Notes on *Designing social inquiry: scientific inference in qualitative research* (King, Keohane, Verba 1994)

Ian Stewart

June 26, 2017

1 The *Science* in Social Science

- In social science, quantitative and qualitative researchers have different styles but the same underlying approach to understand social phenomena.

- Quantitative researchers seek to build statistical models that abstract the causes of large-scale phenomena in a replicable way, while qualitative researchers seek to unpack the specific details of small-scale phenomena.

- The two styles of research can be combined productively, such as using quantitative methods to test causal effects and later examine qualitatively the specific details behind any spurious causal effects.

- Building abstract and formal models of social phenomena helps to illuminate causal effect in a scientific way, even if the methods are qualitative. The conclusions that researchers draw may not always be certain, but the validity and certainty of the conclusions is assured by scientific discipline (not dogma).

- Scientific research has four characteristics: (1) inference as a goal, (2) public procedures, (3) uncertain conclusions, (4) method as content.

  1. Research should not accumulate facts but should make explanatory inferences about the world that go beyond the particular situation of study. These inferences may be descriptive, using observed knowledge to learn about unobserved knowledge, or causal, using observed knowledge to understand causal effects between variables.

  2. The procedures used to achieve inference should be public so as to allow public judgment of scientific validity and replicable study.

  3. The estimates of causal effect provided by research should account for inherent uncertainty in the real world, as it is impossible to know anything perfectly.

4. Methods and rules, rather than specific details, are the main contributions of scientific research.

- The apparent complexity of a phenomenon depends on the complexity of the underlying theory, and even complex phenomena that resist easy explanation can provide insight on the methods used to study such phenomena (i.e. methods as content).

- One method to simplify a complex phenomenon is to classify it as a particular kind of event in order to generalize and make comparisons to similar events. Another method is to compare the phenomenon with a counterfactual and to assess their respective outcomes. In studying the mass extinction of dinosaurs, one may assess the relative outcomes of a proposed meteor collision, as compared to alternatives, by gathering evidence that supports the proposed cause or contradicts the counterfactual. This requires specifying the causal theory behind the phenomenon and improving the data used to assess the outcome.

- The process of social science research is not deterministic but flexible, with a research design that can be revised according to a stable set of expectations. Such revision can result from data that do not fit the original theory, and research is often facilitated by data collection before theory proposition.

- Research design consists of four components that are developed in tandem: (1) research question, (2) theory, (3) data, and (4) use of data.

  1. Deciding on a *research question* often requires creative intuition and may be driven by personal interest, such as a citizen of an postcolonial country wanting to research a similar kind of country. The research question should have consequences for the real world and should contribute to an identified body of scientific literature, thus advancing the field. Choosing a research question can be guided by prior untested hypotheses or assumptions in the field of study, or by similar hypotheses in a related field of study.

  2. Theory proposes an answer to the research question and a reason for the answer, while remaining consistent with prior evidence. A theory should be falsifiable, or capable of being proven wrong, and should allow a researcher to generate as many observable implications as possible in order to maximize the change of proving it wrong. A theory should also be concrete and stated precisely. A theory does not have to simple (parsimonious), if the situation under study is not simple. While gathering data before proposing a theory can be useful, it can also lead to overfitting and may not provide useful evidence for the theory. A researcher can always revise the theory to be less restrictive but if the theory is changed to be more restrictive, more data should be collected to test the new theory.

3. Data may be collected without a clear research question in mind, but following guidelines for collection can help ensure the data's future utility. First, researchers should record the process of data collection and processing. Second, researchers should collect data from as many observations as possible to ensure that the theory is addressed by the maximum amount of evidence available. This can include collecting observations of the dependent variable from new circumstances or additional dependent variables to test different outcomes of the theory. Third, researchers should maximize measurement validity, or make sure that they are measuring what they assume they are measuring. Fourth, the data collection methods should be reliable and produce similar results when applied at different times. Fifth, the data and analysis should be replicable, and the reasoning behind all research choices should be clear to other researchers.

4. Researchers can improve existing data by (1) minimizing *bias*, or variability, in inference in data collection and analysis, which includes avoiding pitfalls such as selection bias; and (2) maximizing *efficiency*, or information relevancy, which includes examining the data from multiple perspectives such as different levels of aggregation.

- The rest of this book covers four main themes: (1) connecting theory and data through observable implications, (2) maximizing theoretical leverage; (3) reporting uncertainty; (4) skepticism.

    1. The observable implications of the theory under study will shape the data collection process, since the implications need to be proven by the data.

    2. Science that tests a theory's hypothesis with as few variables as possible can be said to maximize the leverage over the hypothesis. Researchers should enumerate all possible observable outcomes of their hypothesis, including at different levels of analysis (e.g., macro- vs. individual-level), to assess the widest range of information available and maximize the leverage of the hypothesis.

    3. Researchers should measure and report the degree of uncertainty in their estimates of causal effect, because the social world is probabilistic.

    4. Researchers should be skeptical and evaluate alternative hypotheses when assessing a potential causal effect, such as a reversal of cause and effect. In this way, causal inference is a continuous process of refining the observed phenomenon through repeated approximations.

2 Descriptive Inference

- In studying a particular phenomenon, social scientists look to describe and explain the phenomenon.

- Studies may seek to provide an explanation of a particular phenomenon or a general class of phenomena.

- Accurate interpretation of the phenomenon often requires prior knowledge about the surrounding culture (e.g. winks vs. twitches), and this interpretation is necessary to abstract general knowledge from the specifics of unique events.

- Finding relevant observations is hard!

- The world is so messy that we must often simplify our descriptions to fit an interpretable model.

- Comparative case studies can provide causal insight if the same variables are collected and compared systematically across cases.

- Scientists gain insight on hypotheses and theories through inference, "the process of using the facts we know to learn about facts we do not know" (46). However, the hypotheses and theories may evolve over the course of study.

- Developing an algebraic model to describe observational data allows scientists to test hypotheses more precisely.

- When collecting data, it is important to document how the data was created and processed.

- The process of determining variables to study will vary depending on the level of granularity and detail required, e.g. whether to binarize a continuous variable.

- Extracting the essential patterns from data requires statistics that (1) focus on the outcomes to be explained and (2) simplify the data for the audience.

- Descriptive inference helps us understand an unobserved phenomenon, by distinguishing the systematic factors that can explain the phenomenon from the unsystematic factors. For example, a study of voting results may consider a wide range of factors, from voter turnout to weather, and would need to determine which factors contributed reliably to the outcome.

- Since we often can only observe one outcome out of many possible outcomes, we distinguish the realized variable (actual) from the random variable (hypothetical). A goal of inference is to understand the systematic features of the random variable, such as expected voter turnout in an average election, and most work begins by assuming all variables are nonsystematic.

- We can understand the random variable as the result of something between a stochastic or deterministic process ("Probabilistic World" vs. "Deterministic World").

- Judging the methods of descriptive inference relies on the criteria of unbiasedness, efficiency, and consistency.

- A method of inference produces *unbiased* estimates if the estimates vary nonsystematically (i.e. without bias).

- *Bias* may refer either to the data, such as data that are collected under a biased set of conditions, or to the estimator, such as the method of computing the average. A formal example of an unbiased estimator is the sample mean, which is guaranteed to produce estimates close to the ideal mean μ based on the definition of expected value.

- An *efficient* estimator has a relatively small variance across multiple replications (e.g. compare the variance of a sample with few observations versus a sample with many observations).

- A *consistent* estimator has smaller variance, and higher accuracy, with more observations.

- More data isn't always better! It is harder to account for confounding factors with a larger sample size, such as factors that throw off measurements (e.g. biased government data), which may result in inefficiency. In this sense, it may make sense to conduct an in-depth case study of one observation rather than a large-scale shallow study of many observations.

- The most productive study design may require coordinating between a small-scale study and a large-scale study, since a small study's careful analysis can correct the bias of the large-scale study's estimator.

- Choosing the right study design can be guided by the balance between bias and efficiency, e.g., studying a large left-leaning voting population would produce a more biased but more efficient estimation than a small, evenly balanced population.

- Metrics such as the mean squared error formalize the tradeoff between bias and efficiency: for estimator g and parameter γ, $MSE(g) = V(g) + E(g - \gamma)$.

- Connection! bias : accuracy :: efficiency : precision

3   Causal inference

- Descriptive inference can only do so much for social science, and causal inference can help make stronger claims as long as the uncertainty of the claims is properly addressed.

- Causal inference analysis seeks to isolate the systematic features of the *realized causal* effect of the *treatment* on the *outcome*, i.e. the difference between the actual and *counterfactual* outcome, while accounting for *control* variables.

- The control variables should include all variables that are not consequences of the treatment variable: e.g., for predicting the success of an incumbent candidate, control for past performance of the political parties but not incumbency consequences such as name recognition.

- The fundamental problem of causal inference is that a causal effect can never be observed with certainty because the counterfactual can never be perfectly realized.

- Causality is a systematic feature of any random variable, such that an unobserved causal effect is a realization of the random causal effect. For instance, the causal effect of incumbency on a particular election is the realization of a random causal effect across elections in general.

- Example causal effect calculation: Mean causal effect of A vs. B for unit $i = E(\text{Random Causal Effect for unit } i) = E(Y_i^A - Y_i^B) = \mu_i^A - \mu_i^B$

- Causal inference may study both the mean and variance of the causal effect, for instance to better understand the risks behind a certain treatment.

- Understanding causal effect does not require perfect understanding of the causal mechanisms, because the effect of the treatment on the outcome will hold regardless of the different intermediate paths that it may take.

- Causal effects are not invalidated by multiple causality, or the interdependence of independent variables on the outcome, as long as the counterfactual situation is clearly defined. For instance, testing the effect of college graduation on job income would require holding seniority constant, e.g. comparing a graduate with someone who worked for four years instead of college.

- Understanding causality does not require an understanding of whether the causality is symmetric or asymmetric, i.e. whether removing the treatment returns the outcome to its original state (symmetric), because causal effect is hypothetical. However, performing causal inference may require accounting for symmetry.

- To estimate causal effects, we must ensure (1) *unit homogeneity*, i.e. all units within the same treatment/control group have the same expected value for the outcome; and (2) *conditional independence*, i.e. explanatory variable values are independent of the outcome variable values.

- The *constant effect assumption* is a weaker version of unit homogeneity, in which we assume that the causal effect is constant across units even if the expected values vary. For instance, the effect of incumbency on voting percentage may be constant (e.g. always +10%) even if the voting percentage varies across similar units.

- For *conditional independence* to hold, studies often rely on random selection to ensure that the explanatory variables are independent of the outcome variable, but random selection is not necessary as long as potential confounding variables are accounted for (e.g. accounting for conflict not just by looking at residential segregation but also the confounding factor of ideology).

- Formal causal model: where Y_i equals the outcome for unit i, X_i equals the treatment for unit i and β equals the difference in means between treatment and control ($\mu_i^T - \mu_i^C$), then $E(Y_i) = X_i\beta$.

- Similarly to estimators like μ, causal inference estimators are judged based on the criteria of unbiasedness and efficiency. For instance, we know that the least-squares estimate $b = \frac{\sum_i Y_i X_i}{\sum_i X_i^2}$ is an unbiased estimator of β based on the definition of expected value, i.e. $E(b) = \frac{\sum_i X_i E(Y_i)}{\sum_i X_i^2} = \beta$. Similarly, we know the efficiency of b based on $V(b) = \frac{1}{\sum_i X_i^2} \sum_i X_i^2 V(Y_i) = \frac{\sigma^2}{\sum_i X_i^2}$, which demonstrates that the estimator's variance increases with more variance in the explanatory value (σ^2) and decreases with a larger range of explanatory values ($\sum_i X_i^2$).

- A causal theory, i.e. an explanation for the cause of a phenomenon, should be internally consistent with respect to the set of hypotheses that comprise the theory (i.e. non-contradicting) and should adhere to the following criteria.

  1. Falsifiable: future study should be able to prove the theory wrong in order to advance scientific knowledge, but the theory can still serve future studies if it explains a significant portion of the world and may be modified, not discarded, in the face of contradictory evidence. In general, theories should be constructed to explain as much as possible with as few assumptions as possible (*parsimony*), but if a more complicated theory explains a much larger set of phenomena (*leverage*) then we should choose accordingly.

  2. Internally consistent: a theory should not generate contradictory hypotheses. Consistency can be guaranteed with mathematical *formal models*, which often make simplifying assumptions, but these models must be augmented to adapt to empirical study.

  3. Carefully selected dependent variables: the variables should (1) not change the outcome variable, (2) vary across observations, and (3) represent the full range of variation to be explained.

4. Concrete: the theory should deal with observable concepts that can be empirically evaluated. Abstract concepts such as culture should be handled through concrete indicators, despite a potential gap between the concept and the indicator, which the researcher must justify. The theory should also be described in concrete rather than vague terms, even at the risk of being wrong.

5. Maximally encompassing: a theory should explain as much of the world as possible, and researchers should be able to leverage a theory developed in a specific situation to other situations. This rule often conflicts with the "concrete" rule.

4 Determining what to observe

- Causal inference requires conditional independence and unit homogeneity (or constant effect), which may be violated by data limitations but can also be worked around through intentional selection of observations.

- An indeterminate research design reveals nothing about the causal hypotheses and can result from two situations:

  1. More inferences to make (i.e. more explanatory variables) than observations collected: the inclusion of too many explanatory variables precludes the researcher from making inferences because of the infinite possible combinations of causal values to be assigned to explanatory variables. This problem may be addressed through limiting the explanatory variables or collecting an extra, limited set of observations.

  2. Multicollinearity among explanatory variables: when two explanatory variables are perfectly correlated then no inferences can be drawn because there is insufficient variation in the variables. This problem can be addressed through collecting more observations that add variation to the data, changing the level of analysis, or limiting the explanatory variables (without omitting important variables).

- In large-n situations, random selection can reduce the risk of selection bias but is often impossible due to the indeterminately large space of possible observations (e.g. a comprehensive list of policymakers is likely to be biased in some way).

- In small-n situations, random selection can create bias if the wrong samples are selected (e.g., in 2-observation study with 3 ordinal values for outcome, have a $\frac{2}{3}$ chance of selection bias if the high and middle or middle and low values are selected).

- In small-emphn situations, non-random selection can create bias if the observations drawn do not cover the full range of possible dependent variable values, but this

problem can be avoided by supplementing the detailed case studies with more shallow, wide-scale observations of relevant variables.

- *Selection bias* results from researchers selecting observations that support the intended outcome based on their combination of independent and dependent variables instead of allowing variation in the dependent variable. A weaker form of selection bias results when the range of variation in the dependent variable is restricted (e.g. only studying , and allowing a wider range of variation in the dependent variable allows stronger causal claims to be made. All forms of selection bias lead to an underestimation of causal effect.

- Example of avoiding selection bias: in study of presidential involvement in foreign policy meetings, we can avoid problem of secret meetings (i.e. unobservable data) by comparing presidential action after meetings as well as presidential action in general, to determine prevalence of independent variable (e.g. threat of force) in the after-meetings data, therefore correcting for bias.

- More data doesn't fix selection bias! Social scientists need to also consider how external processes can induce bias. E.g., in studying relationship between endorsement and voting outcome, we need to consider how endorsement works - parties often ignore soon-to-lose candidates, thus selecting only positive outcomes.

- Selecting on the explanatory variable does not induce bias because it does not affect the dependent variable, although it may reduce the certainty of the causal effect. E.g., in studying the effect of discrimination on student performance, choosing schools with significant vs. little discrimination will not introduce bias as long as we consider the effect of the other explanatory variables. Including an irrelevant variable as a control can help avoid selection bias.

- Overestimation of causal effect can occur if the causal effect varies across observations and some observations have a significantly greater effect than others, which results in an artificially high average causal effect.

- Non-random selection of observations can include:

  1. Selecting on the explanatory variable, while controlling for other variables;
  2. Selecting on a range of dependent variable values, which may begin with an exploratory analysis on high/low dependent values then selecting a broader range of independent values without regard for dependent;
  3. Selecting on both dependent and independent variable values, e.g. if the independent or dependent variable has a skewed distribution, we need to account for as wide a range in the other variable as possible;

9

4. Selecting on constant values of independent or dependent variable, e.g. if the relationship between independent and one particular value of the independent is known from prior work (e.g. only studies of no-fertilizer farms), then selecting on the other value of the independent (fertilized farms) can reveal causal effect.

5   Understanding What to Avoid

- Reminder: unbiasedness means having an accurate estimate, efficiency means having a precise estimate.

- How do we ensure that our estimates are unbiased and efficient? By reducing possible sources of bias and inefficiency: (1) measurement error, (2) omitting explanatory variables, (3) including irrelevant variables, (4) endogeneity.

- Quantitative research is often more precise, qualitative research often more accurate. In both kinds of research we try to minimize *measurement error* that can lead to inaccurate or imprecise conclusions.

- The measurements that researchers make are shaped by their own research design, such as whether to classify government systems by categorical or ordinal values. A researcher should use the measure that is most appropriate for their theoretical motivations.

- Grouping error results from inappropriate discretization of a continuous variable, e.g. bucketing age into "young," "middle aged," and "old." Conversely, converting a categorical variable into an ordinal or continuous variable can create an inaccurate metric.

- Measurement error may be systematic or nonsystematic.

- *Systematic* measurement error refers to the consistent misestimation of measurements and can result from researcher bias toward a particular kind of data or from a flawed research design, such as self-reporting bias. However, systematic errors that affect all observations by the same constant effect cause no bias (e.g. overestimating the salary of all survey respondents by same amount). Accounting for systematic measurement error may require applying judgments from other studies to avoid bias toward data that supports the theory under study, e.g. having outside researchers code observations without knowing about the theory.

- *Nonsystematic* measurement error results from random fluctuation in measurement and has different consequences for explanatory versus dependent variables.

- Nonsystematic error in the dependent variable reduces the efficiency but not bias in the estimation of a causal relationship. Addressing this kind of error requires increasing the amount of information concerning the dependent variable, such as more accurate analysis of existing data or collection of additional data.

- Nonsystematic error in the explanatory variable can result in a weaker (closer to zero) apparent causal relationship. This means that an apparently weak positive or negative causal relationship may actually be stronger than expected after accounting for error in the explanatory variable.

- *Omitting* an explanatory variable will not cause bias if the explanatory variable is uncorrelated to the dependent or other explanatory variables, although if the omitted variable is correlated with the dependent but not the other explanatory variables then the causal inference will be less efficient. Formally, with explanatory variable X_1, omitted variable X_2 and dependent variable Y, the estimator b_1 is unbiased if X_2 has no correlation with Y or with X_1, i.e. if $B_2 = 0$ or $F = 0$ in $E[b_1] = \beta_1 + F\beta_2$.

- Omitting a correlated explanatory variable can cause the causal inference to produce an estimate higher than expected (positive correlation between omitted and explanatory) or lower than expected (negative correlation between omitted and explanatory). For example, if unemployment and political repression are negatively correlated, then omitting unemployment will cause political repression to have an apparently smaller effect on the dependent.

- Researchers can avoid the problem of variable omission by looking outside the existing data for additional possible variables and to control simultaneously for all variables (e.g. instead of sequential variable analysis). However, researchers should not control for explanatory variables that are consequences of the key causal variable: for example, adding a pre-voting poll question as a control variable ("do you plan to vote Democrat?") would reduce the apparent impact of the key explanatory variable, e.g. income, on the dependent variable of voting.

- Deciding what explanatory variables to include should revolve around a theoretical model rather than trial and error. More variables means less chance of omitted variable bias but also means the possibility of confusing the causal effect of the key explanatory variable.

- Example analysis: determining correlation between a person's education and their political participation. Including race as a control variable might reduce the apparent causal relationship but could also reveal the interaction between race and education (i.e. a valuable contribution), or could reveal that the control variable is more predictive of the dependent (i.e. that the key causal variable is not as important as originally hypothesized). We might also include sub-variables of education,

11

e.g. participation in civics courses, or variables that precede education, e.g. general intelligence.

- *Including* an irrelevant variable may result in research with indeterminate conclusions, and a higher correlation between key explanatory and irrelevant control variables results in less efficient causal inference (i.e. higher variance estimator). For example, estimating the correlation between repression and coup likelihood would be confused by including the irrelevant variable of relationship with prior colonial rule (e.g. breaking violently), because the number of observations in each repression-colonial rule category would be lower than otherwise (due to correlation) and would result in a less efficient causal inference.

- *Endogeneity* results from explanatory variables being the consequence rather than cause of the dependent variable, i.e. a reversed causal direction. For example, a politician's expected future vote may determine the service they perform during their time in office, rather than the expected opposite direction (service determines future vote).

- Researchers can handle endogeneity in five ways: (1) correcting a biased inference, (2) restricting the dependent variable to actual consequences of the explanatory variable, (3) treating it as an omitted variable problem, (4) restricting the observations to those without endogeneity problems, (5) restricting the explanatory variables to their exogenous parts.

  1. Correcting the bias involves estimating the causal effect of the explanatory variable, estimating the direction of bias on the dependent variable, and re-estimating the causal effect while taking the bias into account. For example, in the service-vote example, the correlation may be positive but if the bias is negative then the overall effect will be smaller than expected, therefore a corrected estimate would have a higher causal effect than expected.

  2. Restricting or "parsing" the dependent variable requires identifying and measuring only the consequential aspects of the dependent variable. For example, studying the service-vote correlation may require restricting the vote to only the percentage of vote affected by incumbency (i.e. a small percentage that is unlikely to correlate with the amount of service performed).

  3. Reformulating endogeneity as an omitted variable problem requires reconfiguring the dependent and explanatory variables to account for an omitted variable that is causally prior to both. For example, rather than fixate on the relationship between proportional representation and parliamentary fragmentation, a researcher might bring in the causally prior variable of social fragmentation to explain both.

4. Restricting observations to those without endogeneity problems often requires comparative analysis to determine which observations are exogenous. A researcher studying the effect of preconceived ideas on policy can compare situations in which the ideas may have been influenced by anticipation of policy (endogenous) with situations in which the ideas could not have been influenced by anticipation of policy (exogenous), e.g. comparing the adoption of Stalinist doctrine across Eastern Europe (endogenous) and China (exogenous).

5. Restricting the explanatory variable to its exogenous components entails a reconsideration of how each component relates to the dependent variable. In a study of the relationship between religious activity and political participation, a researcher may isolate the components of religious activity that precede political participation such as learning how to make a speech, which in turn are influenced by structural considerations such as church governance.

- In observational study as opposed to experimental study, endogeneity affects how observations are produced and how the explanatory variable is assigned. In studying a country's arms budget and its likelihood of going to war, the reverse causality plays a significant role in assigning values for the explanatory variable (i.e. higher likelihood of going to war produces larger arms budget). Therefore, if we cannot collect data randomly over a large number of samples, we should select observations and assign explanatory variables to avoid bias and inefficiency, or at least understand the magnitude and direction of bias.

- To control the research situation, a researcher must simultaneously handle measurement error, omitted or irrelevant variables, and endogeneity. One way of controlling these factors is by matching observations on control variables. In cases where it is not possible to match observations perfectly (i.e. a lot of cases), a researcher may rely on the difference in units over time. For example, comparing the policies of colonial countries before and after colonial withdrawal with control non-colonial countries can entail comparing their relative differences across time, i.e. difference-in-differences, and this procedure controls for lack of homogeneity across space/time.

- Matching procedures must be careful to account for omitted variables that account for more difference between matched units than the main explanatory variable alone, such as variables that have an extreme value for one of the matched units. A study of prisoners that matches the participants with non-prisoners should account for omitted variables that kept the non-prisoners out of jail, e.g. strong religious beliefs.

- Researchers often differ in whether units should be matched to maximize variability ("most different") or maximize similarity ("most similar"). Regardless of matching technique, research should rely on observations that provide a higher degree of leverage over the causal hypothesis.

13

- Methodology problems often happen in clusters, and researchers should be prepared to modify their research design to handle these overlapping problems.

6 Increasing the Number of Observations

- A study with only a single observation ($n = 1$) can be improved by subdividing the observation into multiple sub-observations.

- A single-observation or "crucial" case study can (1) provide support for a theory if it is an especially difficult case that the theory proves or (2) contradict a theory if it is an especially easy case that the theory fails to support. However, this is often an unlikely scenario because of (1) alternative explanations (the effect of different values in the explanatory variable), (2) measurement error in the variables, and (3) non-deterministic factors affecting the observed causal effect.

- One example of a crucial case study going wrong is reasoning by analogy, which seeks to compare a single pair of observations through analogy (Y is to country A as Z is to country B). The comparative approach relies on many observations, even if imperfectly paired, and provides stronger causal evidence.

- Assuming the same linear relationship that we have used throughout, the number of observations necessary for inferring the causal effect of causal explanatory variable x_1 relies on four factors.

  1. Higher values of fundamental variability, σ^2 i.e. variability in the dependent variable, require a larger sample size, and inefficient estimators, e.g. estimators with random measurement error, require a larger sample size.
  2. Lower uncertainty of causal inference, $V(b_1)$, requires a larger sample size, and some studies (e.g. new area) may be willing to tolerate more uncertainty in the inference in exchange for new insight.
  3. Lower relative collinearity between causal and control variables, R_1^2, requires a larger sample size, because stronger correlations between causal and control variables require more observations to determine the actual causal effect on the dependent variable. However, careful selection of observations to minimize collinearity can also produce accurate causal estimates.
  4. Lower variance of the values of the causal variable, $S_{x_1}^2$, requires a larger sample size. Again, careful selection of observations with a wide range of causal values can produce accurate causal estimates without a large sample size.

- These four factors yield the required number of observations

$$n = \frac{\sigma^2}{(1 - R_1^2)S_{x_1}^2 V(b_1)}.$$

14

- If the causal effect is not linear, the researcher should focus attention on the middle values, rather than the extreme values, of the causal explanatory variable, because these values may provide an unexpected amount of explanation for the dependent variable.

- Don't forget the main goal of social science! Rather than try to explain a single complex phenomenon, social scientists attempt to abstract causal effects from particular phenomena to test a particular hypothesis, which requires building case studies to test a causal effect rather than simply describing a case.

- In the case of a small-n study, a researcher should consider increasing the number of observations by (1) extracting similar observations from additional units, (2) extract new dependent variable data from the same units, and (3) extract new dependent variable data from additional units.

  1. Extending the research design to additional units may require taking a more cursory look at other observations outside the original scope (e.g. shallow analysis of other countries), dividing the original observations into sub-units (e.g. switch from nation to city), or studying the original units across time. Researchers should only choose new units if they are fit to test the replication of the hypotheses: in a study of federal tariffs, it would be inappropriate to choose state-level units. Researchers should also be careful to minimize dependence between the units, because dependence may reflect a confounding variable and may reduce the certainty of the causal effect. For instance, data from neighboring counties may reflect geographic interdependence, and data from across time may reflect temporal influence of past events on present conditions.

  2. Extracting new dependent variable data means looking for a different effect from the same cause, i.e. a new hypothesis. This can allow researchers to evaluate multiple effects of the same cause, such as studying both federal and state-level impacts of tariff policy.

  3. Extracting new dependent data from new units requires deriving new hypotheses and may apply to units at a higher or lower aggregate level. Researchers should use this method when they lack observations of a meaningful social process and need to adjust their hypotheses: in studying nuclear war, a researcher would probably need to look at threats of war rather than actual war. Changing the level of aggregation and the dependent variable often entails new explanatory variables, such as changing from a state level (policy) to an individual level (psychology). This change is referred to as *process tracing* or searching for evidence about the individual causal mechanisms behind a macro-level outcome. Researchers should trace causal mechanisms in such a way to give opportunities to refute the theory being tested, which includes choosing observations inde-

pendent of outcome (avoiding endogeneity) and not omitting key explanatory variables.

- Both quantitative and qualitative social scientists need to take careful consideration of research design and rely on logic when making a case for causal inference. *mic drop*