

Defining the inquiry

Defining the inquiry

An inquiry is a summary of a theoretical model. Suppose your theory is that X affects Y , but that the effect is bigger for units with higher values of a third variable C . One inquiry might be the average level of Y . A second might be the average treatment effect of X on Y . A third might be the difference in the effect of X on Y for units with high values of C versus units with low values.

In general, an inquiry is a summary function I that operates on the model M . When we summarize the model with the inquiry, we obtain an “answer under the model.” We formalized this as $I(m) = a^m$. You can think of the difference between I and a^M as the difference between a question and its answer. I is the question we ask about the model and a^m is the answer. Alternatively, you can think of I as the “estimand” (that which is to be estimated) and a^m as the value of the estimand.

In this book when we talk about inquiries, we will usually be referring to single-number summaries of the models. Some common estimands are descriptive, such as the means, conditional means, correlations, partial correlations, quantiles, and truth statements about variable in the model. Others are causal, such as the average difference in one variable when a second variable is set to two different values. You can think of a single-number inquiry as the atom of research question.

While most inquiries are “atomic” in this way, some inquiries are more complex than a single-number summary of the model. For example, the best linear predictor of Y given X is a two-number summary: it is the pair of numbers (the slope and intercept) that minimizes the total squared distance between the line and each value of Y . Why stop at two-number summaries? We could imagine the best quadratic predictor of Y given X (three-number summary) or the best cubic predictor (four), and so on. We could have an inquiry that is the full conditional expectation function of Y given X , no matter how wiggly, nonlinear, and nuanced the shape of that function – it could in principle be a 1,000 number summary of the model, or much more.

Complex inquiries can also be something along the lines of “should I reject M ” as a causal model of the world? A researcher might articulate a handful of important questions about the model that all have to come out a certain way or the model itself should be rejected. These complex inquiries are made up of a series of atomic inquiries – we’re interested in the sub-inquiries only insofar as they help us understand the real inquiry – is this model of the world a good one or not.

Kinds of Inquiries

There are two basic kinds of inquiries I about a model M , descriptive and causal. One way to think about the difference is that descriptive inquiries are about the nodes of a dag and causal inquiries are about the edges. A descriptive *inference* is a conclusion about a latent variable Y^* on the basis of a measured variable Y . The feature we seek to describe – our inquiry – is some summary of Y^* like its mean or perhaps its covariance with a second latent variable X^* . When we do descriptive research, we draw inferences about features of the nodes of the latent causal model M . A causal *inference* is a conclusion about the edge between two latent variables X^* and Y^* . You can’t just observe edges – you can see nodes but you can’t see edges.

Inquiry	Description	Inquiry	Description
$E[X]$	The average value of the variable X	ATE	
$V[X]$	The variance of X	CATE	
$\text{Cov}(X, Y)$		SATE	
		PATE	

Inquiry	Description	Inquiry	Description
		CACE / LATE	
		ATT	
		ATC	

Descriptive estimands

Descriptive estimands can also require inference, not simply measurement.

Simplest case level estimand:

$$X = 1$$

“Yes”

An example of a descriptive estimand is:

$$E_{i \in N}(Y)$$

$$E_{i \in N}(Y|X = 1)$$

This descriptive estimand is to be distinguished from the counterfactual estimand:

$$\Pr(Y = 1|X \leftarrow 1)$$

where \leftarrow is interpreted to mean that X is “set” to the indicated value.

- Moments; covariance

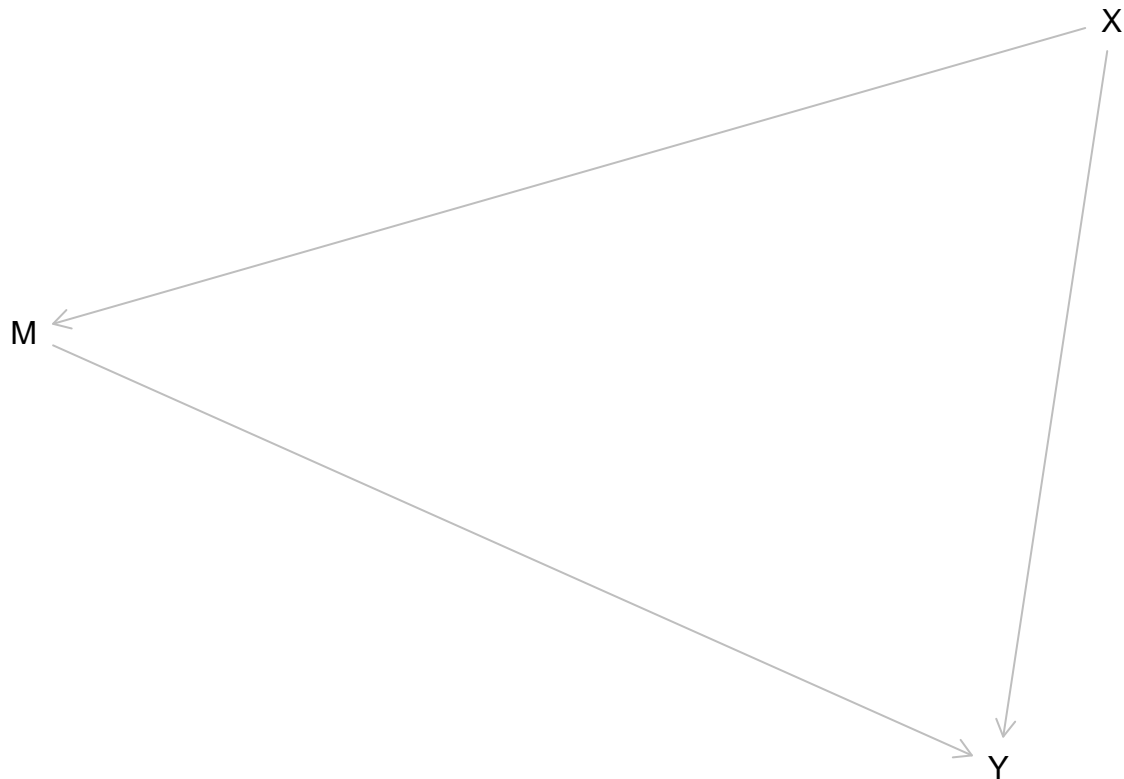
Example of a descriptive inquiry:

```
declare_population +
declare_estimand(COV = )
```

Causal Inquires

Causal inquiries are summaries that involve multiple potential outcomes. To describe these we will imagine a simple model, in which Y is thought to depend on X and M and M in turn is thought to depend on X . We represent this barebones model using a DAG in figure @ref(XMYgraph).

Inquiry	Description	Inquiry	Description
$E[X]$	The average value of the variable X	ATE	
$V[X]$	The variance of X	CATE	
$\text{Cov}(X, Y)$		SATE	
		PATE	
		CACE / LATE	
		ATT	
		ATC	



[Note, this section has a huge jumble of letters that will need to be harmonized!]

Causal inquiries like the average effect of A on B in a causal model like $A \rightarrow B \leftarrow U$ are helped enormously by good descriptive inference about the nodes A and B, but the focus is on the edge between them. You typically can't learn about the edge by doing descriptive inference on the nodes only. If we measure A and B and find that they covary, we can't be sure that $A \rightarrow B$ because it could be that $A \leftarrow U \rightarrow B$. This problem goes by the familiar phrase that "correlation doesn't imply causation," which is true and is a problem that can't be wished away. But "correlation doesn't imply causation" also kind of misses the point. The point is that you can't see causality because it involves counterfactuals, which are imaginary and unseen. You have to infer causality on the basis of design.

Mediation Inquires:

Consider $X \rightarrow B \leftarrow A$ and $X \rightarrow A$. In this DAG A is a mediator. There are three causal inquiries:

$X \rightarrow B$ $X \rightarrow A$ $A \rightarrow B$.

If we randomize X, we can learn about $X \rightarrow A$ but not $X \rightarrow B$, because $X \rightarrow B$ is confounded. We obviously don't learn anything about $A \rightarrow B$ because for that we'd need to randomize A. Mediation analysis tries to get all three for the price of one random assignment. At best we get $X \rightarrow A$ and the total effect of X on B, which is not just $X \rightarrow B$, but instead $X \rightarrow B + (X \rightarrow A) * (A \rightarrow B)$. Mediation analysis tries to get something for nothing.

Conditional Causal Inquiries Complier average is also a CATE RDD is also a CATE

Example of a causal inquiry:

```
declare_population +  
declare_potential_outcomes +  
declare_estimand(ATE = )
```

How should you select Inquiries?

Social scientists are picking research questions it can be hard to know where to start.

Theory is imaginary. We have to spend time *imagining* it. That's what theorists do (with different formalizations). They imagine models (M). Models are complex, there are many things to know about them, that is, there are many inquiries.

Data are real. We have to spend time *making* them. That's what empiricists do (with many different measurement procedures). They make data real by creating them. Data come in many forms, there are many things to know about creating data. There are many ways to summarize them.

Empiricists have to be concerned about research design. We have to develop empirical strategies to demonstrate I . We have to learn a lot about how to select the data of strategy and the answer strategy in a way that demonstrates an inquiry about a model that anybody cares about. So empiricists have to learn about models and inquiries (theory) as well as about data strategies and answer strategies (empirics).

The sine qua non of a good research question (I) is that there is a feasible design ($MIDA$) that could answer it. That means picking a good question I does not just involve theory.

Ofcourse, you should study M to understand which I s are worth knowing. But you should also study D and A in order to learn how to demonstrate I .

It can be a difficult mapping, but we think to get started it's helpful to – at a minimum – write down the M and I of any causal model that interests you to get started thinking about selecting strong D s and A s. The goal is to learn how to map $I(M)$ to $A(D)$. Then return to theory to find new important inquiries, then write down a new model inquiry data strategy and answer strategy. This process is how we make progress on M , that is, we bring M closer to W , thereby making M truer.

A good inquiry is informative about the model. - Existence of a node (descriptive mean) - Model suggests two variables should be correlated (descriptive correlation) - Arrow between a node (experiment) - CDF / PDF of a variable

A bad inquiry: - does not exist - does not address a “central” part of the model - does not summarize the theory

You are responsible for your inquiry.

Choice of inquiry is not valueless.

Grab bag of ideas we don't know where to put

- Study dependent estimands like Expected ATT vs realized ATT. Yes, your I depends on D , but in a weird way. Since you can write down all possible ATTs in the model, you actually *can* write the estimand without realizing D , you just don't know which of the multiverse of estimands you've written down will actually eventuate. Since beliefs about the data strategy (it will create a class of subjects who will be treated) can sort of be hacked into the model, I wonder about how fundamental this wrinkle is.
- Complex counterfactuals
- Parametric estimands (Dimensionality reduction, Model parameters)
- Vector-valued inquiries (set of predictions for N units or a function like the CEF),

- QCA estimand.
- estimands like “TRUE”
- Models as estimands (is this different from the CEF?)
- Estimand scope: What is the set of units which you want to learn the answer about? Know what ATE averages over
- Unknown estimands: Inquiries for discovery
- Auxilliary Inquiries: these inquiries are not the main substantive focus of the design, but they are features of the model that have observable implications and can be checked. For models that include mediating roles for a variable, need to demonstrate that Z affects M at a minimum. Balance tables demonstrate a required assumption that Z is orthogonal to pre-treatment characteristics. A full DAG specification can enumerate the full set of conditional independencies that are implied by the model; these can be checked and verified.

ADVICE SECTION

[note to all, this is a nuts stream-of-consciousness. I want there to be some advice because I don’t want to punt, but I know this advice will have to change.]

Since there are so many things we *don’t know* about the social world, researchers are positively overbrimming with interesting theories to test. Then again, most of us are a little bit bewildered when it comes to actually formulating a research question that would simultaneously be worth knowing the answer to and that would be feasible to answer. This section considers and critiques some common research question formulation advice. Perhaps unsurprising, our overarching perspective is that we can’t compromise on research design – scientific conclusions depend essentially on strong research designs. That said, choosing research questions is much more art than science. Moreover, the choice of what to study is not valueless in the sense that scholars’ research agendas are reflections of their values and their politics. While scientists seek to reveal objective truths, we also have to choose **which** truths to investigate, and that choice is value-laden.

The false choice between theory-first and method-first approaches.

Some research questions are boring and trivial – no doubt you have encountered such research questions when reading bad research. Figuring out the best way to answer unimportant questions can be a stimulating intellectual puzzle, but the answers to those questions themselves will not advance scientific knowledge. Some research questions are exciting and vital – those are the ones we want to answer using credible research designs. Giving weak or noncredible answers to important research questions is *dangerous* and unethical.

Sometimes scholarly debate about research design treat these two desiderata (good questions and strong designs) as forming a one-dimensional spectrum. To the left we have boring studies done well and to the right we have interesting studies with weak empirics. This “spectrum” view manufactures a tradeoff between the quality of the question and the strength of the research design. It tricks students into thinking that they have to make compromises on the research design (i.e., control for enough things and hope for the best) in order to do great things.

The reality is that research is hard and sometimes it is so hard that we must abandon them. Good questions might be too damn hard to answer and we must deliberately walk away from them even though they are interesting. Bad questions aren’t worth the effort, even if they are feasible – walk away from them too. Both interesting questions and strong research designs to answer them are required.

Advice about how to solve this dilemma often takes one of two forms: theory-first approaches and method-first approaches. Under the theory-driven approach, you read a lot of others’ work in order to formulate a good research question; you then search among the available methods to find the “right” method for the question. This approach is risky because it is very hard to abandon a good research questions, even if none of the available methods is good enough. The worst case scenario for the theory-first is a research process that takes forever, scholars falling prey to the sunk cost fallacy and falling to abandon a hopeless project, and a final

research project that relies on heroic assumptions and makes no real progress on the interesting research question.

The method-first approach involves learning a method well first, then looking for interesting projects to apply it to later. This frequently means *practicing* the method for the method's sake. To become an expert interviewer, you must conduct many interviews. To become an expert experimentalist, you must randomize many treatments. To become an expert observational data analyst, you must download many datasets and develop a sense for when inferences are reliable and when they are not. The worst-case scenario for the method-first approach is that scholars never stop *practicing* research to actually finally *conduct* research on important topics. What good is it to be a top-flight archivist if those skills are never applied to good questions.

Obviously, a mix of the two approaches is required, and most of variation in advice given to young scholars isn't just one or the other, but the relative weight given to the two. As a team of authors, we disagree amongst ourselves about the right weights. That said, the downsides to boring but well-done research don't seem as bad to us as the downsides of interesting but poorly-done research. Bad studies on important questions are used by disingenuous and bad faith actors to push for policies that are bad for society. The ethical stakes are higher for important questions than they are for unimportant questions. For that reason alone, we think it is vitally important that researchers master their methods so as to assiduously avoid making unknowably incorrect claims.

People who are looking in the forest for mushrooms often don't see a mushroom for a long period of time. After a while, they acclimate. They get their "eyes on," and successful finds seem to be around every bend. In this analogy, the theory building process is going for a walk in the forest and methods training is learning to spot mushrooms – you need get your eyes on answerable research questions worth answering.

In our experience, that means that the most important thing for young scholars to do is practice. Practice both searching for interesting questions and practice research design. Practicing research design often involves honing one's skills on less than interesting questions. Looking for interesting questions entails reading other scholars' work carefully and generously. Even when their research designs are not strong, their questions might nevertheless be interesting. You may learn from reading bad work on interesting questions that some research questions are impossible to answer. That's OK – now you've learned about the constraints on what we can learn. Operating within the constraints of what is knowable means you can anticipate when a poor research design on an interesting question is unlikely to lead to progress.

Novelty

Another route for finding research questions is to try to determine what is unknown to the research community, then to seek to fill that gap. This approach can be quite generative, but it is not without pitfalls. Establishing primacy – that a finding is "new" and that the knowledge generated by the study was previously unknown – is hard. It's also risky, since peer reviewers have access to Google Scholar. False primacy claims are especially damaging during the review process because they are easily refuted and demonstrate ignorance of the published record.

What does it mean for a research finding to be novel? There are two main cases. First, a finding is new if the empirical pattern was predicted by an earlier theory but had yet to be confirmed. That is, there was a theoretical prediction that had not yet found its empirical verification. The empirical finding is novel in that it confirms the previously unconfirmed theoretical prediction. Second, an empirical finding is also novel if the empirical reality was *not* predicted by the theory. If the empirical finding is trustworthy, it means that the original theory is incomplete or incorrect in a particular way. Findings that contradict theoretical predictions are also novel. Empirical findings are not novel when the theoretical predictions that pertain to the finding have already been well-documented.

This puts the advance of science in a funny position. Because we are worried that many of the advances made at the forefront of the scientific endeavor are incorrect or spurious, we really value replication of novel results. When an empirical claim in favor of a theoretical proposition is made in Place A at Time 1, we immediately want to know if it holds at Place B at Time 2 as well.

Mixing methods

A common piece of advice for young scholars who are interested in a topic is to mix methods. The thinking goes like this: I have an important question but none of my methods perfectly answers the question. I'll take a multifaceted approach that gets that many different aspects of the question which I hope to aggregate into a whole.

We caution against this approach. Mastering one method is difficult; mastering three is very difficult. Even when mastery of the methods is not an obstacle, the problem that often occurs is that the three research designs actually address *different* research questions. That is, not only do the three research designs use different data strategies and answer strategies, they also employ different models and inquiries. As a result, mixed methods work often gives three incomplete answers to three related, but ultimately different, research questions.

Things to incorporate

- “model” inquiries in which I’m trying to rule out some Ms. Process tracing is