Impact Evaluation (IE) Concept Note Template

[IE Title]

[Country]
[IE code]
[Date]

Keywords: [1] Choose one of more keywords/categories that describe your IE.[2] **(R)**

---

[1] Please refer to JEL classification codes http://papers.ssrn.com/sol3/displayjel.cfm .
[2] The concept note is aligned to Ethical clearance (E) and Registry (R) indicative requirements. These indicative requirements are referenced throughout the document.

Table of Contents

IE PROFILE INDICATORS

No.	Indicator	Description
1	IE code	IE code (hyperlink to IE portal)
2	IE Title	Legal title of the IE
3	IE TTL	As in IE portal
4	IE Contact Person	Name and affiliation/unit
5	Region	AFR/EAP/ECA/ LCR/ MEN/SAR
6	Sector Board/Global Practice	Use standard abbreviations
7	WBG PID (if IE is evaluating a WBG operation)	PID (hyperlink to operations portal)
8	WBG Project Name (if IE is evaluating a WBG operation)	Legal project name
9	Project TTL (if IE is evaluating a WBG operation)	TTL in operations portal
10	Intervention	<Policy/intervention to be evaluated> e.g., Financial literacy in high schools
11	Main Outcomes	<Most important outcome(s)> e.g., student financial knowledge, attitudes and behavior
12	IE Unit of Intervention/Randomization	Main level of treatment assignment/randomization if it applies (e.g., school)
13	Number of IE Units of Intervention	e.g. 900 schools
14	IE Unit of Analysis	Lowest level of analysis on which outcomes are measured (e.g., student)
15	Number of IE Units of Analysis	E.g., 20,000 students
16	Number of Treatment Arms	Number of treatment arms
17	IE Question 1 (Treatment Arm 1)	What is the impact of <intervention> on <outcomes>?
18	Method IE Question 1	Main method to estimate IE Question in (17). E.g., "Random assignment at the school level"
19	Mechanism tested in IE Question 1	If the treatment arm tests a particular mechanism, classify it accordingly, e.g., information, incentives, behavioral biases, constraints, accountability measures (For general discussion about mechanisms see http://www.itg.be/internet/ds/tde/doc/Astbury%20%26%20Leeuw%20.pdf). If the treatment arm does not identify a particular mechanism, write "Package"
20	IE Question 2 (Treatment Arm 2)	What is the impact of <intervention variation 1> on <outcomes>? E.g., What is the impact of a parental workshop on financial literacy on student knowledge, attitudes and behavior?
21	Method IE Question 2	Method IE Question in (20): e.g., "Random assignment at the parent level"
22	Mechanism tested in IE Question 2	See (19)
23	IE Question 3 (Treatment Arm 3)	See (20)
24	Method IE Question 3	See (20)
25	Mechanism tested in IE Question 3	See (18)
25	Gender-specific treatment (Yes, No)	Yes if it is a gender-specific intervention
27	Gender analysis (Yes, No)	Yes if there is stratification/power for gender-specific analysis
28	IE Team & Affiliations	Name 1 (Organization/Unit Affiliation, Role); Name 2 (Organization/Unit Affiliation, Role)...
29	Estimated Budget (including research time)	Total in USD
30	CN Review Date	Month-Year
31	Estimated Timeframe for IE	Month-Year to Month-Year
32	Main Local Counterpart Institution(s)	E.g., Ministry of Education

1. EXECUTIVE SUMMARY

(1 page)

- Describe the proposed IE **in non-technical language in one paragraph or less.** This could be an abstract of your IE. Include broad motivation/background and policy/research contribution. (E,R)
- Present IE questions and main outcome(s) the intervention aims to affect.
- Briefly explain how you are proposing to test your main evaluation question(s).

2. BACKGROUND AND KEY INSTITUTIONAL FEATURES

(1 page)

- Present an overview of the local context.
- Identify and define the problem: what is the policy/research problem this IE is proposing to study? Which groups are affected by the problem?
- Describe the intervention whether existing or new, implementing organization, institutional setting and any important consideration.
- Describe the intervention geographic/demographic scale and scope: Does it represent the "mode" of delivery in the country? **(R, E)**

3. LITERATURE REVIEW (E)

(1 page or less)

- Describe most relevant literature/scientific background specifically linked to your problem/evaluation question(s).

4. POLICY RELEVANCE

(1/2 page or less)

- Assess the extent to which the study may influence policy and institutional capacity at the national, regional, and international level. Explain how you plan to track the policy influence of your study (see Appendix on i2i sample indicators of IE influence on program/policy. These indicators, which are currently under revision, will be collected through Grant Monitoring and Reporting on annual basis from all i2i supported IEs).

5. THEORY OF CHANGE (E)

(1 Figure and 2-3 paragraphs)

- Describe the main elements of the intervention, and the hypothesized causal chain from inputs, through activities and outputs, to outcomes.
- Describe the main assumptions and other factors underlying the causal chain (internal and external).

A theory of change describes how the intervention is expected to affect the outcomes of interest (based on theory) but it does not demonstrate whether the intervention causes the observed outcomes. It usually includes the most important outcomes (intermediate and final) that are critical to the causal chain, even if not all will be measured (see Appendix for example).

A theory of change sets the structure for the hypotheses, evaluation questions, and outcomes of interest. It also lists key indicators for developing the implementation protocol and IE monitoring system aimed at understanding what is being evaluated, and whether the critical intervention activities/components were implemented/taken up as planned.

6. HYPOTHESES/EVALUATION QUESTIONS (E,R)

(1/2 page)

- List the hypotheses derived from your theory of change.
- List the main evaluation question(s) to be addressed by the proposed study. Evaluation questions connect the specific intervention/treatment variation to the outcomes of interest, and end with a question mark. They should be in the following format: What is the impact of <intervention/intervention variation> on <outcomes>? E.g., What is the impact of a parental workshop on financial literacy on student knowledge, attitudes and behavior?
- **You may have a broad evaluation question** based on the knowledge gap and the strategy proposed. **However**, **the number of specific questions in this section should be <u>perfectly aligned</u> to the number of your treatment arms** (i.e., if you have 3 treatment arms you should have three specific evaluation questions). Each question can be evaluated on a vector of outcomes (i.e., you may organize them as sub-questions). Methods to answer sub-questions on heterogeneous treatment effects and spillovers should be described in the methods section.
- Describe how the evaluation questions were derived.

7. MAIN OUTCOMES OF INTEREST (E,R)

(1 table)

- Briefly list and define main outcomes of interest (primary and secondary/intermediate) as in Table 1.
- Further details on how the outcomes will be measured/collected will go in the data collection section.

Table 1. Main Outcomes of Interest

Outcome Type	Outcome Name	Definition	Measurement Level
Primary/Secondary			

8. EVALUATION DESIGN AND SAMPLING STRATEGY (E,R)

(2 pages or less)

- Present the main features of the proposed evaluation design to address the evaluation question(s).
- Describe precisely the identification strategy (e.g., trial design including clustering, factorial, stratification details) for each evaluation question.
- Report all inclusion/exclusion criteria to define the target population/population studied, providers, settings, and clusters (as relevant).
- Report any **ethical issues** that may arise concerning the evaluation design and the sampling strategy (not related to data collection).

8.1 TREATMENT AND CONTROL GROUPS

- Provide specific description of features of each control and treatment arm (one paragraph per arm).

8.2 SAMPLE SIZE CALCULATIONS

- Present the sample size estimates. Describe how the sample size was determined, including the sampling frame, and main assumptions including Minimum Detectable Effect (MDE), variance estimates, intra-cluster correlation, and units per cluster (if applicable).

9. DATA COLLECTION (E,R)

(1 page if basic, 1-2 pages if include all sections for registration and ethical clearance)

- Describe main instruments for data collection

9.1 QUANTITATIVE INSTRUMENTS

- Describe how primary and secondary outcomes (from section 7) will be measured, their timing and frequency.

9.2 MANAGEMENT OF DATA QUALITY

- Describe methods used to enhance the quality of measurements (e.g., multiple observations, training of surveyors), electronic data collection, protocols for quality assurance.

9.3 ETHICAL ISSUES

- Describe if this IE will require ethical approval, informed consent procedures, and important ethical considerations related to data collection.

9.4 QUALITATIVE INSTRUMENTS

- Provide a description of all qualitative instruments (if applicable).

9.5 IE Implementation Monitoring System (R)

- Describe the IE implementation monitoring system, particularly, what specific indicators and system will be used to follow up the studied population, their treatment participation, treatment *actually* delivered and received based on activities, and outputs (see the theory of change section).

10. Data Processing and Analysis

(~ 1-2 pages)

10.1 Data Coding, Entry, and Editing[3] (E)

- Describe planned methods for data entry, and for handling missing data, imputations.

10.2 Model Specification for Quantitative Data Analysis

- Describe the statistical method(s) that will be used to compare groups for primary and secondary outcomes (the specific equation should be included), any transformations to quantitative data. Specify whether the standard errors will be clustered or corrected.
- Specify what IE parameter of interest will be estimated (e.g., ITT, TT, MTE, LATE).
- Describe how you plan to address multiple hypothesis testing.
- Describe methods for additional analyses, including spillovers and subgroup analyses.
- Provide a list of any variables to be collected to check balance and correct for potential selection due to attrition, non-response, take-up rate issues (all theoretically important variables to be measured at baseline, including, those thought to be related to participation/dropout/non-response and the outcomes of interest).
- Lay out a strategy to follow up, test and correct for (if required) sources of bias (e.g., non-random attrition, non-response, endogenous take-up).
- State if you plan to register this IE (see selected links below)
    - AEA RCT Registry (https://www.socialscienceregistry.org/)
    - 3ie Registry (http://www.3ieimpact.org/evaluation/ridie/)

11. Study Limitations and Risks (E)

(1/2 page)

- Provide an assessment of risk and threat to internal validity (related to previous section)
- Discuss issues related to external validity, particularly (i) representativeness of the sample; (ii) representativeness of the institution(s) delivering the intervention, and (iii) feasibility that the intervention can be scaled up.

---

[3] This subsection is optional

12.   IE MANAGEMENT (E,R)

(All tables)

12.1 EVALUATION TEAM AND MAIN COUNTERPARTS

- Provide list of all IE team members with their position, affiliation, and responsibilities (including lead researcher, other research team members, and all project staff involved in the IE work, and main implementing agency counterparts).

Table 2. IE Team and Main Counterparts

Name	Role	Organization/Unit
	Principal investigators (specify Lead Researcher)	
	Other IE team members (specify IE TTL and Field Coordinator)	
	WBG Project staff involved in the IE (if the IE is related to a WBG project, specify Project TTL)	
	Main implementing and policy counterparts	

12.2 WORK PLAN AND DELIVERABLES

Table 3. Milestones, Deliverables, and Estimated Timeline

Milestones	Deliverables	Completion Date
Peer-reviewed Concept Note	Methodology note	May 15, 2013
Data collection plan and pilot	TORs Questionnaires	
Data collection (Baseline)	Cleaned data Dictionaries	
First data analysis	Presentation Data file Do files Baseline report	

Implementation of intervention aligned to evaluation	Rollout plan Monitoring reports verifying treatment and control status	
Follow-up data collection plan	TORs Questionnaire	
Data collection (Follow-up)	Cleaned data Dictionaries	
Final report and policy notes	Technical note Policy note Data file Do files	
Dissemination of findings	Presentations	May 30, 2017

12.3 BUDGET

(1 paragraph)

- Present total budget and disaggregated by staff time, data collection, and travel. Include all sources of funding, both Bank-executed and client-executed (BB resources, trust fund and grants, FBS, EFO, project financing for the IE, such as data collection, and other client financing). Estimate and include all research/staff time (not only the time charged).

Table 4. Total Budget per Category

Category	USD	%
Staff		
STC		
Data Collection		
Travel		
Total		

- Attach detailed budget (see excel file template).

13. PLAN FOR USING DATA AND EVIDENCE FROM THE STUDY

(1 paragraph)

- Describe communication, participation, and dissemination strategy (potential users of findings, media channels) at all stages of the IE (design, baseline analysis, mid-corrections, follow-up analysis, and final results).

REFERENCES

APPENDIX

i2i INDICATORS OF IE INFLUENCE ON PROGRAM/POLICY

Indicator	Definition	Example of Specific Output (required)
Quality of Policy Design		
Rationalized policy design (Yes=1, No=0)	IE improved design based on clear understanding of the underlying theory of change (causal links between the intervention components and the outcomes) and highlighted areas of uncertainty and critical assumptions.	The IE helped the [matching grant project] develop better communication and technical assistance strategies to address [common take-up/participation issues]
Introduced structured learning (multiple treatment arms) (Yes=1, No=0)	IE included multiple treatment arms to determine causal mechanisms, and compare the effectiveness of different interventions in practice	The IE introduced [alternative drug supply chains] to improve delivery of [medicine to the facilities in Zambia]
Adopted solution from existing IE results (Yes=1, No=0)	IE exposed program providers to existing evidence across sectors and/or countries and affected agreement on what components of intervention might work, and what might need to be removed using existing evidence.	The IE on [Malawi's gender reservation] was adopted by [Mozambique one year after this IE was designed]
Quality of Implementation		
Increased take-up (Yes=1, No=0)	IE improved information and/or incentives that resulted in increased proportion of people that use a program (out of those targeted by the program.)	The IE motivated [the court to enforce the digitalization of court cases] and increased the proportion of [judges moving from paper-based to computer-based processes in Senegal]
Improved delivery (Yes=1, No=0)	IE ensured that the treatment reached the treated in an effective and timely manner. IE clarified what the treatment entailed, to whom it will be delivered, when and where it will take place and provided feedback to the implementation agency to act upon differences between planned and executed treatment.	The incentives provided to increase adherence to treatment also increased [computer use by judges in Senegal]
Improved timeliness of implementation (Yes=1, No=0)	IE helped create a better design that improved efficiency of implementation, translating into better planned disbursement schedule and faster speed of disbursement. IE field coordinators followed/facilitated implementation.	A technical working group, formed as part of the IE [used to define changes in the regulatory framework for health facility inspections in Kenya] helped speed up project implementation.
Quality of Data		
High-quality baseline survey (Yes=1, No=0)	IE provided a full high-quality baseline survey (including covariates, with sufficient sample size, and representative of policy-affected population) available to policymakers and researchers creating or building on instruments for policymaking even before the IE started.	

High-quality follow-up survey(s) (Yes=1, No=0)	IE included comparable follow-up panel data for project.	
Improved administrative data (Yes=1, No=0)	IE data requirements for planning of data collection rounds strengthened the M&E function and reporting of key indicators.	The IE data requirements [convinced the Minister of Education to introduce unique school identifiers] that improved their administrative data.
Availability of Data as a Public Good (Yes=1, No=0)	The location of IE project data is known and accessible or it was uploaded in micro-data catalog.	
Capacity Building		
Client participated in IE workshop(s) (Yes=1, No=0)	The client attended IE workshop(s) and was connected to a broader global network of practitioners, policymakers, and experts in a set of relevant areas.	Rio, 06/2011 - DIME-FPD Impact Evaluation of Finance and Private Sector Development.
Baseline results discussed with client (Yes=1, No=0)	If baseline results were discussed with client please specify the client type, how it was discussed with the client, and the location/event of the discussion.	The IE team held a [workshop, meeting, VC] to present baseline findings on [Date]
IE results discussed with client (Yes=1, No=0)	IE analysis and results were discussed with the client to understand their policy relevance and application. Please specify the client, how it was discussed with the client and the location/event of the discussion.	The IE team held a [workshop, meeting, VC] to discuss IE results on [Date]
Training provided for data analysis (Yes=1, No=0)	IE improved skills of local institutions and/or staff to develop and implement IE, general monitoring and other data analysis independently (through discussions, technical assistance, workshops, and other training channels.)	The IE field coordinator led trainings on CSPro, Stata and SPSS to the Minister of Agriculture, allowing the team to use the data for their programming and improve their data-entry function.
Quality of Policy Decisions		
Baseline informed policy design/implementation (Yes=1, No=0)	IE baseline data was used by governments and other stakeholders to stimulate policy dialogue and/or help identify problems and solutions.	The IE identified [incentive problems in the size of women groups in Ethiopia] and helped project [introduce specialization in group composition]
Adopted the results of testing causal mechanisms or packages based on the IE (Yes=1, No=0)	IE evidence from experimental testing of alternative mechanisms or packages was used by governments or other stakeholders to inform policy decisions.	The IE identified more efficient [supply chain system] and government agreed to scale up at [national level in Zambia]
IE results were used to motivate scale-up/scale-down of policy at national level (Yes=1, No=0)	IE results reported success (or insufficient) impact of the intervention in achieving desired outcomes and were used by governments and/or other agencies/stakeholders to motivate scale-up (scale-down) of policy at national level.	The IE tested the [financial literacy program] and the [government] decided to scale up/scale down at the [national level]
Quality of Dissemination		
Number of presentations to policymakers of IE results	The number of presentations on the IE results given to policymakers.	
Number of presentations to academics of IE results	The number of presentations on the IE results given to academics.	

THEORY OF CHANGE EXAMPLE

Inspections Regimes in Health Care and their Impact on Patient Safety Standards and Quality of Care in Kenya

The evaluation targets all health facilities in three counties of Kenya (around 1,000 or 10% of the total number of health facilities in the country). It will have three arms, each of which is expected to affect quality of care and patient safety for all the population in their catchment areas. The details on the treatment arms are discussed in Section IV. The control group will be the "Business-as-usual" Low-intensity Health Inspections Regime, the first treatment arm will be a High-intensity Public Health Inspections Regime with Public Disclosure of the health facility's patient safety score, and the second arm will be a Private System of Supportive Supervision.

There are three common components across these arms: (1) a regulatory framework accompanied by clear guidelines on the minimum patient safety standards that facilities are expected to comply with; (2) a monitoring system to track compliance with minimum patient safety standards over time and; (3) a scoring and information report card system to publicize health facilities' compliance with minimum patient safety standards. Each treatment arm will include 2 or 3 of these components. Intervention activities are defined by these three components as follows:

(1) **A regulatory framework accompanied by clear guidelines on the minimum patient safety standards:** Activities in this component include the streamlining of the Checklist, the development of a detailed implementation manual, a score system to grade health facilities and a system of warnings and sanctions for non-compliant HFs. All these activities have been taking place during the last several months, from a process that started some years ago (see Checklist Process Note in Appendix 2 for further details). The new regulatory framework is schedule to be completed by the end of June 2014 (See Draft of Checklist in Appendix 3).

(2) **A monitoring system to track compliance with minimum patient safety standards:** This component includes activities related to the definition of the parameters for the inspections and the supportive supervision, including who the inspectors/supervisors will be, how often the inspections/supportive supervision will take place, and the system to enforce the warnings and sanctions developed in (1). Several of these elements have been defined, and a planning meeting with the participant counties and the MOH is schedule for August/September 2014.

(3) **A scoring and information report card system to publicize health facilities' compliance with minimum patient safety standards**: The scoring system is part of the regulatory framework from (1), and the report card system will be developed consistent with that scoring in a way that effectively communicates to patients the status of the health facilities.

Figure 1 presents a simplified theory of change behind the broad intervention to be evaluated, including its main components/inputs, activities, outputs, and the hypothesized causal chain to select outcomes of interest.

The main assumption behind this theory of change is that the combination of two or more of these components leads to inspection/supportive supervision systems that provide incentives for health facilities to comply with/increase patient safety standards (i.e., clear rules of the game, good information systems, a working system of warnings and sanctions, an effective supportive supervision, as well as a consumer liability system through the scorecard grading that create costs/benefits that incentivize health facilities to improve patient safety).

For instance, activities and outputs from component (1) are expected to affect the knowledge of the health facilities in terms of quality and patient safety, which is a necessary (although not sufficient) condition to affect compliance with the standards. Activities and outputs in component (2) are expected to directly affect compliance by creating incentives (and costs) for noncompliance. Finally, the third component is expected to affect consumer demand, which in turn may lead to changes in provider behavior and a reallocation of demand to facilities with higher patient safety scores In the long-term, better quality of care contributes to improvements in the health outcomes of the population served by health facilities.

We examine the impact of the interventions on a triad of measures that takes into account our setting with public and private providers: Quality and Patient Safety in the market, Quantity in the market and Prices in the market. The first set of outcomes are intermediate outcomes in the continuum of patient safety and quality of healthcare services measured through (A) compliance with a quality of care and patient safety checklist; (B) compliance with key infection control measures in doctor-patient interactions; (C) accuracy of diagnosis and treatment as evaluated through the use of standardized patients; (D) prevalence of unnecessary or harmful medication, and (E) prevalence of substandard drugs. Consumer and provider behavior are further measured through (A) patient loads in different health facilities and (B) prices charged to patients for consultations and laboratory procedures.

Figure 1. Theory of Change of a High-stakes and High Consumer Liability Health Inspection Regime