Appendix 1. Definitions of important terms.

Additionality
The difference between the value of an outcome after the implementation of a policy, and its value in a counterfactual scenario in which the policy is not enacted.

Analysis
A step in the research process that involves describing and then making inferences based on a set of data.

Baseline
A measurement of an outcome made prior to an intervention, which is then compared with measurements made after the intervention in order to evaluate the effects of the intervention.

Cascade effect
The indirect effects that an intervention has on an outcome in areas other than that for which the intervention was intended.

Case study
An observational study that involves the in-depth study of a single, or relatively few, observations.

Case study meta-analysis
A quantitative synthetic study that uses content analysis to systematically codes findings from previous case studies in order to compare them with a new dataset.

Categorical variable
A variable with a range of possible values that cannot be meaningfully ordered.

Census
An exhaustive sample that contains every member of the population of interest.

Cluster sample
A sample collected by first dividing a population into (frequently geographically defined) clusters, selecting a sample of these clusters, and then selecting samples of observations within each cluster.

Concept
A noun or adjective of theoretical/scientific significance that has been given a specific definition by a research community.

Congruence testing
A deductive qualitative analytical strategy of testing a theory by comparing the observational expectations it generates for a case with the observations made on this case.

Confounding variable
A variable that is associated with both a dependent variable and an independent variable and is thus a threat to causal inferences made regarding these two variables.

Content analysis
The process of coding qualitative information and data to produce quantitative data.

Convenience sample
A sample of observations that is selected primarily or exclusively based on the accessibility of these observations and the convenience of accessing them.

Correlational study
An observational study that involves the (usually statistically) comparative analysis of a large number of observations.

Covariation
A relationship between two variables in which an increase in one is associated with an increase or decrease in the other.

Counterfactual
An alternative scenario to which a realized scenario is compared in order to evaluate the significance of a causal factor that changes across the scenarios.

Cross-sectional comparison
A comparison of multiple observations, each of which is a measurement of a different entity at the same point in time.

Deduction
The process of developing testable hypotheses as the observational implications of a theory, and testing these hypotheses, and thus the theory, with empirical data.

Deductive validity
The accuracy with which a general principle or theory is applied to a specific case or context.

Dependent variable
A variable that is viewed as an outcome to be explained.

Direct observation
Measurement strategy in which the researcher directly observes the subject of observation, and either directly (via video or audio) records the subject, or records the values of qualitative or quantitative variables describing the subject.

Ecological fallacy
Inaccurately assuming that the characteristics of a population or group are representative of subgroups within that population or group.

Ecological validity
The accuracy with which findings from a highly controlled project, usually an experiment, can be generalized to more complex, real-world environments.

Embedded case study
A study that combines a case study of one unit of analysis, as well as a correlational study of a unit of analysis nested within the cases study.

Endogeneity
A situation in which a supposed dependent variable causes an independent variable to change.

Environmental History
A largely qualitative and historical approach to analysis that examines the historical relationship between humans and their natural environment.

Ethicality
The extent to which a research project upholds important ethical standards.

Ethnography
A highly inductive, fieldwork-based approach to social science that focuses on developing a thick understanding of a particular culture.

Experiment
A study that randomly assigns observations to control and treatment groups in order to isolate the effects of the treatment on the treatment group.

Expert sampling
A type of purposive sampling in which human respondents are identified based on their expertise in a subject matter relevant to the research project.

External validity
The extent to which characteristics of a sample can be generalized to describe the larger population or other related population (also called internal validity).

Factorial experiment
An experiment with more than one treatment, and subsequently multiple distinct subgroups to reflect each possible combination of the presence or absence of each treatment.

Feasibility
The extent to which a research project can be feasibly accomplished given the resource constraints facing a research team.

Focus group
A data collection strategy that involves a researcher holding a meeting with multiple respondents at once in order to engage with them and observe their interactions.

Framework
A set of concepts along with a set of statements describing their relationships.

Geographic information systems
Hardware, software and analytical operations designed to collect, process and analyze (primarily vector-based) spatial data.

Hypothesis
An observational implication of a theory. A statement that describes an empirically observable pattern that would be expected if a theory were correct.

Independent effect
An effect that an independent variable has on a dependent variable irrespective of the values of other variables.

Independent variable
A variable that is viewed as a cause of an outcome.

Induction
The formation of general principles or theories based on patterns or regularities found in a set of data.

Inductive theory-building
The process of inferring generalized relationships among a set of variables based on the (qualitative or quantitative) analysis of a particular case or set of observations.

Inference
A conclusion that explains a set of data by combining the data with something else, such as prior knowledge, a theory or model, or set of assumptions.

Informal interview
Occurs when researcher talks informally with subjects without any structured way of guiding the discussion or recording data.

In-person instrumentation
The use of a technological device to record data about an environment, or to take samples from this environment.

Interaction effect
The effect that two independent variables have on a dependent variable based on a non-additive interaction between them.

Internal validity
The validity of an inference connecting two or more variables in a causal relationship.

Interval/ratio variable
A variable with a range of possible values that includes a set of numeric values that can be compared in absolute terms.

Leakage
A process in which forbidding certain behaviors or outcomes in one jurisdiction creates incentives for these activities to spread elsewhere. This complicates causal inference.

Linear relationship
A relationship between an independent variable and a dependent variable that doesn't change in nature or magnitude across the range of either variable.

Longitudinal comparison
A comparison among multiple observations, each of which is a measurement of the same entity at distinct points in time.

Measurement validity
Quality of a variable based on (1) the fidelity of this variable to the concept it operationalizes, and (2) the accuracy with which this variable is measured to produce data.

Mediator variable
A variable that serves as a proximate cause of a dependent variable by mediating the effects of an underlying independent variable on this dependent variable.

Mid-range theory
A theory that is relevant for an identifiable set of cases or type of case, but not all cases.

Model
A theory, or set of theories, expressed in a formal language (e.g. with graphs, mathematics).

Moderator variable
A variable that affects the strength of a relationship between two other variables, thus producing an interaction effect.

Multistage sample
A sample that is obtained via two selection procedures rather than just one.

Narrative path analysis
A qualitative analytical method that explores the historical path that a system has taken and how this path affects the system's current situation and the prospects for future dynamics (rigidities, adaptations).

Natural experiment
An observational study that mimics an experiment, in which some observations are naturally exposed to a condition, while other highly similar observations are not.

Necessity
A condition in which an independent variable is required for a dependent variable to be present.

Negative feedback
The self-dampening effect of a relationship between two variables in which the first positively affects the other, and the other negatively affects the first.

Negative relationship
A relationship in which an increase in an independent variable causes a decrease in a dependent variable.

Network analysis
Analysis of a system as a network of nodes connected by links. Involves estimation of network-level and node-level properties and attempts to associate them with important outcomes.

Nonlinear relationship
A relationship between an independent variable and a dependent variable that changes in nature or magnitude across a particular threshold within the range of each variable.

Objective data
Data obtained without the involvement of human respondents, produced by the use of some physical data collection device or direct observation.

Observation
An instance of a unit of analysis or observation (e.g. a resource user) that is measured and analyzed.

Observational study
A non-experimental study that does not involve any active intervention or control on the part of the researcher.

Operationalization
The process of constructing a variable out of a concept by assigning it a range along which it can vary.

Ordinal variable
A variable with a range of possible values that can be meaningful ordered but cannot be compared in absolute terms.

Participant observation
A data collection strategy in which the researcher becomes an active participant in a study system in order to understand complex, day-to-day nuances in that system which otherwise may be hard to predict or understand.

Participatory rural appraisal (PRA)
A set of techniques that build on rapid rural appraisal by formally incorporating community stakeholder input into the appraisal process.

Permanence
The ability of a policy intervention to maintain its effects over time, particularly after the intervention itself has ceased.

Population
A set of observations of theoretical interest, about which the researcher wants to tell a scientifically rigorous story.

Positive feedback
The self-reinforcing effect of a relationship between two variables in which each has the same effect on the other.

Positive relationship
A relationship in which an increase in an independent variable causes an increase in a dependent variable.

Practical importance
The extent to which a research question and associated research project have some practical importance to society and the environment.

Primary data
Novel data collected by a researcher for a specific project.

Proportional sample
A stratified sample that has numbers of observations in each strata that reflect their proportions within the larger population.

Proximate cause
A cause that is in some way closest to an outcome of interest.

Purposive sample
A non-random sample that is obtained by purposively selecting observations from a population. Usually used in small-n research.

Qualitative analysis
An analysis of non-numerical data, usually either via content analysis to create quantitative data, or inferences made via direct observation and experience.

Qualitative comparative analysis
Estimation of the necessity and sufficiency of combinations of factors to produce an outcome.

Qualitative literature review
A non-quantitative synthetic study that summarizes findings from a particular research program or discipline.

Qualitative modeling
The process of developing a qualitative model of a system that divides it up into constituent components and describes the relationships between them, without quantification.

Qualitative variable
A variable that can take on any text value.

Quantitative analysis
An analysis that examines the associations among quantitative (categorical, ordinal, and interval/ratio) variables.

Quasi-experiment
An experiment in which the assignment of observations to control and treatment groups is non-random.

Random sample
A sample that is obtained by randomly selecting observations from a population. It is usually, but not necessarily, representative of that population.

Rapid rural appraisal (RRA)
A set of multidisciplinary techniques primarily conducted by development professionals to expediently collect data in rural areas by balancing between formal surveys and completely unstructured interview approaches.

Reliability
The consistency with which variables are measured across data collectors.

Remote instrumentation
Measurement method in which a researcher remotely manages a data collection technology that records features about the subject of observation.

Remote sensing
Hardware, software and analytical operations designed to collect, process and analyze (primarily raster-based) spatial data.

Representative sample
A sample that is representative of the population from which it was selected. This is needed in order to generalize findings regarding the sample to the larger population.

Research design
The type of study (e.g. correlational, experimental) that a research project employs to address its research questions.

Research question
A question of some practical and scientific interest that may be descriptive or causal, in which case it generally asks what factors affect an important outcome.

Sample
A subset of observations from a population that is selected for measurement and analysis.

Sample size
The number of observations within a sample.

Sampling bias
A bias that results from a sampling strategy that systematically selects for particular types of observations more than others, leading to a non-representative sample.

Sampling frame
A list of observations within a population that are available for sampling.

Sampling strategy
The process of selecting a sample of observations from a population of theoretical interest.

Secondary data
Data collected previously for different research projects that are used by a researcher to address a new question.

Self-administered survey
Measurement strategy in which human subjects fill out questionnaires without the presence of the researcher.

Semi-structured interview
An interview in which the researcher conducts an interview with the help of an interview guide.

Snowball sample
A sample that is obtained by identifying subsequent observations based on interactions with a preliminary set of observations.

Spillover effect
The process whereby an intervention applied to one set of observations ends up affecting other observations, complicating causal inference. Also called diffusion.

Spurious relationship
A correlation or association between two variables that does not actually indicate a causal relationship between these variables.

Statistical analysis
Calculations applied to quantitative data to descriptively summarize a sample and to make inferences about the population from which the sample is drawn.

Statistical validity
The validity of statistical inferences based on the satisfaction of important supporting statistical assumptions.

Stratified sample
A sample obtained by dividing a population into subgroups and sampling within each one of those subgroups.

Structured interview
Occurs when a researcher conducts an interview with the help of a written questionnaire that is filled out as the subject responds to the questions it contains.

Subjective data
Data obtained by eliciting the perceptions of human respondents.

Sufficiency
A condition in which the presence of an independent variable guarantees the presence of a dependent variable.

Synthetic study
A study that relies exclusively, or at least primarily, on secondary data to synthesize findings from a set of existing research projects, results, or case studies.

Systematic review
A synthetic study that qualitatively and/or quantitatively analyzes secondary data from a set of studies in order to examine the effects of an intervention.

Theory
A statement that describes (1) a (causal) relationship between two or more concepts, and (2) a mechanism by which this relationship occurs.

Triangulation
Comparing the results of different methods in order to check whether they produce similar or different results.

Type 1 error
Inferring that a pattern or cause exists when it does not.

Type 2 error
Inferring that a pattern or cause does not exist when it does.

Underlying cause
A cause that affects an outcome via a proximate cause.

Unit of analysis
The object of an analysis. A category about which a research question is posed, instances of which are frequently compared to test hypotheses.

Unit of measurement
The unit, such as meters or pounds, which is used to count or measure an interval or ratio-level variable.

Unit of observation
A data source that is used to obtain information about a unit of analysis. Can also be a unit of analysis.

Unstructured interview
Occurs when a researcher conducts an interview without any supporting materials.

Variable
An operationalized concept that has been assigned a level of measurement and range of possible values.