Designs: Small N Studies

Definition: some, but only few cases with a limited number of units of observation (= small N).

- Central problem: How to select the cases?

- Random selection, as propagated by King, Keohane and Verba in order to achieve descriptive and causal inference, is impossible.

- Therefore a conscious selection of cases is necessary that avoids selection bias.

- Design depends on strategy (x-centered or y-centered)

x-centered strategy

Three principles:

(1) selection on the independent variable to achieve a robust ceteris-paribus assumption,
(2) selection under basic information of the universe of cases and
(3) selection of different cases than those from which the hypotheses were derived.

y – centered strategy

In a y-centered strategy, the cases select themselves. Here analytical narratives are relevant in disciplining the specifics through general theory.

In the current methodological debate a combination of both is favoured:

Steffen Ganghof: First develop an explanation in a y-centered design for one or a few cases, then test this explanation in an x-centered strategy for a large N.

Bernhard Kittel: for macro-quantitative research questions of the type „Do institutions matter?“, start with an x-centered design for a large N to test for potential explanations. Second, these explanations should be tested for a small N with a short causal chain.

(„A Crazy Methodology? On the Limits of Macroquantitative Social Science Research“, 2005)

Criteria for case selection

Principle:

Maximise experimental variance, minimise error variance and control for external variance. (Peters 1998: 30)

► Experimental variance: Choose cases such that x and y show large variance.

► Error variance: Carefully select measurement concepts and make use of a code book.

► External variance: Take care of those factors that are assumed to be constant, to be ceteris paribus.

All designs that seek to follow this principle are variations of John Stuart Mill's „System of Logic" (1843).
=> Eliminate factors which are irrelevant for the dependent variable.

► Method of agreement (= Konkordanzmethode)
=> dependent variable constant, elimination of those factors that cause variance on the dependent variable.

► method of difference (= Differenzmethode)
=> variation on the dependent variable, elimination of those factors that have no effect on the dependent variable.

Mill's other procedures are only relevant in experimental designs (e.g. method of residuals, method of concomitant variation).

Logical preconditions for Mill's designs: deterministic causality and relatively simple causality, i.e. no multiple causalities.

With a given case selection external variation cannot be controlled. Nevertheless such designs are very often used:

- Most similar with different outcomes = MS-DO strategy

- Most different with similar outcomes = MD-SO strategy

Therefore, if one follows such strategies it is extremely important to choose the most similar or most different cases on the basis of a large number of variables in the universe of cases, that is, on the basis of quantitative cluster analysis.