Causal Inference, Mill's Method and Process-Tracing

Mill’s Methods

- John Stuart Mill: "System of Logic" (1843).
- “method of agreement” (“positive”) and the “method of difference” (“negative”).
- The *agreement*/positive method attempts to identify a similarity in the independent variable associated with a common outcome in two or more cases.
- The *difference*/negative method attempts to identify independent variables associated with different outcomes.
- "method of concomitant variations": Instead of observing merely the presence or absence of key variables, concomitant variation measures the **quantitative variations** of the variables and relates them to each other, a method that is in some sense a precursor to statistical methods.
- "Since the logics associated with Mill’s methods are integral to the strategy of controlled comparison, one must scrutinize studies that employ this strategy. One must judge how well the investigator has managed to achieve “control” among the cases"

First, the causal relation being investigated must be a **deterministic regularity** involving **only one condition** that is either *necessary* or *sufficient* for a specified outcome.

Second, *all causally relevant variables must be identified prior to the analysis* (whereas Mill’s methods are applicable only for explaining single-cause hypotheses).

Third, *cases that represent the full range of all logically and socially possible causal paths must be available for study*.

In the *method of agreement*, the investigator employs the *logic of elimination to exclude as a candidate cause (independent variable) for the common outcome (dependent variable) in two or more cases those conditions that are not present in both cases*. A cause or condition that survives this method of elimination can be regarded as possibly associated (“connected,” in Mill’s terminology) with the case outcome.

Omitted variable bias: "An inherent *weakness* of this method of causal inference is that *another case may be discovered later in which the same outcome is not associated with the variable that survived the elimination procedure in the comparison of the two earlier cases*."

Thus, the possibility remains that the common condition identified for the similar outcome in two cases may turn out to be a “false positive.”

In the *method of difference*—in which two cases having different outcomes are compared—the investigator *employs the logic of elimination to exclude as a candidate cause* (independent variable) for the variance in the outcome (dependent variable) *any condition that is present in both cases*. On the face of it, *the logic is quite simple: a condition present in both cases cannot account for the difference in case outcomes*. However, conditions that were not present in both cases can

only be regarded as possibly causally associated with the variance in case outcomes, for these conditions may not be present in other cases with the same outcome

Hence, inferences in both methods of agreement and difference may be spurious and invalid.

On the other hand, if a much larger number of independent variables are included, we may well encounter the problem of underdetermination (also known as “too many variables, too few cases”).

This logic of causal inference for small-n comparisons is highly problematic if the phenomenon being investigated has complex, multiple determinants rather than—as in the simple examples of Mill’s methods discussed above—a single independent variable of presumed causal significance

Another major difficulty in employing the logic of elimination occurs when different instances of the phenomenon under investigation have alternative determinants—what Mill referred to as the problem of “plurality of causes.” This condition is termed “equifinality” in general systems theory and is also sometimes called “multiple causality.”

With the method of agreement we cannot be certain that the outcome is associated only with a given independent variable. *If that phenomenon is subject to plurality of causes, we may sooner or later encounter one or more additional cases in which the outcome occurs in the absence of the conditions with which it was earlier associated*

Case Studies and Theory Development in the Social Sciences

Types of Process Tracing

1. DETAILED NARRATIVE

The simplest variety of *process-tracing takes the form of a detailed narrative or story presented in the form of a chronicle that purports to throw light on how an event came about*. Such a narrative is highly specific and *makes no explicit use of theory* or theory-related variables. It may be supportable to some extent by explanatory hypotheses, but these remain tacit. Historical chronicles are a familiar example of what is at best an implicit, atheoretical type of process-tracing.

1. USE OF HYPOTHESES AND GENERALIZATIONS

In a *more analytical form of process-tracing*, at least parts of the narrative are accompanied with *explicit causal hypotheses* highly specific to the case without, however, employing theoretical variables for this purpose or attempting to extrapolate the case’s explanation into a generalization.

The Limits of Process-Tracing

1. There are two key constraints on process-tracing. Process-tracing provides a *strong basis for causal inference only if it can establish an uninterrupted causal path linking the putative causes to the observed effects*, at the appropriate level(s) of analysis as specified by the theory being tested.

2. *Another potential problem for process-tracing is that there may be more than one hypothesized causal mechanism consistent with any given set of process-tracing evidence*. The researcher then faces the difficult *challenge of assessing whether alternative explanations are complementary in the case*, or whether one is causal and the other spurious.

Others

In the same way, process-tracing can ameliorate the limitations of John Stuart Mill's methods of agreement and difference. For example, process-tracing offers a way of assessing hypotheses regarding causal relations suggested by preliminary use of Mill's methods, as in Theda Skocpol's study. More generally, process-tracing can identify single or different paths to an outcome, point out variables that were otherwise left out in the initial comparison of cases, check for spuriousness, and permit causal inference on the basis of a few cases or even a single case.