bank (2022.05.26)	Demographic and Health Surveys, Multiple Indicator Cluster Surveys,	https://datacatalog.worldbank.org/search/dataset/0037712/world-development-indicators	Capacity building
---	-------------	--	---	---	-------------------

Household surveys,
UN Population Division
[Development-Indicators](#)

Target 17.10: Promote a universal trading system under the WTO

Indicator 17.10.1 is the worldwide weighted tariff-average.	1988 – 2020	World Development Indicators - World Bank (2022.05.26)	The World Integrated Trade Solution platform	https://datacatalog.worldbank.org/search/dataset/0037712/world-Development-Indicators	Trade
---	-------------	--	--	---	-------

Target 17.11: Increase the exports of developing countries

Indicator 17.11.1 is developing countries' and least developed countries' share of global exports.	1960 – 2020	World Development Indicators - World Bank (2022.05.26)	National accounts data - World Bank / OECD	https://datacatalog.worldbank.org/search/dataset/0037712/world-Development-Indicators	Trade
--	-------------	--	--	---	-------

Target 17.12: Remove trade barriers for least-developed countries

Indicator 17.12.1 is the average tariffs faced by developing countries, least developed countries and small island developing States.	1988 – 2020	World Development Indicators - World Bank (2022.05.26)	The World Integrated Trade Solution platform	https://datacatalog.worldbank.org/search/dataset/0037712/world-Development-Indicators	Trade
---	-------------	--	--	---	-------

Target 17.13: Enhance global macroeconomic stability

Indicator 17.13.1 is the Macroeconomic	1960 – 2020	United Nations Sustainable	International Monetary Fund (IMF)	https://sdgs.un.org/goals	Finance and trade
--	-------------	----------------------------	-----------------------------------	---	-------------------

Dashboard. Annual inflation of consumer prices.		Development Goals - United Nations (2022-07-07)			
Gross public sector debt, central government, as a proportion of GDP (%)	1995 – 2021	United Nations Sustainable Development Goals - United Nations (2022-07-07)	World Bank	https://sdgs.un.org/goals	Finance and trade
Merchandise exports as a share of GDP	1960 – 2020	Our World in Data	Our World in Data based on World Development Indicators - World Bank (2022.05.26)	https://datacatalog.worldbank.org/search/dataset/0037712/World-Development-Indicators	Finance and trade

Target 17.14: Enhance policy coherence for sustainable development

Indicator 17.14.1 is the number of countries with mechanisms in place to enhance policy coherence of sustainable development.	2020 – 2020	United Nations Sustainable Development Goals - United Nations (2022-07-07)	UN Statistics Division	https://sdgs.un.org/goals	Systemic issues
---	-------------	--	------------------------	---	-----------------

Target 17.15: Respect national leadership to implement policies for the sustainable development goals

Indicator 17.15.1 is the extent of use of country-owned results frameworks and planning tools by providers of development cooperation.	2016 – 2018	United Nations Sustainable Development Goals - United Nations	OECD and UN Development Programme	https://sdgs.un.org/goals	Systemic issues
--	-------------	---	-----------------------------------	---	-----------------

		Nations (2022-07-07)			
--	--	----------------------	--	--	--

Target 17.16: Enhance the global partnership for sustainable development

Indicator 17.16.1 is the number of countries reporting progress in multistakeholder development effectiveness monitoring frameworks.	2016 – 2018	United Nations Sustainable Development Goals - United Nations (2022-07-07)	OECD and UN Development Programme	https://sdgs.un.org/goals	Systemic issues
--	-------------	--	-----------------------------------	---	-----------------

Target 17.17: Encourage effective partnerships

Indicator 17.17.1 is the amount of US dollars committed to (a) public-private partnerships and (b) civil society partnerships.	2000 – 2020	United Nations Sustainable Development Goals - United Nations (2022-07-07)	World Bank	https://sdgs.un.org/goals	Systemic issues
--	-------------	--	------------	---	-----------------

Target 17.18: Enhance availability of reliable data

Indicator 17.18.1 is the proportion of sustainable development indicators produced at the national level with full disaggregation when relevant to the target, in accordance with the Fundamental Principles of Official Statistics.	2004 – 2017	World Bank Data on Statistical Capacity		https://data.worldbank.org/data-catalog/data-on-statistical-capacity	Systemic issues
--	-------------	---	--	---	-----------------



Indicator 17.18.2 is the number of countries that have national statistical legislation that complies with the Fundamental Principles of Official Statistics.	2019 – 2021	United Nations Sustainable Development Goals - United Nations (2022-07-07)	PARIS21 SDG Survey via United Nations Global SDG Database	https://sdgs.un.org/goals	Systemic issues
Indicator 17.18.3 is the number of countries with a national statistical plan that is fully funded and under implementation.	2019 – 2021	United Nations Sustainable Development Goals - United Nations (2022-07-07)	PARIS21 SDG Survey via United Nations Global SDG Database	https://sdgs.un.org/goals	Systemic issues
Target 17.19: Further develop measurements of progress					
Indicator 17.19.1 is the dollar value of all resources made available to strengthen statistical capacity in developing countries.	2016 – 2019	United Nations Sustainable Development Goals - United Nations (2022-07-07)	UN Statistics Division	https://sdgs.un.org/goals	Systemic issues
Population census	2004 – 2017	World Bank Data on Statistical Capacity	World Bank Microdata library. Original source: United Nations Statistical Division (UNSD), 2010 World Population and Housing Censuses Programme	https://data.worldbank.org/data-catalog/data-on-statistical-capacity	Systemic issues
Birth registration	2000 – 2021	World Development Indicators -	State of the World's Children - UNICEF	https://datacatalog.worldbank.org/search/dat	Systemic issues

		World Bank (2022.05.26)		dataset/0037712/World-Development-Indicators	
Death registration	2015 – 2019	A. Karlinsky, International Completeness of Death Registration 2015–2019 (2021)		https://github.com/akarlin/sky/death_registration	Systemic issues

* All data were downloaded from: <https://sdg-tracker.org/global-partnerships> (on 1/3/2023)

** Most data seem to have been retrieved from their original data source June–July 2022.

*** Some indicators have additional supporting charts available on the SDG–tracker website.

About the American Institutes for Research

Established in 1946, with headquarters in Arlington, Virginia, the American Institutes for Research® (AIR®) is a nonpartisan, not-for-profit organization that conducts behavioural and social science research and delivers technical assistance to solve some of the most urgent challenges in the U.S. and around the world. We advance evidence in the areas of education, health, the workforce, human services, and international development to create a better, more equitable world. The AIR family of organizations now includes IMPAQ, Maher & Maher, and Kimetrica. For more information, visit [AIR.ORG](https://www.air.org).



AIR® Headquarters

1400 Crystal Drive, 10th Floor
Arlington, VA 22202-3289
+1.202.403.5000 | [AIR.ORG](https://www.air.org)

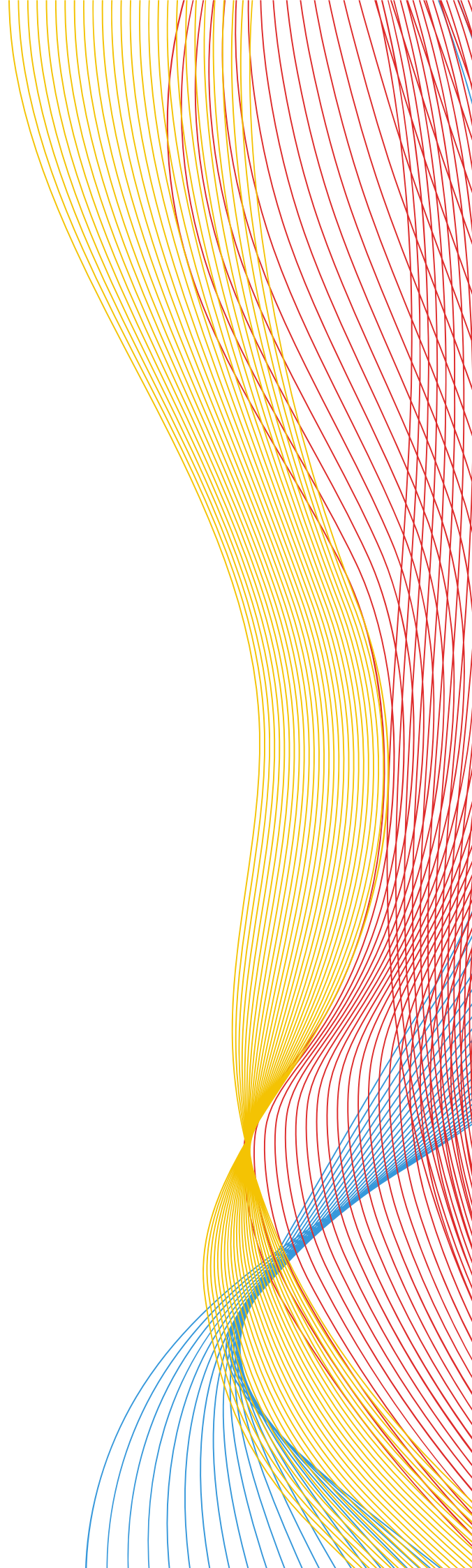


The Global
SDG Synthesis
Coalition



AIR[®]

Advancing Evidence.
Improving Lives.



17 PARTNERSHIPS
FOR THE GOALS



 sdgsynthesiscoalition.org

 @SDGSynthesis

 The Global SDG Synthesis Coalition