

Trade, Investment and Tourism Evidence Gap Map

Protocol – Preliminary

This version: October 20, 2026

Authors:

Knowledge and Learning Division (VPS/KLD)

Trade and Investment Division (PTI/TIN)

Productivity, Trade and Innovation Sector (VPS/PTI)

Table of Contents

1. Background	3
2. AI-enabled pipeline development.....	6
3. Study objectives and questions.....	7
3.1 Objectives.....	7
3.2 Research questions.....	7
3.3 Framework development and scope	7
3.4 Criteria for including or excluding studies (PICOS)	8
3.5 Other inclusion and exclusion criteria	18
4. Search strategy	18
5. Screening protocol	21
5.1 Title and abstract screening	21
5.2 Full-text retrieving.....	22
5.3 Full-text screening	22
6. Data extraction and critical appraisal.....	23
6.1 Basic study and publication information	23
6.2 Filters.....	23
6.3 Data extraction training and pilot	25
6.4 Dealing with multicomponent interventions	25
6.5 Quality checks	26
7. Analysis and reporting	26
8. Engagement and communication plan.....	26
Appendix A. Alignment with the catalogs of results.....	28
Appendix B. Scopus database search strategy	30
Appendix C. Grey literature sources and repositories	33
Appendix D. Title and abstract screening protocol	34
Appendix E. Full-text screening protocol for impact evaluations	35
Appendix F: Data extraction template.....	37
Appendix G. Title and abstract screening rules applied by the AI model..	38
Appendix H. Full-text screening and coding rules applied by the AI model	39

1. Background

1.1 The problem, condition, or issue

The IDB Group Institutional Strategy 2024-2030 highlights the importance of consolidating the institution as the knowledge bank of Latin America and the Caribbean by producing cutting-edge research on development issues in the region and serving as a bridge to development research from the rest of the world that could be useful to its members. For this issue, improving the feedback loop between operations and knowledge is of utmost importance, providing easy access to empirical evidence, insights from past projects, and tools for generating new knowledge from operations.

The Trade and Investment Division (TIN) of the IDB's Productivity, Trade and Innovation sector (PTI) operates across all 26 borrowing members of the IDB Group. To ensure impact, the operations in this division are formulated around a robust theory of change grounded in the best available knowledge, including frontier empirical evidence.

1.2 State of affairs and policy responses

TIN currently has a team of specialists dedicated to supporting its operational portfolio by providing empirical evidence to inform its operations and by participating in the corresponding impact evaluations. In addition, in 2019, the division developed a Sectoral Framework Document (SFD) that identifies the main challenges of the Latin American and Caribbean Region across five key areas for the trade and integration sector. The SFD also compiles global evidence on the effectiveness of programs and initiatives that address these problems, providing a comprehensive analytical foundation for the sector's strategic priorities. Similarly, the Climate Change and Sustainable Development Sector developed a Tourism SFD in 2022, focusing on several issues relevant to the topic.

Finally, the division is developing its indicators for the IDB's Catalog of Results Indicators, which provides a standardized set of metrics and definitions to systematically capture progress across the main strategic areas, enabling consistent diagnosis, monitoring, and evaluation of trade and integration outcomes.

1.3 Importance of developing this evidence gap map

Given the institutional mandate to promote evidence-based decision-making and to leverage the division's current efforts, the IDB signed an agreement in 2024 to co-produce two evidence gap maps with support from the International Initiative for Impact Evaluation (3ie) and IDB's Knowledge and Learning Division (KLD). Both maps were launched in 2025: one with almost 430 studies for the citizen security division and another for the transportation division, which encompasses more than 700 entries.

These maps are intended mainly for sectoral specialists and teams working in operations and knowledge-oriented tasks in headquarters and country offices. EGMs could also be helpful for country economists and their teams, development effectiveness specialists, and research teams, among others. However, everyone inside and outside the IDB Group will have access to this resource.

Building on these recent efforts, and since the Bank has now acquired the technical capacity to produce EGMs, the TIN and KLD divisions have decided to develop the first fully in-house EGM designed, created, and implemented by the IDB.

The interventions and outcomes set out in this map are highly aligned with the IDB's portfolio of trade, investment, and tourism operations, with the indicator categories used to measure their effectiveness and impact, as well as current priorities within the division and the sector. This EGM will focus on a similar type of interventions to those already captured in the existing SFD but grouped into four macro categories: (a) trade facilitation; (b) export promotion; (c) attraction of Foreign Direct Investment (FDI); and (d) tourism. Also, outcome variables will be grouped into six categories: (a) access; (b) quality; (c) affordability; (d) service management; (e) sector governance; and (f) socio-economic outcomes. This EGM will serve as an input for the 2027 Trade, Investment and Tourism Thematic Framework Documents (TFD).

1.4 Related Evidence Gap Maps

A review of existing EGMs reveals that while there is a growing body of synthesis work on private-sector development, infrastructure, and finance, no EGM explicitly and comprehensively focuses on trade, investment, and tourism interventions (Table 1). The closest existing maps cover adjacent domains but do so either from a sector-agnostic perspective or with a focus on specific instruments (e.g., infrastructure or financial services). Even the most relevant example, the SDG Partnership EGM, includes export and investment promotion only as part of a broader framework, without the granularity required to inform policy design in these areas. Tourism, in particular, is almost entirely absent as a standalone domain.

This gap highlights a clear opportunity and justification for the proposed EGM on Trade, Investment, and Tourism interventions. By systematically organizing evidence across these interconnected domains, the proposed EGM will fill a critical void in the evidence architecture, particularly by incorporating policy interventions central to development practice yet underrepresented in existing synthesis efforts—such as customs modernization, non-tariff measures, export and investment promotion agencies, and tourism competitiveness strategies. In doing so, it will not only complement existing EGMs but also provide a novel, policy-relevant framework that better aligns with the operational needs of institutions like the IDB, where trade, investment, and tourism interventions are often jointly designed and implemented.

Table 1: Description of related EGM

Title	Institution	Year	Direct relevance	What it includes	What it misses
SDG Partnership Pillar Evidence & Gap Map	AIR (SDG Synthesis Coalition funded)	2023	Closest direct overlap	Includes a trade component with trade liberalization, trade sanctions, infrastructure for trade, anti-dumping measures, FDI and export promotion	No inclusion of trade facilitation or tourism interventions, and trade and investment promotion pillars are broad
Building Stronger Transport Policy: An Evidence Gap Map	IDB + 3ie	2025	Indirect	Transport infrastructure, regional logistics , transport regulation and policy	No explicit trade, investment or tourism interventions
Evidence and Gap Map: Transport Interventions in Low- and Middle-Income Countries	Campbell + CEDIL	2022	Indirect	Roads, ports, shipping, aviation, divided into infrastructure, incentives, and institutions	No explicit trade, investment or tourism interventions
Private Sector Engagement (PSE) Interventions EGM	3ie (GIZ funded)	In progress	Indirect	Private sector interventions: finance of and with private sector actors, capacity development, public-private partnerships, matchmaking and consulting	Broader focus. No clear structure around trade, foreign investment, or tourism interventions.

2. AI-enabled pipeline development

Screening, full-text retrieval, coding, and data extraction will be conducted through an AI-enabled pipeline developed in-house. The pipeline consists of purpose-built scripts that apply the eligibility criteria of this protocol through structured, rule-based prompts submitted to large language models (LLMs). Each prompt asks the model to follow the same sequence of decisions as a trained reviewer, so that every exclusion can be traced to a specific criterion.

Human reviewers will remain responsible for defining the taxonomy and prompts, constructing the gold-standard samples used to validate model performance, reviewing samples of model decisions, and resolving ambiguous or failed cases. The pipeline is designed to be re-run at a predefined frequency, so that newly identified studies enter the same screening, coding, and extraction procedures and the map can be maintained as a living evidence resource.

2.1 Gold standard-standard samples

Model performance at each stage will be benchmarked against gold-standard samples of studies screened and coded manually by trained reviewers. The manual screening of these samples will follow the decision guides in Appendix D (title and abstract) and Appendix E (full text), which reflect the eligibility criteria of this protocol. The gold standards will be kept fixed during prompt development, so that successive prompt versions, and any future change in the underlying model, can be compared against the same benchmark.

For the title and abstract stage, the gold standard will consist of a random sample of records stratified by study type (impact evaluations, systematic reviews, and grey literature). Trained reviewers will screen the records independently and in duplicate, working in pairs across batches, and will classify each record as “include,” “unsure,” or “exclude,” recording the reason for any exclusion. Inter-reviewer reliability will be reported using both observed agreement and a chance-corrected measure. All discrepancies, including “unsure” ratings and differences in exclusion reasons, will be resolved through structured discussion, and the reconciled consensus decision (include or exclude) will constitute the gold standard.

For the full-text stage, the gold standard will consist of a sample of retrieved studies selected to cover the range of interventions and outcomes in the framework, rather than drawn at random, so that less frequent categories are also represented. Two trained coders will independently screen each study and code its interventions, outcomes, and study designs following Appendix E and the framework in Section 3, and their decisions will be reconciled into a single gold-standard classification. The same sample will also be used to check the variables produced by the extraction script before the full run.

3. Study objectives and questions

3.1 Objectives

Provide state-of-the-art empirical evidence to support the operations of the IDB's Trade and Investment Division and its knowledge agenda, promoting strategic selectivity and strengthening impact.

3.2 Research questions

Two general questions should be addressed to cover the methodological purpose of an EGM:

- What are the main concentrations and gaps in causal evidence in the trade and investment literature, focusing on the types of interventions where the IDB provides loan operations?
- What are the main characteristics of the empirical evidence on the effects of interventions in this sector in terms of geographic distribution, time, and methods used to evaluate the effectiveness of these interventions?

In addition, to integrate the use of EGMs into the core of operations at the IDB, the following additional question is proposed:

- To what extent does the existing empirical evidence align with the strategic priorities and intervention types identified by the Trade and Investment Division, and the Productivity, Trade and Innovation Sector?

Together, these questions aim to both map the current state of causal evidence and ensure its operational relevance for decision-making.

3.3 Framework development and scope

The Intervention-Outcome (I-O) framework from this EGM leverages existing knowledge from

- i. The 2019 Trade and Integration Sectoral Framework Document (SFD)
- ii. The 2022 Tourism SFD
- iii. The portfolio of IDB loan operations from INT and TIN since 2005, and from other areas, when related to tourism.
- iv. SPD's Catalog of Results Indicators for Trade and Investment, and Tourism.
- v. The flagship publication *Out of the Border Labyrinth: An Assessment of Trade Facilitation Initiatives in Latin America and the Caribbean* (Volpe, 2016).
- vi. Current division's strategic priorities.

Aligning with these sources is essential to facilitate navigation within the EGM, enhance usability, and ensure consistency with the IDB's taxonomy and operations.

Note: This map covers trade, investment, and tourism interventions and does not include interventions that fall under the scope of other IDB–3ie evidence gap maps. In particular, interventions related to domestic transport, road and railway infrastructure, and ports belong to the Transportation EGM and are therefore excluded here, even when their effects are measured in terms of trade flows or trade costs.

3.4 Criteria for including or excluding studies (PICOS)

Population (P)

Population defines who or what the interventions are intended to affect. It specifies the target group, setting, or context to which the evidence applies. Clearly defining the population helps ensure that the EGM remains focused and that included studies are comparable across contexts.

This EGM will include evidence for all countries and all units of analysis. Depending on available resources, the set of countries in this stage may be smaller in the full-text screening or data extraction stages.

Interventions (I)

Refers to the specific action, program, policy, or initiative whose effectiveness is being assessed. It represents what is implemented or changed to influence outcomes in the target population.

We will group interventions into four main categories: (1) trade facilitation initiatives; (2) export promotion; (3) attraction of Foreign Direct Investment (FDI); and (4) tourism. Table 2 outlines the specific interventions within each category and provides definitions for each.

Note that, technically, there are no standalone interventions for transversal topics such as institutional capacity, digital transformation, or skills development. Instead, interventions addressing these themes are embedded throughout the framework and classified by their primary program, goal, or objective. For example, a program focused on training firms or individuals to improve their abilities to access new markets or destinations would be classified under “Export promotion policies”. To address and visualize these areas, we propose adding a tag for studies that cover one or more of these transversal categories. Information for these tags can be extracted during the data extraction stage.

Finally, studies relying solely on aggregate measures of trade openness, such as the ratio of exports plus imports to GDP, broad trade liberalization indices, or average tariff levels are excluded from the map, unless they evaluate a discrete, identifiable intervention that can be meaningfully assigned to a category within the framework. Similarly, studies that exploit changes in product-level tariffs should only be included if they are linked to an intervention category within the framework. For example, if changes in trade openness or tariffs are due to a trade liberalization episode or a specific trade policy, the studies should be included. On the contrary, if these changes are just changes observed in a dataset without a clear policy behind them, the studies should be excluded.

Table 2: Intervention framework

Trade facilitation	
Intervention	Description
Trade agreements and tariff policy	Policy actions that change tariffs and the rules of market access between countries, whether through negotiated agreements or unilateral measures. This category includes bilateral, regional, plurilateral and multilateral agreements such as free trade agreements, customs unions, common markets, economic partnership agreements, accession to the WTO, preferential Trade Agreements, Economic Union, Complete Economic Integration, Agreements and international agreements on trade facilitation such as the World Trade Organization (WTO) Trade Facilitation Agreement. It also covers non-reciprocal preferential schemes such as GSP, AGOA or Everything But Arms, and unilateral tariff changes, trade restrictions and trade wars, including retaliatory tariffs, import bans and quotas. Studies are classified here when the intervention evaluated is the agreement or the tariff measure itself (its signature, entry into force, depth, specific provisions, or the imposition or removal of tariffs and restrictions), whether the outcomes concern tariffs, non-tariff measures, rules of origin or trade flows.
Regulatory and institutional trade facilitation reforms	Domestic regulatory and institutional reforms that reduce border and trade frictions through changes in legal frameworks, procedural simplification, governance arrangements and inter-agency coordination. This category includes the harmonization of national rules and standards, simplification of customs and border procedures, changes in documentary requirements, and organizational reforms of border agencies. It focuses on the legal and institutional environment of trade rather than on digital tools or the technological automation of customs functions. Domestic reforms carried out to implement a trade agreement are classified here when the study evaluates the specific reform and not the agreement as a whole. Trade adjustment assistance and compensation programs are also classified here, as policy responses to the distributional consequences of trade policy changes.
Border management and infrastructure	Investments and operational programs that aim to improve the efficiency and capacity of cargo and passenger border crossings through physical infrastructure development, layout redesign, and on-site operational enhancements. This category also includes canals and other artificial waterways; however, ports are excluded as they are part of the IDB-3ie Transportation EGM.
Digitalization and automation of customs and border management	Interventions that implement digital systems, electronic platforms, and automated technologies to modernize customs and border management processes. These interventions improve enforcement, transparency, and efficiency through the use of information technology, data analytics, and automation. This category includes both standalone customs digitalization and integrated digital platforms connecting multiple border agencies.
Digital trade and data regulation	Regulatory frameworks governing the digital aspects of trade, including rules on data flows, digital transactions, and the legal environment enabling cross-border digital commerce.
Export promotion	

Export promotion policies	Programs that support firm internationalization and export development, including strategic design, modernization of export promotion organizations, and implementation of export support instruments. Programs can include a combination of market development and promotional activities, and capability building for exporting, including digital skills for internationalization when that is the main objective.
Export market linkages	Interventions that connect firms to buyers, markets, and networks, improving market information and commercial relationships, including only digital platforms that connect domestic exporters with foreign buyers or commercial partners, facilitating transactions, partnerships, and market entry.
Trade finance instruments for exporters	Mechanisms and programs designed to expand access to financial instruments that support trade and export activities, particularly for small and medium-sized enterprises (SMEs) and underserved market segments.
Quality, standards, and export compliance support	Policies and programs that strengthen quality infrastructure and regulatory frameworks at both the institutional and firm levels to support compliance with international standards, certification, and conformity assessment requirements. This category includes interventions that build the capacity of public institutions (such as standards bodies, accreditation agencies, testing and inspection authorities) as well as private firms to meet international quality, safety, and technical standards. Examples include alignment with international standards, accreditation of testing bodies, modernization of technical regulations, and training or technical assistance for firms to comply with export requirements and access international markets.
Attraction of Foreign Direct Investment (FDI)	
Investment promotion and facilitation programs	National or regional strategies and programs that attract, facilitate, and retain domestic or foreign investment. These policies often include strengthening investment promotion agencies, investor outreach, site selection support, regulatory guidance, aftercare services, digitalization of investor services, and capacity-building and skills programs for investment promotion. It also covers investment agreements, including bilateral investment treaties, investment chapters in broader economic agreements, investment facilitation agreements such as the World Trade Organization (WTO) Investment Facilitation for Development Agreement, and the investor protection and dispute settlement provisions they contain. They seek to improve the investment climate, foster strategic sector growth, and promote linkages between foreign investors and local firms.
Investment incentives and special regimes	Instruments that provide targeted advantages to attract or anchor foreign investment in specific locations, sectors, or activities. Includes Tax incentives for foreign investors, Investment subsidies or grants, Special Economic Zones (SEZs) and free zones, Priority investment regimes, Greenfield investment incentive packages
Tourism	
Tourism firm competitiveness and service quality	Policies and programs that strengthen the competitiveness, productivity, and export capacity of tourism firms by improving service quality, business practices, and operational performance. These interventions target tourism enterprises—particularly MSMEs—across the tourism value chain and support firm upgrading through skills development, quality standards, digital capabilities, access to finance, innovation, and integration into domestic and international tourism markets.

Tourism market development and destination management	Interventions that promote tourism demand and strengthen the strategic management, governance, and marketing of tourism destinations. These initiatives aim to attract foreign visitors, enhance destination visibility and branding, and improve the coordination and operational management of tourism through institutional arrangements, data-driven tools, and digital platforms. The category encompasses: (i) destination marketing campaigns and branding strategies, including behavioral interventions aimed at shaping traveler perceptions and decisions; (ii) smart destination management programs and visitor intelligence systems; (iii) capacity building programs targeting local residents and service providers in the tourism sector; (iv) the regulation and management of short-term rental and gig economy platforms; and (v) the institutional arrangements required to govern and sustain these programs.
Investment in tourism assets and destinations	Policies and investments that support the development, upgrading, and sustainability of tourism-related assets and destinations, enabling private investment and enhancing long-term attractiveness for international tourism markets. These interventions focus on physical, natural, and cultural tourism assets, as well as investment frameworks, incentives, and resilience features that directly affect tourism supply and destination viability. The category includes: (i) tourism infrastructure investment and upgrades to physical and natural assets; (ii) investment promotion strategies targeting the tourism sector, including FDI attraction efforts, investment climate reforms, and the work of investment promotion agencies oriented toward tourism; (iii) financial incentives for tourism investment such as tax incentives, hotel investment schemes, and special tourism zones; (iv) sustainability and resilience measures that protect the long-term viability of tourism destinations; and (v) institutional frameworks that govern tourism investment and asset management.

Comparators (C)

Refers to the benchmark or reference condition against which the intervention's effects are assessed. In impact evaluations, this typically represents the counterfactual: what would have happened to the same population in the absence of an intervention, program, or policy.

Comparators can include no intervention, an alternative intervention, or a different level or intensity of the same program. We will include any type of comparator but exclude studies that rely solely on simulation or forecasting models, ex ante impact assessments, or scenario analyses.

Outcomes (O)

Refers to the specific variables or effects that the intervention aims to influence and that are measured in the included studies.

Our proposal includes grouping outcomes into six main categories: (a) access; (b) quality; (c) affordability; (d) service management; (e) sector governance; and (f) other socio-economic outcomes. Table 3 outlines the definition of outcomes for each category. All these outcomes are aligned with the Catalogs of Results Indicators for both Trade and Investment, and Tourism (See *Appendix A* for this alignment).

To avoid excessive proliferation of indicators, some thematic areas are intentionally grouped at a higher level of aggregation. For instance, labor

market-related outcomes are consolidated under a single outcome category, rather than disaggregated to the level of detail that would be used in the Catalog (e.g., wages, employment, and formalization separately). This approach prioritizes parsimony and comparability across sectors while capturing the core dimensions of labor market effects relevant to the analysis.

Table 3: Outcome framework

Access	
Outcomes	Description
Extensive margin of trade	Refers to the expansion in the number of firms, products, or markets (e.g., countries of origin or destination) participating in international trade in goods and non-tourism services, both exports and imports. Examples include increases in the number of exporting or importing firms, diversification of traded products, or the entry into new foreign markets or supplier countries. Includes trade in goods and services.
Intensive margin of trade	Measures the deepening of trade activity among existing participants in international trade in goods and non-tourism services, including higher export or import volumes/quantities, values, or shares per firm, product, or destination. It captures how much existing traders exchange rather than how many participate. Includes trade in goods and services.
Foreign Direct Investment outcomes	Includes inflows and outflows of foreign direct investment, as well as the number, size, or value of foreign investment projects established. May also include measures of reinvested earnings and investor retention, FDI leads, leads converted into actual investments, or firms established in a country by foreign investors.
Visitor flows and engagement	Measures the volume and composition of domestic and international visitor flows entering a destination or tourism zone, reflecting demand scale and attractiveness. It also captures the depth of visitor engagement and occupancy rates, including average duration of stay, repeat visits, and the diversity of destinations, attractions, or activities experienced during trips. This category also includes willingness and intention to travel or to visit.
Quality	
Outcomes	Description
Service quality	Refers to the objective quality, reliability, safety, and performance of trade and tourism-related services and infrastructure. Includes measurable, system-level indicators such as adherence to service standards, certifications in quality, frequency of disruptions or delays, accident rates in tourism infrastructure, cargo losses, service availability, continuity, interoperability of platforms, coordinated operating hours, and system resilience. This outcome captures whether improvements in efficiency and scale are achieved without compromising operational performance, safety, or reliability, based on observable and verifiable metrics rather than perceptions.
User satisfaction	Assesses users' or visitors' perceptions and subjective evaluations of their overall experience with services or programs related to tourism, trade, or investment, including satisfaction, perceived safety, comfort, treatment by service providers, and overall experience quality. This category refers to satisfaction with the experience or service received as a whole, rather than satisfaction with specific intervention components or tools (e.g., individual platforms, procedures, or instruments). These outcomes are typically measured through surveys, interviews, ratings, reviews, or stated preferences, and may influence behavioral responses such as likelihood to return, recommend, or continue using a service. This

	category reflects perceived value and experience and excludes objective measures of system performance or service reliability.
Affordability	
Outcomes	Description
Travel time and costs frictions	Measures time and cost frictions affecting trade flows along the supply chain and from the initial to the final destination, including physical transport and logistics costs (travel time, freight costs, corridor expenses) as well as administrative and regulatory costs at the border, such as customs clearance time, inspections, documentation requirements, and compliance costs. Can also include costs associated to investment flows.
Expenditure in tourism activities	Captures the average expenditure per visitor or per trip on tourism-related activities, such as accommodation, transport, food, and cultural or recreational services. This outcome reflects tourism demand and local value generation, measured through realized spending, and is distinct from price levels, inflation, or cost-of-living effects on residents.
Prices and inflation	Measures changes in price levels and inflationary pressures affecting goods, services, and assets, including traded and non-traded items. This includes domestic prices of specific goods or services, import–export price pass-through, overall inflation, and cost-of-living effects in local areas, such as prices of food, energy, hospitality services, land and housing, and rental markets. Indicators may include measures of affordability, such as expenditure shares or price-to-income ratios. This outcome captures prices, rather than income distribution or inequality outcomes.
Service management	
Outcomes	Description
Firm entry, exit and formality	Measures changes in firm dynamics, including entry, exit, survival, and formalization rates. Captures entrepreneurship, business turnover, and firm growth or contraction patterns. It also reflects the extent to which policies encourage the transition of firms from the informal to the formal sector, signaling improvements in the business environment and productive renewal.
Income, sales and profits	Reflects economic performance and financial outcomes at the firm, sector, regional, or national level. Includes firm-level measures such as sales, revenues, profitability, or income growth, as well as aggregate indicators such as GDP, GNP, or satellite accounts capturing trade- or tourism-related value added. This outcome captures levels of income generation, rather than efficiency, prices, or distributional effects.
Access to finance and capital accumulation	Measures changes in access to finance and capital accumulation that support trade, tourism supply, and productive activities. Includes improvements in access to credit and financial markets and instruments, as well as physical and financial capital formation, such as investment in machinery, equipment, infrastructure, and capital market development. Indicators may capture financing conditions and the realized accumulation of capital stocks.
Labor market outcomes	Captures effects on labor market conditions and job characteristics, including total and formal employment, wages, job creation or destruction, labor productivity, and occupational mobility linked to trade, tourism supply, or investment activities. It may also include changes in labor informality, seasonal or temporary employment, job openings, or workforce participation. When relevant, this category includes gender-differentiated labor market outcomes and gender gaps in labor indicators.

Access to better inputs and services	Refers to firms' ability to obtain higher-quality, cheaper, or more reliable production inputs — such as raw materials, components, intermediate goods, or services — as a result of improved trade integration, tourism, investment, and explicit integration into global and regional value chains (GVCs). This category captures outcomes related to firms' participation in GVCs through upstream and downstream linkages, including access to imported or domestically supplied intermediate inputs, supplier diversification, enhanced supply chain efficiency, improved logistics and infrastructure, and reduced input constraints.
Technology Adoption and Digitization	Captures the uptake and use of new technologies, digital tools, or systems by firms or institutions, including the digitization of administrative and operational processes. This category includes the adoption of automation, digital platforms, e-government systems, advanced logistics or information systems, and quality management technologies. It also covers institutional and administrative digitization outcomes, such as the digitalization of procedures, removal of in-person requirements, or implementation of online service delivery systems. These changes may act as intermediate institutional outcomes and can lead to downstream effects on efficiency, processing times, costs, transparency, or access, which are captured separately under outcome categories.
Productivity and competitiveness	Measures improvements in the efficiency of production or service delivery, typically captured through higher output per unit of input (labor, capital, or total factor productivity). It can include sophistication of exports, TFP, labor productivity, measures of quality at the firm-product level, or econometric estimates of productivity.
Innovation	Includes the development or introduction of new or significantly improved products, processes, or business models. It includes R&D investment, patenting activity, product differentiation, and collaboration with research institutions. May also include other measures of product, process or organizational innovations.
Sector governance	
Outcomes	Description
Fiscal revenue and tax compliance	Captures the impact of trade, investment, and tourism-related activities on public revenue generation and fiscal performance. Includes changes in tax collection, customs revenue, tax bases, compliance rates, and fiscal balance.
Market concentration	Refers to changes in the structure, competitiveness, and contestability of markets. Includes outcomes such as market share concentration, firm-size distribution, entry barriers or indices such as the Herfindahl-Hirschmann or measures of mark-ups/mark-downs. It also captures the presence or reduction of monopolies, oligopolistic dominance, collusion, and cartels.
Reforms or approved agreements	Legal, regulatory, or institutional reforms formally approved to strengthen trade, tourism or investment governance, improve coordination, enhance transparency, or modernize sector regulation. Captures institutional or policy-level changes related to trade or investment, such as the number or scope of preferential trade agreements, bilateral investment treaties, or implemented regulatory reforms.
Socio-economic outcomes	
Outcomes	Description
Household welfare and poverty	Outcomes related to household material well-being, including changes in income, expenditure, or poverty status. This category includes absolute and relative measures of economic resources and living standards, as well as poverty and extreme poverty rates. Indicators may capture both average household conditions and disparities among households, but exclude aggregate macroeconomic growth measures.

Inequality and segregation	Outcomes related to inequality, segregation, and disparities across population groups, capturing differences in economic, social, or spatial outcomes rather than average levels. This category includes gaps in income, wealth, education, health, employment, or access to opportunities between groups defined by income, gender, ethnicity, location, or other relevant characteristics. Indicators may include measures such as the Gini index, income or wage gaps, differential employment rates, education or literacy gaps, life expectancy differences, residential or spatial segregation, measures of gentrification (different from changes in housing prices), and unequal access to services or economic opportunities.
Environmental quality and emissions	Includes outcomes related to air quality (e.g., PM2.5, PM10, NO ₂), noise levels (e.g., dB), greenhouse gas emissions (e.g., CO ₂ , O ₃), and other measurable environmental indicators, such as water quality, waste generation, ecosystem degradation or restoration (e.g., coral reef health, coastal or terrestrial ecosystem condition), biodiversity indicators, and land-use or habitat changes.
Skills	Outcomes related to the development, acquisition, or strengthening of individual skills and competencies, capturing changes in human capital that may affect individuals' economic and social opportunities. This category includes indicators such as completion of upskilling or reskilling programs, educational attainment and progression, certifications or credentials obtained, among others. Outcomes may reflect both participation and achievement in formal education, vocational training, or other capacity-building activities.

Study designs (S)

Specifies the types of studies that are eligible for inclusion based on their methodological rigor and ability to establish causal inference. Given the IDB's focus on impact and evidence-based decision-making, EGMs prioritize studies that employ experimental or quasi-experimental designs.

We propose to include only quantitative effectiveness literature, focused on impact evaluations and systematic reviews that use attributional, causal designs to evaluate the effects of a clearly defined development intervention delivered in a real-world setting.

We will only include studies that implement at least one of the following study designs widely used to evaluate intervention effectiveness (Gertler et al., 2016):

1. Prospective studies that allocate participants to treatment and control groups using random assignment or experimental methods:
 - i. Randomized controlled trials (RCTs), with assignment at individual, household, community, or other cluster level, and quasi-RCTs using prospective methods of assignment (such as alternation).
 - ii. Natural experiments with clearly defined intervention and comparison groups, which exploit natural randomness in implementation assignment by decision makers (e.g., public lottery) or random errors in implementation.
2. Quasi-experimental designs where treatment arms are created without random assignment:
 - Regression discontinuity designs (RDD), either sharp or fuzzy designs, and other derived methods (i.e., kink RDD, differences in discontinuity). Instrumental variables (IV), either estimated with 2SLS, 3SLS, SUR, or GMM techniques. This category may include Arellano-Bond and

Arellano-Bover estimators if they are included as instruments in the econometric specification.

- Selection models, or models that correct for selection bias or endogeneity by explicitly modeling the selection process through a latent variable structure. Includes the Heckman two-step, endogenous switching regression, endogenous treatment-effects models, bivariate probit with selection, and the Roy model.
- Difference-in-differences (DID) and fixed-effects (FE), methods that exploit variation within units over time or across groups to control for unobserved heterogeneity. Includes difference-in-differences (DiD) and its extensions (staggered adoption DiD, stacked DiD, event studies), two-way fixed effects (TWFE), high-dimensional fixed effects, two-way Mundlak regressions (TWM), correlated random effects, and within-estimators. These methods identify causal effects by comparing changes over time or across treatment and control groups, controlling for time-invariant unobserved confounders.

Note: The following methods are not classified as fixed-effects: One-way fixed effects only, or year/time dummies alone (no cross-sectional unit FE); DEA, stochastic frontier analysis, or Malmquist index (efficiency methods), even if some fixed effects appear these are not a two-way fixed-effects treatment design by themselves; cross-sectional regressions with dummies and no panel/time variation; "Before-and-after" comparisons with no control group; gravity models without solid two-way fixed effects.

- Synthetic control methods, including their extensions: synthetic differences in differences, and generalized or augmented synthetic control methods.
- Weighting and matching approaches, that control for observable confounding by constructing comparable treatment and control groups based on observed characteristics. Includes non-parametric approaches such as covariate matching, coarsened exact matching, propensity score matching, kernel and radius matching, and Mahalanobis distance matching. Also includes parametric and semi-parametric approaches such as inverse probability weighting (IPW), entropy balancing, regression adjustment, and doubly robust estimators.
- Interrupted time series (ITS) models that estimate the impact of an intervention by comparing the level and trend of an outcome before and after the intervention point. Includes segmented regression, ARIMA-based ITS, and controlled ITS with a contemporaneous comparison group. A minimum of three pre-intervention time periods is required for inclusion.

In panel datasets, additional estimation strategies are often employed to address time dynamics, autocorrelation, and endogeneity, particularly when outcomes are persistent over time. These strategies include random effects models, feasible generalized least squares, and dynamic panel estimations. While these methods support inference, they only yield causal effects when

combined with exogenous variation, valid instruments, or a robust identification strategy. The same applies to some gravity models. Therefore, unless there is a clear identification strategy given by one of the methodologies above, these studies should be excluded (but not necessarily in the title and abstract screening stage).

Gravity models estimated using OLS, PPML, or any other estimator are eligible for inclusion provided that they incorporate a clear causal identification strategy drawn from the qualifying methods listed above. The estimator itself (PPML to handle zero trade flows and heteroscedasticity) does not constitute an identification strategy. What matters is the source of exogenous variation used to identify the causal effect of the intervention. Gravity models that rely solely on cross-sectional variation, time-invariant distance measures, or aggregate trade openness indices without an explicit identification strategy should be excluded at the full-text screening stage.

The same rationale should apply to other theoretical or emerging methods, such as machine learning, Bayesian estimation, or simulation-based methods, or papers proposing general equilibrium models with calibration techniques and without a clear identification strategy.

Observational studies, evaluations, and case studies that do not meet the methodological conditions described above, such as before-after studies without a comparison group or cross-sectional studies that do not adequately address selection bias or confounding, will not be included.

A systematic review is a synthesis of research evidence on a particular topic, obtained through an exhaustive and transparent search across multiple academic databases and other relevant sources. The search process is systematic and reproducible, often including studies in different languages to avoid bias, while maintaining a clear focus on a specific intervention and pre-defined outcomes of interest. Systematic reviews apply widely accepted scientific strategies to minimize bias at every stage, and indeed they assess the quality and reliability of included studies by classifying them according to their design and the credibility of their findings.¹

Includable systematic reviews can implement meta-analysis, meta-regressions, and narrative and descriptive analyses. Meta-analysis are studies that quantitatively pool estimates from multiple primary studies using statistical methods and they include standard meta-analysis, meta-regression, and network meta-analysis. On the other hand, narrative and descriptive systematic reviews are studies that synthesize evidence from multiple primary studies without statistically pooling estimates. Finally, review studies that do not fit the above categories, includes bibliometric analyses, vote-counting reviews, and mixed-method reviews that combine qualitative and quantitative synthesis in ways that do not constitute a formal meta-analysis.

¹ Critical appraisal of the included systematic reviews will not be conducted in the first stage of the map. It may be carried out in later stages or updates, if required

The reviews included on this map go beyond a mere “state of the art” synthesis, since they aim to evaluate the effectiveness of interventions and to estimate their causal impact on relevant outcomes. Reviews that have included study designs or methods not eligible for this map will be retained if at least one eligible study design is included and reports results for at least one relevant intervention and one relevant outcome. In addition, systematic reviews do not need to include a meta-analysis to be included in the map, since meta-analysis is often unsuitable when interventions are highly heterogeneous².

Finally, we will exclude the following study types: qualitative studies, feasibility studies, acceptability studies, and studies examining willingness to pay for goods, services, processes, and business models. We acknowledge that the study types excluded from this map may contain valuable information; however, the focus of this EGM is to map existing rigorous evidence of intervention effectiveness.

3.5 Other inclusion and exclusion criteria

Language

Studies published in English, Spanish, Portuguese or French will be included, although the search terms will be in English. For grey literature, the four languages will be considered for the search if the institutions of interest include language filters.

Publication date

Studies should be included if they were published in 2005 or later. Given the field's evolution, the late 1990s and early 2000s are widely recognized as a turning point in economics, commonly referred to as the “Credibility Revolution.” During this period, there was a growing emphasis on improving the reliability of empirical research by adopting more rigorous designs and increasing the use of experimental and quasi-experimental methods to evaluate policies and interventions. Given this, the 2005-or-later threshold has a low likelihood of missing eligible studies while also limiting the overall breadth of the evidence mapping project, ensuring the exercise remains manageable and within our current resource constraints.

4. Search strategy

The EGM will employ a systematic search strategy, following best practices for literature searching in evidence synthesis as outlined by MacDonald et al. (2024). The search strategy is designed to minimize publication bias by covering both academic bibliographic databases and relevant grey literature sources, including the websites and repositories of international organizations and development agencies.

Electronic searches will be conducted in Scopus, a large multidisciplinary abstract and citation database that provides broad coverage of peer-reviewed journals. The search strategy combines terms related to trade, investment, and tourism policies and programs with terms identifying causal inference

and impact evaluation methods. The full search string used in Scopus is provided in Appendix B.

The search strategy will pursue four main objectives: 1) Maximize recall while keeping methodological rigor; 2) Capture causal and quasi-experimental evidence; 3) Align with the EGM intervention and outcome framework and 4) Ensure transparency and replicability. The design balances sensitivity and specificity, thus it is structured around two mandatory blocks that ensure that the search retrieves studies that evaluate clearly defined policy interventions within the EGM scope and apply rigorous causal inference methods. The strategy is structured around two blocks:

Block A – Intervention / Sector Block: This block captures policy interventions and sectoral terms related to trade, investment, and tourism. Intervention keywords were derived directly from the EGM intervention framework, developed in collaboration with the Productivity, Trade and Innovation Sector team. For each intervention category, a core set of policy-specific terms was identified. Where necessary, limited expansions were included to capture variations in terminology used in applied economic research or specific names of well-known programs. This approach balances conceptual alignment with empirical sensitivity.

For each intervention category, a core set of policy-specific terms was defined. These terms were selected to ensure close alignment with the policy instruments captured in the EGM framework rather than with broad economic outcomes.

Block B – Impact Evaluation Methodology Block: Block B ensures that the search captures studies employing rigorous causal inference methods.

The methodology block includes terminology associated (but not limited to) with: Randomized controlled trials (RCTs) and field experiments; natural experiments; Quasi-experimental methods; Difference-in-differences (DiD); Regression discontinuity designs (RDD); Instrumental variables (IV); Synthetic control methods; Matching and related counterfactual approaches; and econometric panel estimators commonly used in impact evaluation. Furthermore, because not all relevant studies mention their methodology in the title or abstract, the keywords also include broader terms referring to impact evaluations, such as “causal design” and “impact evaluation”.

Requiring the presence of methodological terms at the search stage strengthens rigor and reduces the retrieval of purely descriptive or correlational studies. Hence, this block is intentionally comprehensive to capture variations in methodological terminology across journals and disciplines.

We will first assess the volume and relevance of results retrieved using Block A AND Block B³. The volume and relevance of retrieved records will be assessed at this stage. This adaptive structure allows the strategy to preserve

³ The boolean structure will follow: (Block 1 – Methodology Terms) AND (Block 2 – Intervention Terms)

sensitivity while retaining the option to improve precision if needed. In addition, the search strategy will be limited to subject areas such as social sciences, econometrics, business, environmental sciences, energy, agricultural sciences, decision sciences, psychology, and the arts. It will also exclude publications such as opinion editorials, conference papers, letters, data papers, notes, and reviews.

A separate search strategy will be developed to identify systematic reviews, in line with evidence synthesis best practices, that recommend distinguishing between searches for primary studies and for evidence syntheses. This complementary search will retain the same intervention block used in the main search strategy but will incorporate an adapted methodological filter to capture systematic reviews and meta-analyses. This approach is intended to ensure comprehensive coverage of high-level evidence while maintaining consistency with the EGM's overall conceptual framework.

To identify relevant grey literature, the team will conduct targeted searches of databases and institutional repositories of organizations identified by PTI/TIN and VPS/KLD. The full list of institutions included in this search is provided in Appendix C. Search procedures for these sources depend on the architecture of each and will be documented throughout the search.

In addition, approximately 87 publications manually curated by the IDB Knowledge and Learning Division as part of the Trade and Investment and Tourism SFDs will be incorporated into the search results. These studies will be processed using the same screening protocol as for all other records to ensure consistency and minimize selection bias in identifying relevant evidence.

To assess the performance of the search strategy, we will conduct a validation exercise using a curated benchmark list of well-known studies on trade and investment policies that are indexed in Scopus and employ credible causal inference methods. The benchmark list included seminal contributions on export promotion, trade agreements, foreign direct investment (FDI), and trade liberalization.

The search strategy is expected to retrieve approximately 70% of the benchmark studies, which will be considered a reasonable balance between sensitivity and specificity. Studies that are not retrieved are expected to be those in which the intervention of interest or the study design is not explicitly referenced in the title or abstract, or where alternative terminology is used. This level of recall would be consistent with expectations for large-scale, multidisciplinary searches in Scopus and suggests that the search strategy is sufficiently robust to identify a broad, policy-relevant evidence base while maintaining a manageable number of records for screening.

Search results will be stored in three bibliographic (RIS) files, one each for impact evaluations, systematic reviews, and grey literature. For future updates, a search script will query the Scopus and Overton APIs using the search strategy described here and append newly identified records to these files, keeping a backup of the previous version before each update.

Duplicate records will be identified automatically, first by exact DOI match and then by text similarity (TF-IDF cosine similarity). Records with a similarity score of 0.95 or above will be treated as duplicates and excluded. Records with scores between 0.71 and 0.94 will be flagged as potentially similar so that they can be verified manually.

5. Screening protocol

5.1 Title and abstract screening

Title and abstract screening will be conducted by a screening script that classifies each record using an LLM (currently Claude Haiku 4.5, Anthropic) and a structured prompt whose rules, presented in Appendix G, operationalize the decision guide in Appendix D. Before any model call, deterministic filters will resolve records that do not require model inference: records published in 2005 or earlier, duplicate records, and non-eligible publication types (e.g., editorials or commentaries) will be excluded, while records without an abstract will be included and verified at the full-text stage. For the remaining records, the model will assess three hierarchical criteria in sequence: (1) publication language; (2) whether the study evaluates an intervention within the EGM framework; and (3) whether it plausibly uses an eligible causal design or, for systematic reviews and meta-analyses, addresses an effectiveness question. The assessment will stop as soon as a criterion leads to exclusion, and each decision will be stored together with its exclusion reason.

Two principles will guide how the model handles uncertainty. First, a conservative default: when eligibility cannot be determined from the title and abstract, the record will be included and assessed against the full text, since false positives can still be removed at later stages while false negatives are difficult to recover. Second, a specificity requirement: studies must evaluate an intervention within the EGM framework, rather than merely mention trade, investment, or tourism as context or use related variables as controls.

Before the full corpus is screened, the prompt will be validated against the title and abstract gold standard described in Section 2.1. Prompt versions will be compared against the gold standard using recall, precision, accuracy, F1, and specificity, with recall treated as the priority metric because it measures the risk of missing eligible studies. Prompts will be refined iteratively by examining recurring false negatives and false positives and adding clarifications, disambiguation rules, and worked examples, without changing the underlying eligibility criteria. A prompt version will be used for full screening only when both recall and accuracy exceed 0.90 on the gold standard and remain stable across bootstrap sub-samples. Major prompt changes will be reviewed by the team to avoid over-fitting to the gold-standard sample.

When the search is updated, the script will screen only newly added records, preserving the decisions already made. The gold-standard validation will be repeated whenever the prompt, the eligibility criteria, or the underlying model changes, since model performance cannot be assumed to persist across model versions or providers.

5.2 Full-text retrieving

A retrieval script will attempt to download the full text of every study included at the title and abstract stage. The script will query a cascade of seven sources in sequence and stop at the first successful retrieval: the Elsevier and Wiley publisher APIs, Unpaywall, the IDB Knowledge Graph, Crossref, targeted URL retrieval for grey literature, and Semantic Scholar. Its modular structure will allow additional sources and institutional subscriptions to be incorporated as they become available.

Studies that cannot be retrieved automatically will be flagged for manual retrieval through institutional library subscriptions. Manually obtained files will be registered in the same pipeline and processed in the same way as automatically retrieved documents. The retrieval status and source of each study will be recorded, and retrieval rates will be reported alongside screening results. Because retrieval depends on open-access coverage and available subscriptions, studies whose full text cannot be obtained will be documented as not retrieved and will not proceed to full-text screening.

5.3 Full-text screening

Full-text screening and coding will be performed in a single step by a full-text script that uses an LLM (currently Claude Sonnet 4.6, Anthropic) and a structured prompt whose rules, presented in Appendix H, operationalize the criteria in Appendix E and the intervention, outcome, and study design framework of this protocol. Full texts retrieved as PDF, XML, or HTML files will first be converted to plain text. The model will then identify the eligible interventions, outcomes, and study designs reported in each study. A study will be included if and only if the model identifies at least one eligible value in each of these three dimensions; otherwise, it will be excluded. Files that cannot be read or processed will be recorded and retried in subsequent runs.

Because full-text coding requires distinguishing among closely related categories, the prompt will supplement the framework with explicit disambiguation rules developed from errors observed during testing. For example, it will distinguish physical border infrastructure from procedural or regulatory reforms, and it will require the model to check whether an outcome fits a more specific category before assigning a general one. Multiple classifications will be allowed when a study evaluates more than one intervention or outcome. This full-text screening protocol also applies to systematic reviews; however, only those that report the study designs of their underlying publications will advance to the next stage.

Before full deployment, the prompt will be validated against the full-text gold standard described in Section 2.1. Performance will be assessed for the include/exclude decision using recall, precision, accuracy, and F1, and separately for the coding of interventions, outcomes, and study designs using micro-averaged precision and recall, since a study can receive several codes in each dimension. Trained reviewers will review samples of model decisions and resolve ambiguous or conflicting cases, consulting a senior reviewer where needed. As at the title and abstract stage, the validation will be repeated whenever prompts, taxonomies, or model versions change.

6. Data extraction

Data will be extracted from all studies included at the full-text stage using the data extraction template in Appendix F. Extraction will combine two sources. First, a dedicated extraction script will use a separate LLM call (currently Claude Sonnet 4.6, Anthropic) and a structured prompt, with explicit definitions and coding rules for each variable, to extract information from the full text. Second, the remaining variables will be derived without model calls from study metadata and external sources, including Scopus, the SCImago Journal Rank, and the World Bank country classifications.

All AI-supported extraction will be implemented within a human-in-the-loop framework, in which trained researchers define the variables, test the prompt, and review samples of the extracted information. The team will document any deviation from this protocol arising from the use of AI-assisted methods to ensure transparency and replicability.

The data extraction will cover the following broad areas:

6.1 Basic study and publication information

This coding will capture the general characteristics of each study: title, authors, year of publication, journal or publisher, DOI and URL, abstract, language, open-access status, journal rank, and whether the study is an IDB publication or has IDB-affiliated authors. It will also capture the study location (country, region, and income level), the interventions, outcomes, and study designs coded at the full-text stage, the unit of observation, and the timeframe between the intervention and the measured outcomes. For each outcome, the extraction will record the direction of the estimated effect (positive, negative, mixed, not significant, or not available) and its role in the study (main result or mechanism). Only studies written in an eligible language and reporting at least one main outcome will be included in the map. Effect sizes for evidence synthesis will not be extracted.

6.2 Filters

The online map will display all the included studies in its default view. However, it will offer options to filter studies based on specific criteria, enabling users to view a subset of the evidence base. Table 4 shows the filters proposed for this map. The filters developed for this project will be reviewed and potentially refined in light of any feedback received by PTI/TIN, as the targeted main users of the EGM, and insights from project implementation.

Table 4: List of variables that will be used as filters to classify included studies

Filter	Options for Dropdown	Explanation
Intervention category	Trade facilitation Export promotion Attraction of Foreign Direct Investment (FDI) Tourism	Menu will allow users to identify studies evaluating interventions in a specific intervention group of the framework.
Outcome category	Access Quality Affordability Service management Sector governance	Menu will allow users to identify studies measuring outcomes in a specific outcome group of the framework.

	Socio-economic outcomes	
Country	All countries 15+ countries (studies covering more than 15 countries)	Menu will allow for the identification of studies which occurred in specific countries.
Region	East Asia and Pacific (EAP) Europe and Central Asia (ECA) Latin America and the Caribbean (LAC) Middle East and North Africa (MENA) South Asia (SAR) Sub-Saharan Africa (SSA) North America Multi-region (studies covering more than 15 countries)	Menu will allow for the identification of studies which occurred in specific regions based on the World Bank classification.
Income Level	Low income Lower middle income Upper middle income High income	The menu will allow for the identification of studies that occurred in countries of a specific income category based on the World Bank classification.
Study type	Impact evaluation Systematic review	Users may want to focus on a specific type of publications.
Study design	For impact evaluations Experimental methods Regression discontinuity design Instrumental variables Selection models DiD and Fixed Effects Interrupted time series Weighting and matching methods Synthetic control methods For systematic reviews Meta-analysis and meta-regressions Narrative and descriptive reviews Other reviews	The menu will allow for the identification of studies using a specific study design.
Unit of observation	Individual Household Firm/organisation Establishment Sector Product Area or polygon Municipality District/State/Provincial Country Other Not specified	Menu will allow users to identify studies by the level at which the data are analyzed.
Timeframe	0–1 years 1–2 years 2–5 years 5–10 years More than 10 years Not available	Menu will allow users to identify studies by the time elapsed between the implementation of the intervention and the measurement of the outcome.

Cost information	Cost information (program costs and/or cost per participant) Cost-benefit/cost-effectiveness analyses None	Menu will allow for the identification of studies presenting cost evidence.
Heterogeneity	Size of the firms Size of the geographical unit Economic sector Gender Race/ethnicity/community None	Menu will allow users to identify whether the study integrates a heterogeneity component. The heterogeneity could be measured both at the intervention or the outcome level.
Cross-cutting themes	Institutional capacity Digital transformation Skills development None	Menu will allow users to identify whether the intervention includes a cross-cutting component related to institutional strengthening, digitalization, or skills development.
Year of publication	2006-2025	Menu will allow users to review evidence from specific time periods.
Language	English Spanish Portuguese French	Menu will allow users to identify studies written in a specific language.
Journal Rank	Q1 Q2 Q3 Q4 Not indexed/Other	Menu will allow users to review evidence from higher quality journals based on the SCImago Journal Rank (2025).
Open access	Yes/No	Menu will allow users to identify studies whose full text is freely available.
IDB study	Yes/No	Menu will allow users to identify studies with at least one author affiliated with the IDB.

6.3 Data extraction training and pilot

Before the full extraction run, the extraction prompt will be piloted on a sample of included studies, and trained coders will check the extracted variables manually against the full texts to confirm that each field is captured correctly. The team will refine variable definitions and coding rules in response to recurring errors before committing to a full run. The project team will meet periodically to resolve queries and agree on refinements to the variable definitions.

6.4 Dealing with multicomponent interventions

Multicomponent interventions are those that include several activities or components spanning different intervention categories. For example, an intervention may simultaneously implement export promotion activities and streamline customs procedures through a single reform package. Studies coded under more than one intervention will be flagged as multicomponent in the extraction dataset.

In the online map, studies whose interventions span more than one intervention group will be displayed in a separate “Multi-component intervention” row, so that each study is counted only once across intervention groups. Studies whose interventions fall within a single group will be shown under each relevant intervention category.

Common combinations of interventions will be reviewed with PTI/TIN to assess whether specific bundles warrant their own category in future updates of the map.

6.5 Quality checks

Throughout the extraction process, the project team will check the extracted data for accuracy and consistency. Records for which extraction fails, or whose files cannot be processed, will be flagged in the extraction dataset and excluded from the map until they are successfully processed. Trained coders will review samples of extracted records, and targeted reviews will be conducted when inconsistencies are detected or when prompts, taxonomies, or model versions change. This quality check process is in place to ensure that the extracted data is accurate and does not assess the quality of the study itself or the evidence presented in it.

7. Analysis and reporting

The IDB’s trade and investment division will be in charge of a posterior analysis of the resulting EGM, for which the Knowledge and Learning division could provide support, if needed. This EGM will serve as an input for the 2027 Trade, Investment and Tourism Thematic Framework Document.

8. Engagement and communication plan

Nothing planned at the moment. To be confirmed.

References

Campbell Collaboration, & Centre of Excellence for Development Impact and Learning (CEDIL). (2022). *Evidence and gap map: Transport interventions in low- and middle-income countries*.

Inter-American Development Bank (IDB) & International Initiative for Impact Evaluation (3ie). (2025). *Building stronger transport policy: An evidence gap map*.

International Initiative for Impact Evaluation (3ie). (forthcoming). *Private sector engagement (PSE) interventions evidence gap map*. Funded by GIZ.

MacDonald, H., & Young, S. (2024). New search guidance for Campbell systematic reviews. *Campbell Systematic Reviews*, 20(4), e70005.
<https://doi.org/10.1002/cl2.70005>

SDG Synthesis Coalition. (2023). *SDG partnership pillar evidence and gap map*. American Institutes for Research (AIR).

The Campbell Collaboration. (2021). *Campbell systematic reviews: Policies and guidelines*. <https://campbellcollaboration.org>

Volpe Martincus, C. (2016). *Out of the border labyrinth: An assessment of trade facilitation initiatives in Latin America and the Caribbean*. Inter-American Development Bank.

Appendix A. Alignment with the catalogs of results

The following table presents a crosswalk between the Tourism and Trade and Investment Catalogs of Results Indicators and the EGM outcomes framework. Both catalogs were reviewed, and each cell indicates the specific catalog subcategory to which a given EGM outcome is mapped. Columns correspond to the main categories in the catalog (Access, Quality, Affordability, Service Management/Efficiency, and Long-Term Results) while rows reflect the outcome structure of the EGM framework. Some outcomes span multiple subcategories and, in certain cases, cut across more than one result category, reflecting their multidimensional nature. Although Sector Governance is included as an outcome area in the EGM framework, it does not appear as a standalone category in the published version of the Catalogs; however, it is retained for its analytical relevance to the sector.

1. Access					
Outcomes	Access	Quality	Affordability	Service Management/efficiency	Long-term results
Extensive margin of trade	Coverage/ Availability/ Use				
Intensive margin of trade					Economic impact
Foreign Direct Investment outcomes	Coverage/ Availability/ Use				Economic impact
Visitor flows and engagement	Coverage/ Availability/ Use + Use Intensity				
2. Quality					
Service quality		Quality general			
User satisfaction		Interpersonal/ User satisfaction			
3. Affordability					
Travel time and costs frictions		Quality general	Affordability	Process Optimization/ Operational Performance	
Expenditure in tourism activities	Use intensity				
Prices and inflation					Economic impact
4. Service management/efficiency					
Firm entry, exit and formality	Coverage/ Availability/ Use				Economic impact + Employment
Income, sales and profits				Commercial / Financial Performance	Economic impact
Access to finance and capital accumulation					Economic impact
Labor market outcomes					Economic impact + Employment

Access to better inputs and services				Operational performance efficiency/ Capacity building	
Technology adoption				Operational performance efficiency/ Capacity building	
Productivity and competitiveness				Operational performance efficiency/ Capacity building	
Innovation				Operational performance efficiency/ Capacity building	
5.Sector governance					
Fiscal revenue and tax compliance					
Market concentration					
Reforms or approved agreements					
6. Socio-economic outcomes					
Household welfare and poverty					Long-term results
Inequality and segregation					Long-term results
Environmental quality and emissions					Long-term results
Skills					Long-term results

Appendix B. Scopus database search strategy⁴

/* Block A – Intervention / Sector Block: */

TITLE-ABS-KEY("internationalization of companies" OR "internationalization of firms" OR "WTO addition" OR "WTO accession" OR "trade agreement*" OR "bilateral investment" OR "free trade agreement*" OR "preferential trade agreement*" OR "regional trade agreement*" OR "economic integration agreement*" OR "economic integration treaty" OR "customs union*" OR "trade liberalization agreement*" OR "trade liberalization" OR "bilateral trade agreement*" OR "multilateral trade agreement*" OR "plurilateral trade agreement*" OR "trade cooperation agreement*" OR "economic partnership agreement*" OR "facilitación del comercio" OR "reforma aduanera" OR "modernización aduanera" OR "gestión de riesgo aduanero" OR "ventanilla única de comercio exterior" OR "puesto fronterizo integrado" OR "zona franca" OR "trade facilitation" OR "customs reform" OR "border procedures" OR "trade procedures" OR "administrative simplification" OR "risk-based inspection" OR "risk management in customs" OR "customs risk management system*" OR "risk-based selectivity" OR "WTO Trade Facilitation Agreement" OR "trade facilitation agreement" OR "capacity building in trade" OR "customs training" OR "Authorized Economic Operator*" OR "AEO program*" OR "harmonization of border procedures" OR "inter-agency coordination mechanisms" OR "border infrastructure" OR "border crossing*" OR "ports of entry" OR "seaport*" OR "maritime port*" OR "port terminal*" OR "port infrastructure" OR "container port*" OR "dry port*" OR "port management" OR "customs infrastructure" OR "logistics infrastructure" OR "trade corridor*" OR "land border*" OR "border modernization" OR "national logistic* strateg*" OR "anti-evasion" OR "customs digitalization" OR "customs modernization" OR "customs clearance time" OR ("clearance" W/2 customs) OR ("time" W/3 "customs clearance") OR "time to clear customs" OR "time to clear customs" OR "trade costs" OR "border delays" OR "automated customs systems" OR "e-customs" OR "customs automation" OR "border digitalization" OR "electronic declaration" OR "digital borders" OR "automated risk assessment" OR "channel selectivity tools" OR "non-intrusive inspection technologies" OR "intelligence-based customs controls" OR "automation of customs clearance" OR "automation of inspection processes" OR "single window" OR "national single window" OR "trade single window" OR "interoperability" OR "integrated border management" OR "information sharing systems" OR "ventanilla unica de comercio exterior" OR "declaracion anticipada" OR "digital trade regulation" OR "e-commerce regulation" OR "cross-border data flows" OR "data localization" OR "digital transactions" OR "digital economy policy" OR "data governance" OR "online trade" OR "export promotion" OR "export development programs" OR "trade promotion" OR "trade promotion organizations" OR "trade promotion program*" OR "firm internationalization" OR "export competitiveness" OR "export capacity building" OR "export readiness" OR "export training" OR "internationalization training" OR "export skills" OR "export training program*" OR "export readiness program*" OR "SME internationalization support" OR "firm internationalization program*" OR "export mentoring program*" OR "national export strategies" OR "export acceleration programs" OR "trade fairs" OR "trade missions" OR "buyer-seller matching" OR "business matchmaking" OR "B2B platforms" OR "export market access" OR "market linkage programs" OR "virtual trade fairs" OR "online buyer missions" OR "trade finance" OR "export finance" OR "export credit" OR "export guarantees" OR "trade credit" OR "SME trade finance" OR "letters of credit" OR "factoring" OR "export insurance" OR "quality infrastructure" OR "standards and certification" OR "technical regulations" OR "conformity assessment" OR "accreditation" OR "sanitary and phytosanitary standards" OR "investment promotion" OR "investment climate" OR "investment facilitation" OR "investment promotion agencies" OR "investment promotion agenc*" OR "investment promotion program*" OR "investor services" OR

⁴ This search string was adapted following the best practices from the Transportation EGM (2025) produced by the IDB and Zie.

"aftercare services" OR "single windows for investment" OR "one-stop shop for investment" OR "FDI strategy" OR "investment incentives" OR "investment tax incentive*" OR "FDI tax incentive*" OR "export tax incentive*" OR "tourism tax incentive*" OR "special economic zone tax incentive*" OR "special economic zones" OR "SEZs" OR "free zones" OR "zonas francas" OR "tourism skills" OR "UNESCO World Heritage list inscription" OR "UNESCO World Heritage list inscription" OR "blue flag certificat*" OR "blue flag program" OR "tourism competitiveness" OR "tourism productivity" OR "tourism promotion" OR "tourism destination" OR "destination marketing" OR "tourism governance" OR "tourism infrastructure" OR "tourism investment" OR "tourism development zones" OR "special tourism zones" OR "tourism SEZs" OR "ecotourism" OR "climate-resilient tourism" OR "one-stop border post" OR "pre-arrival processing" OR "post-clearance audit" OR "export subsidy program" OR "tourism cluster" OR "short-term rental regulation" OR "technology extension services" OR "industrial extension services" OR "fábricas de productividad" OR "Exporta Fácil" OR "simplified export scheme*" OR "postal export scheme*" OR "postal export program*" OR "small-value export*" OR "Declaração Simplificada de Exportação" OR "electronic single window" OR "foreign trade single window" OR "virtual single window" OR "two-window arrangement" OR "single entry point" OR "interoperable single window" OR "Authorized Economic Operator certification" OR "AEO mutual recognition agreement*" OR "trusted trader program*" OR "customs risk profiling" OR "coordinated border management" OR "time release study" OR "export permit reform" OR "digital export permit" OR "ConnectAmericas")

/*Block B – Methodological Block*/

AND TITLE-ABS-KEY((match* W/2 ("propensity" OR "coarsened" OR "covariate" OR "neighbor" OR "method" OR "data"))) OR "propensity score" OR "matching econometrics" OR "difference-in-differences" OR "difference in differences" OR "differences-in-differences" OR "double difference" OR "quasi-experiment*" OR "quasi-random" OR ("estimator" AND "evaluat*") OR "instrumental variable*" OR (Heckman W/3 (model* OR approach*)) OR (("two-stage" OR "three-stage" OR "four-stage") W/3 (control* OR function* OR regression* OR "least squares")) OR "2SLS" OR "regression discontinuity" OR "synthetic control" OR "time series" OR "counterfactual" OR ("segment*" W/1 regression) OR (non W/2 participant*) OR (("control" OR "comparison") W/2 (group* OR condition* OR area* OR village* OR household* OR intervention OR site*)) OR (panel W/2 (data OR household* OR model*)) OR (("exploit*" OR "tak* advantage") W/3 (variation* OR exogen* OR heterogen*)) OR (econometric W/2 (model* OR adjust*)) OR (select* W/2 ("bias*" OR "self")) OR "fixed effects" OR "random effects" OR "event study" OR "staggered adoption" OR "causal inference" OR "self-selection" OR "identification strategy" OR "empirical strategy" OR "estimation strategy" OR "random control trial" OR ((impact* OR effect*) W/2 (evaluat* OR assess* OR analy* OR analysis OR estimate* OR causal OR causality OR estimation)) OR ("estimate*" W/3 (impact* OR effect*)) OR (analy* W/3 (impact* OR effect*)) OR (examin* W/3 (impact* OR effect*)) OR (study W/3 (impact* OR effect*)) OR (investigat* W/3 (impact* OR effect*)) OR ("empirical" W/2 (analysis OR evidence OR strategy)) OR ("firm-level" W/2 (evidence OR data OR analysis)) OR ("micro-level" W/2 (evidence OR data OR analysis)) OR ("panel data" W/2 (analysis OR evidence OR estimat*)) OR "impact evaluation" OR "policy evaluation" OR "policy effectiveness" OR "evaluation study" OR "program evaluation" OR "project evaluation" OR "evaluation research" OR "natural experiment" OR "program effectiveness" OR "outcome assessment" OR "treated by" OR "field experiment" OR (experiment* W/2 (design OR study OR research OR evaluat* OR evidence OR variation)) OR (("random" OR "randomized" OR "randomly") W/2 (trial OR assign* OR treatment OR control* OR allocat* OR experiment*))

/*Block C – Other Inclusion/Exclusion Criteria*/

AND PUBYEAR > 2005 AND (EXCLUDE (DOCTYPE,"re") OR EXCLUDE (DOCTYPE,"cr") OR EXCLUDE (DOCTYPE,"no") OR EXCLUDE (DOCTYPE,"ed") OR EXCLUDE (DOCTYPE,"sh") OR EXCLUDE (DOCTYPE,"er") OR EXCLUDE (DOCTYPE,"le") OR EXCLUDE (DOCTYPE,"dp") OR EXCLUDE (DOCTYPE,"tb") OR EXCLUDE (DOCTYPE,"bk")) AND (LIMIT-TO (SUBJAREA,"AGRI") OR LIMIT-TO (SUBJAREA,"ENER") OR LIMIT-TO (SUBJAREA,"ENVI") OR LIMIT-TO (SUBJAREA,"ARTS") OR LIMIT-TO (SUBJAREA,"BUSI") OR LIMIT-TO (SUBJAREA,"DECI") OR LIMIT-TO (SUBJAREA,"ECON") OR LIMIT-TO (SUBJAREA,"PSYC") OR LIMIT-TO (SUBJAREA,"SOCI") OR LIMIT-TO (SUBJAREA,"MULT")) AND (LIMIT-TO (LANGUAGE,"English") OR LIMIT-TO (LANGUAGE,"French") OR LIMIT-TO (LANGUAGE,"Spanish") OR LIMIT-TO (LANGUAGE,"Portuguese")))

Appendix C. Grey literature sources and repositories

To identify relevant grey literature, in its first stage, the EGM will include targeted searches of institutional repositories, organizational websites, and working paper series from international organizations, development banks, and research institutions.

A. Multilateral and international organizations

- Inter-American Development Bank (IDB)
- World Bank Group
- Development Bank of Latin America and the Caribbean (CAF)
- Economic Commission for Latin America and the Caribbean (ECLAC/CEPAL)
- Asian Development Bank (ADB)
- African Development Bank (AfDB)
- European Investment Bank
- United Nations Trade and Development (UNCTAD)
- United Nations Industrial Development Organization (UNIDO)
- International Monetary Fund (IMF)
- World Trade Organization (WTO)
- Organization for Economic Co-operation and Development (OECD)
- United Nations World Tourism Organization (UNWTO)

B. Policy research institutions and think tanks

- Peterson Institute for International Economics

C. Working paper repositories and research networks

- National Bureau of Economic Research (NBER) – Working Papers
- Institute of Labor Economics (IZA) – Discussion Papers
- United Nations University World Institute for Development Economics Research (UNU-WIDER) – Working Papers
- Kiel Institute for the World Economy (IfW Kiel) – Working Papers

Appendix D. Title and abstract screening protocol

This guide was used by trained coders to screen the title and abstract gold-standard sample manually. The rules applied by the AI full-text screening and coding model, which were developed and validated against that gold standard, are presented in Appendix G.

Questions		Decision				Notes
1	Is the study in English, Spanish, Portuguese, or French?	NO	EXCLUDE on publication language	YES	Continue to the next question	If not sure about the language of the full text (e.g., title and abstract shown in multiple languages), continue to the next question. In such cases, this criterion should be verified during full-text screening.
2	Was the study published after 2005?	NO	EXCLUDE on publication year	YES	Continue to the next question	If the record does not indicate the year, continue to the next question. In such cases, this criterion should be verified during full-text screening.
4	Does the study evaluate a trade, investment or tourism intervention listed in the EGM framework?	NO	EXCLUDE on intervention	YES	Continue to the next question	Refer to the protocol for a detailed description of includable interventions
5	Does the study use quantitative experimental or quasi-experimental designs? OR if a review, does it address effectiveness questions?	NO	EXCLUDE on study design	YES	Choose an option from below	Includable designs: randomized studies, matching (incl. PSM), FE (incl. DID), IV, RDD, synthetic control, interrupted time series, and other attributional methods that account for selection bias and confounding. Excluded studies: descriptive and regression studies without a clear identification of causal effects. If unclear at this point, screen at the FT stage. Effectiveness systematic reviews and meta-analyses. Excluded studies: qualitative, descriptive, literature reviews and those failing to describe the methodology adequately.
6	Is the study a duplicate?	YES	EXCLUDE as known duplicate			Choose only one option.
7	Is this record a non-eligible publication type (e.g., commentary, Op-Ed, narrative review, or language mismatch)?	YES	EXCLUDE as other document			
8	Are you unsure about inclusion?	YES	INCLUDE second opinion			

9	Does the study meet all eligible criteria?	YES	INCLUDE on title and abstract	
Whenever the response to the question is UNCLEAR, continue to the next question.				

Appendix E. Full-text screening protocol for impact evaluations⁵

This guide was used by trained coders to screen and code the full-text gold-standard sample manually. The rules applied by the AI full-text screening and coding model, which were developed and validated against that gold standard, are presented in Appendix H.

Questions		Decision				Notes
1	Are you confident you can exclude the paper based on the abstract alone (you should always begin FT screening by reading the abstract first)	YES	MARKER - Exclude on TA	NO	Continue to the next question	This marker will probably be used mostly for incomplete or absent abstracts that got included at the title and abstract screening and once full text is retrieved, screening the complete abstract allows us to confirm or deny its relevance. However, all studies should be checked. Use the notes section to indicate the reason for exclusion.
2	Is there a clear intervention?	YES	EXCLUDE - No intervention	NO	Continue to the next question	Studies describing trade, investment or tourism as a phenomenon rather than aiming to establish a causal relationship between interventions and outcomes should be excluded. For example, studies evaluating the differences in tariffs between countries, but without clearly exploiting a particular trade agreement will be excluded.
3	Does the study evaluate an intervention that does NOT have a trade, investment or tourism component?	YES	EXCLUDE - Not a relevant intervention	NO	Continue to the next question	Refer to the protocol for a detailed description of the included interventions. [Briefly describe includable interventions.]
4	Does the study evaluate the effects on outcomes	YES	EXCLUDE - Not a relevant outcome	NO	Continue to the next question	Refer to the protocol for a detailed description of the included outcomes.

⁵ For systematic reviews, a similar checklist should be used, with the necessary adaptations.

	OTHER THAN [broad categories for outcomes of interest]?					[Briefly describe the includable outcomes. Define what to do with aggregate measures.]
5	The study does not have a control group or does not use a valid impact evaluation method to assess the causal relationship between the intervention and the outcomes.	YES	EXCLUDE - No valid causal inference	NO	Choose an option from below	Includable designs: statistical matching, FE (incl. DID), IV, RDD, synthetic control, and other attributional methods that account for selection bias and confounding. Exclude studies that fail to adequately describe the methodology and those with insufficient clusters (assignment at the cluster level AND there is only one cluster in either the treatment or control condition).
6	Is the study a duplicate?	YES	EXCLUDE as known duplicate			Choose only one option.
7	Are you unsure about inclusion?	YES	INCLUDE second opinion			
8	Does the study meet all eligible criteria?	YES	INCLUDE on full text			

Appendix F: Data extraction template

Code	Subcode
Study Information	Internal document ID
	Title
	Language
	IDB publication
	IDB study (IDB-affiliated authors)
	Journal rank
Publication Information	Study type (impact evaluation or systematic review)
	Search source (journal or grey literature)
	Authors
	DOI
	Abstract
	Journal name
	Year of Publication
	URL
	Open access
Geographic Information	Country name(s)
	Region
	Country income level
Methodological information	Evaluation design
	Evaluation method
	Unit of Observation
EGM framework	Intervention group
	Intervention category
	Multicomponent intervention
	Outcome group
	Outcome category
	Outcome type (main result, mechanism, or not specified)
	Effect direction (per outcome)
	Cost information
Custom fields	Mechanism (up to three)
	Timeframe
	Size of the firms (if apply)
	Size of the geographical unit (if apply)
	Economic sector (if apply)
	Gender (if apply)
	Race/ethnicity/community (if apply)
	Tags: Skills/Digital Transformation/Institutional Capacity

Appendix G. Title and abstract screening rules applied by the AI model

The table below summarizes the rules applied by the title and abstract screening script. Part A lists the deterministic filters applied before any model call. Part B lists the questions the model answers in sequence; the assessment stops as soon as a question leads to exclusion. Throughout Part B, the model applies a conservative principle: when in doubt, it answers YES and moves to the next question.

Step	Rule	Decision
A1. Publication year	Was the record published in 2005 or earlier?	YES → EXCLUDE (publication year)
A2. Duplicates	Does the record match another record by DOI, or by text similarity (TF-IDF cosine similarity of 0.95 or above)?	YES → EXCLUDE (known duplicate)
A3. Publication type	Is the record a non-eligible publication type, such as an editorial or commentary?	YES → EXCLUDE (other document)
A4. Missing abstract	Does the record lack an abstract?	YES → INCLUDE (no abstract); eligibility verified at the full-text stage
B1. Language	Is the title and/or abstract written in English, Spanish, Portuguese, or French?	NO → EXCLUDE (publication language). YES or UNCLEAR → continue to B2
B2. Intervention	Priority check for reviews: if the abstract mentions a “systematic review,” “scoping review,” or “meta-analysis” on a topic within the intervention framework, continue directly to B3. Otherwise: does the study evaluate the effect of at least one of the 13 interventions in the framework? The study must have an identifiable policy lever, policy shock, or intervention. Not eligible: purely theoretical models with no real-world policy; interventions outside the framework (e.g., general transport infrastructure unrelated to trade, capital controls); tourism studies limited to individual tourist psychology with no policy or program; and studies of firm behavior in which no specific policy is the treatment.	NO → EXCLUDE (intervention). YES or UNCLEAR → continue to B3
B3. Study design	Does the study possibly use an eligible design? Impact evaluations: randomized controlled trials and natural experiments, regression discontinuity, instrumental variables, selection models, difference-in-differences and fixed effects, synthetic control, matching and weighting, interrupted time series, and gravity or structural models with a clear identification strategy. Reviews: systematic reviews, scoping reviews, or meta-analyses on topics within the framework. The design does not need to be named in the abstract; any indication of quantitative causal analysis is sufficient. Not eligible unless combined with an eligible design: computable general equilibrium or other simulation models, cointegration as the primary strategy, optimization models, descriptive surveys or correlations, and narrative literature reviews with no empirical synthesis.	NO → EXCLUDE (study design). YES or UNCLEAR → INCLUDE (title and abstract)

Appendix H. Full-text screening and coding rules applied by the AI model

The table below summarizes the rules applied by the full-text screening and coding script. The model reads the full text, which may be in English, Spanish, Portuguese, or French, and codes all and only the interventions, outcomes, and study designs present in the study. Tables or equations that cannot be interpreted as plain text are ignored.

Dimension	Rule	Result
Interventions	Code all that apply among the 13 interventions in the framework (Section 3). Code only the treatment being evaluated, not variables used as controls, period indicators, or background context. Trade openness indices or tariff changes without an identifiable policy event are not eligible. Domestic transport, roads, railways, and ports are out of scope. Disambiguation rules distinguish closely related categories, such as border infrastructure versus regulatory reforms, tourism investment versus destination management, and the export promotion categories.	List of intervention codes (may be empty)
Outcomes	Code all that apply among the 24 outcomes in the framework (Section 3). Code only outcomes for which the study reports an estimated effect, not outcomes mentioned only in the discussion. Before coding a general category such as income, sales, and profits, check whether a more specific category applies.	List of outcome codes (may be empty)
Study designs	Code all that apply among the eight impact evaluation designs and three systematic review types (Section 3). Code only designs explicitly described in the study, not designs of cited papers. One-way fixed effects and before-and-after comparisons without a control group do not qualify as difference-in-differences or fixed effects. Narrative reviews qualify only if they explicitly follow a systematic review protocol (e.g., PRISMA or Campbell). Not eligible: simulation or computable general equilibrium models without an empirical identification strategy, and descriptive case studies.	List of study design codes (may be empty)
Inclusion decision	Is at least one value coded in each of the three dimensions? Systematic reviews advance only if they report the study designs of their underlying publications.	YES → INCLUDE. NO → EXCLUDE
Checks at extraction	Is the study written in an eligible language, and does it report at least one main outcome (rather than only mechanisms)?	NO → excluded from the map