Case Studies and Generalization

PAPER SUBMITTED TO DOCTORAL COURSE DRE 1002

CASE STUDIES METHODOLOGY

Lars C. Kolberg
Department of Public Governance
Norwegian School of Management BI
P.b. 580
N-1302 Sandvika
Norway

lars.c.kolberg@bi.no

Case study approaches, where few units are investigated in depth, were dominant within the fields of sociology, political science and anthropology in the first part of the 20th century (Andersen 1997). However, following the development in increasingly more sophisticated statistical methods and sampling techniques the case study approaches came to be regarded more as a sub discipline at best able to generate different kinds of insight from that of quantitative oriented studies. General knowledge is often assumed to require large-N studies with statistical representativeness.

Eckstein (1975:113ff) elaborates this points neatly: If we think of theories as a curve or line to which empirical observations are expected to correspond, a case study would yield only one observed point, and as we know, an infinite number of lines could be drawn through one single point. QED: a case study cannot discriminate between potential general theories. One problem with this conception is that it assumes induction from empirical observations to theories – i.e. that “theories … can be fully derived from observations by simple inspection or, at any rate, sophisticated data processing” (Eckstein 1975:113). However, thinking of theories in for instance physics, they do not seem to be constructed inductively from observations. The discovery of new ideas is not subject to a particular logic, and their
subsequent testing might occur deductively (Popper 1968). In a deductive reasoning a well-chosen case might yield results that can be compared to the predictions from a theory just as well as a well-constructed experiment. The concept of using control groups in medical research is but one example of this. According to Habermas, finding theories that are contradicted by one's interesting data might be just as interesting as the opposite (Andersen 1997).

The main focus of this paper will be the discussion of various challenges often associated with case studies and generality (section 3) and different strategies for dealing with these (section 4). At first, however, the next two sections (1 and 2) briefly outline the concept of general knowledge and the very concept of a case study.

1 Conceptions of generality

It is argued that the hallmark of scientific research is the goal of making inferences that go beyond particular observations (King, Verba & Keohane 1994:8). This is in essence about generality – making knowledge that is transferable, with external validity.

A study of voting behaviour using mass surveys typically explains the voting decisions of voters in general and says nothing about the vote of a particular voter. A study of the French revolution on the other hand, tends to explain that very event. The goals of generality and specificity are not, however necessarily mutually exclusive: “Indeed, the best way to understand a particular event may be using the methods of scientific inference also to study systematic patterns in similar parallel events” (King, Verba & Keohane 1994:43).

In statistical research we are used to consider whether the findings can be generalised from the sample to the universe. One thing statistics supplies us with is a handy way of reporting findings. The very term ‘statistic’ “is an expression of data in abbreviated form. Its purpose is to display the appropriate characteristics of the data in a convenient format” (King, Verba & Keohane 1994:53). The Sample mean is one example of this. Further, we can calculate what the chances are that the observations in our sample are random i.e. they do not come from a systematic variation in the universe. If the chances are to small we say that the result is significant, meaning that we believe the findings to be general.

Early case studies in anthropology and sociology also aimed at generality. Typically a link between the micro level and the macro level was assumed, i.e.: findings in a micro cosmos somehow were assumed to represent the macro cosmos. In the study of delimited societies where it is possible to collect “all” data, this may hold to some extent. However, it appears to be more problematic in a modern, complex society (Gomm, Hammersley & Foster
Glaser & Strauss’ concept of ‘grounded theory’, combining inductive and deductive reasoning, represented one departure from this assumption (Andersen 1997).

A “weak” form of generalisation often associated with case studies is ‘naturalistic generalization’ (Gomm, Hammersley & Foster 2000). The term implies that case studies offer working hypotheses that may be appropriate for other cases if the ‘fit’ between the cases is close enough. It is apparently up to anyone to judge whether the fit is close or not. Therefore it introduces an element of everyday common-sense reasoning (Gomm, Hammersley & Foster 2000). This might work. After all we often rely on this kind of generalisation, but our everyday generalisations are also shown to be erroneous and indeed biased (Tversky & Kahneman 1974). Using naturalistic generalisation as the form of generality possible to achieve from case studies therefore implies sort of a capitulation: real general conclusions may not be sought and case studies are still an inferior science to statistical research.

**Middle range theories**

One particular question when dealing with the concept of generality is how general general is. Robert K. Merton addresses this in his seminal article on Middle Range Theories (Merton 1967). To him theories are “logically interconnected sets of propositions from which empirical uniformities can be derived”. ‘Theory’ covers a broad spectrum where ‘Middle range theories’ are theories that “lie between the minor but necessary working hypotheses … and the all-inclusive systematic efforts to develop a unified theory that will explain all the observed uniformities of social behaviour, social organisation and social change” (Merton 1967:39). Such theories are closer to the observations than general system theory, making it more convenient to incorporate them in propositions that permit empirical testing. Still, they are generalised – they involve abstraction. Middle range theories deal with delimited aspects of social phenomena. They hold under certain conditions and they can yield predictions that need not be true for the universe, but which holds for certain groups e.g. for certain nations (Lewis in Merton 1967:57). Further, middle range theories might be a necessary condition for more general theory development. Weber’s inquiry into “Das Protestatische Etik und Geist des Kapitalismus” for instance, deals with a severely delimited problem. It covers a certain phenomenon in a particular historical epoch. Still without doubt, it has contributed to more general theories of modes of interdependence in society. The same is true for the important heritage of *concepts* from the early sociologist. Tönnies’ typology of different societal forms as *Gemeinschaft* and *Gesellschaft*, Durkheim’s concept of *Anomi* and Weber’s concept of Bureaucracy to mention but a few have had a long and lasting impact on the social sciences.
Concepts then, represent one way of generating general knowledge. If they cover the essence of the data variation, concepts can be used as “lenses” in further investigation of new data. From concepts, the road to developing hypotheses need not be a long one. Linking concepts in a conceptual model, it is possible to hypothesise about the relationships between the variables. The use of middle range theories, then, facilitates research that is both empirically based and theoretically relevant (Merton 1967:58).

Generality might mean different things and may be sought on different levels. What is important is that it is not something reserved to statistical representative survey studies. Aiming for “middle range generality” from case studies does not imply surrender into loosely edited narratives suitable for discretion-based naturalistic generalisation. It can even be developed into even more general theory. The argument that will follow, however, is that generality from case studies requires careful design and selection of cases in addition to a prior knowledge of theory. Previous to that, the next two paragraphs will briefly present the common usages of the term case study and some challenges usually associated with generality from case studies.

2 What is a case study?

It is possible to identify a number of approaches to what the concept of a case study is. One typical approach takes the (small) number of cases as the defining characteristic. A study of few cases allows for the investigation of many variables, therefore producing in-depth knowledge. Case study research as a research strategy is often confused with types of data collection methods (e.g. fieldwork) and with types of evidence (e.g. qualitative data) (Yin 1981). Case study methods are consequently, often viewed as a support discipline at its best able to deliver ideas that can be tested in quantitative studies.

Case studies are for the same reason often seen as analogous to clinical studies in medicine, whereas large N studies resemble the experiment (Eckstein 1975). Clinical studies attempts to capture the whole individual, while experiments are used for discovering relationships. This is the same as the division often made between intensive (small N, many variables) and extensive (large N, fewer variables) studies (Hellevik 1991). There is however, no necessary reason why case studies should be intensive clinical-like studies aiming at holistic explanations. Case studies might also be used for investigating relationships between variables. This is not a particularly new insight (for instance Eckstein 1975). Still the “ideal types” of qualitative, intensive case-approaches telling unstructured narratives vs. quantitative, extensive studies on representative samples reporting coefficients and
confidence intervals are very enduring. Even the very concept of a case differs between different streams of research (Ragin 1992). In line with realism, cases can be argued to be observable units that are collected or found. A conception of ‘a case’ in line with nominalism on the other hand, would be to argue that the researcher constructs cases. Further, cases can be viewed as analogue to the observational units or as examples from a larger category – instances of something. Even where cases are conceptualised as ‘instances’ researchers differ as to whether the cases are to be understood as instances of an actual population (business companies) or from a conceptual class (bureaucracy) (Abbott 1992). These different approaches give way for different ideas about the size of N, and the logic of case study research. How the cases are approached, are therefore more important than how many they are (Yin 1989). Well-conducted case studies can be seen as (quasi) experiments. This requires an understanding of what the challenges are and how they can be dealt with, matters that will be discussed immediately below.

3 What are the challenges with case studies and generality?

Case studies are often seen as able to answer ‘how’ rather than ‘why’ questions. Researchers within the tradition of “etnomethodology” take this to the extreme, arguing that the purpose of case studies is not general knowledge but subjective actors’ experience. When case studies are associated with ideographic, non-general knowledge, this is often due to a mix of the techniques often associated with conducting case studies and the case study as a research strategy.

Still, on a more general level, it can be argued that generality and complexity are at odds with one another. Ragin (1987) argue that case study approaches value understanding of complexity over generality, while the opposite are true for the variable oriented approach concerned with testing hypotheses derived from general theory: “In the case-oriented approaches … it is clear that the goal of appreciating complexity is given precedence over the goal of achieving generality … The case-oriented approach uses theory to aid historical interpretation and to guide the identification of important causal factors; the variable-oriented strategy, by contrast usually tests hypotheses derived from theory” (Ragin 1987:54). This is a commonly held conception and it supports the notion that case studies are no good for generalisation purposes.

Smelser (1973) recommends the use of case oriented methods when N is too little to use statistical methods. Case strategies are therefore regarded as inferior to statistics. To Ragin (1987) the case strategy is not able to cope with a number of cases sufficient to yield
general results, but it can give valuable insights still: it is “useful for resolving paradox” (Ragin 1987:45). He suggests a synthesis between variable approach and case approach allowing for more cases to be analysed as the remedy: “A proper synthetic strategy should provide an avenue from this criticism [of non-generalisation]. At a minimum, it should allow investigators who wish to protect themselves from this charge to examine more than a few cases” (Ragin 1987:82).

In the following two standard objections will be discussed. The first has to do with the small-N characteristic and non-representativeness of case study data. The second objection has to do with the concepts of causation in case study research.

**Case studies as different from statistical studies**

The first objection, the lack of statistical representativeness in case studies, is relatively easy to dismiss. *Case study cases are not meant to be selected randomly.* Rather it could be viewed as a form of experiment where cases are selected with a view to maximize the variation of the variables under investigation. This is in line with Gomm, Hammersley and Foster (2000) who argue that general conclusions from case studies may be arrived at in two ways: via *theoretical inference* and *empirical generalisation*. Theoretical inference resembles the experiment with cases being selected with an analytical view.

The main problem with *Empirical generalisation* from studied to unstudied cases is that it is potentially subject to high and unknown levels of error. This problem obviously increases with heterogeneity in the population. If it could be reasonably argued that the population to which generalisations was to be made consisted of more or less similar cases, generalisation would be fairly simple. If it is (as it usually is) however, heterogeneity in the target population, random sampling can deal with it effectively assuming random distribution of heterogeneity across population. *Case study approaches do not produce statistically representative data in this manner.* The use of theoretical ideas and careful selection of cases therefore represent another way of dealing with the heterogeneity. The control question would be: “How are the relevant aspects of the case typical / atypical for the population?” And “of what population could it be typical” (Gomm, Hammersley & Foster 2000:105).

In the statistical language, it is often argued that case study research faces a ‘degrees of freedom’ (df) problem (Campbell 1975, King, Verba & Keohane 1994): The estimation of free parameters in a model requires more observations than parameters: one cannot estimate two parameters with one observation. The responses to this are various. One stance is that df is not really a problem because case studies is qualitative different from statistical studies.
Perhaps more constructively it can be argued that df are only a problem in case study research under an inductive logic, as a way of discovering theories (Ebbinghaus 2003). If adding up estimated parameters makes a theory, df might be a problem in case studies. Viewing instead case studies as a means to test hypotheses deductively (Savolainen 1994) the df problem arising from small-N is less acute. Another stance towards df is to acknowledge df as a problem. There is obviously two different strategies to deal with it: To increase N or to reduce the number of estimated parameters (variables).

<table>
<thead>
<tr>
<th>Do case studies face a df problem?</th>
<th>Why?</th>
<th>How to deal with it?</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>Not relevant</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Deductive logic</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>Small N</td>
<td>Increase N: more cases</td>
</tr>
<tr>
<td></td>
<td>Many variables</td>
<td>Reduce variables: comparable cases, analytical control, theory</td>
</tr>
</tbody>
</table>

**Table 1: df and case studies - various stances and responses**

**Case studies and causation**

A second standard objection to case study research is that even well conducted case studies may have problems dealing with causation. Two different aspects of this will be scrutinised in the following: 1) causation as deterministic vs. causation as probabilistic 2) Multiple and conjunctural causation.

There are basically two different types of causation and different methods for dealing with them (Mahoney 2000). *Nominal* methods work within in a deterministic logic, i.e. that the association between variables is either 0 or 1. *Ordinal* methods consider also the degree of the association, and therefore probabilistic associations can be estimated.\(^1\) The standard claim is that case study methods can only handle deterministic causes (Lieberson 1992). However, ‘determinism’ is often understood as “what happened had to happen”. Such an understanding leaves a very small space for case studies. However, determinism can also refer to situations where *causes are necessary and/or sufficient* (Mahoney 2000). A subgroup of this is when causes are *both* necessary and sufficient. This special situation gives an invariant association: When a sufficient and necessary \(x\) is present, \(y\) will always be present.

---

\(^1\) This is however, only briefly treated in the course literature. Although for instance Mahoney (2000) writes insightful about Mill’s method of concomitant variation, the focus in this paper will be on nominal methods following a deterministic logic.
When a sufficient and necessary x is absent y will never be present. Often however, causes will be either necessary or sufficient, which gives a variation of outcomes as illustrated in table 1 below. The presence of x will always result in y if x is sufficient, but might or might not result in y if x is necessary. Similarly but opposite: The absence of x might or might not result in y if x is sufficient, but it will never result in y if x is necessary.

<table>
<thead>
<tr>
<th>Sufficient x</th>
<th>Necessary x</th>
<th>Necessary and sufficient x</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present</td>
<td>Y is always present</td>
<td>Y is present or absent</td>
</tr>
<tr>
<td>Absent</td>
<td>Y is present or absent</td>
<td>Y is always absent</td>
</tr>
</tbody>
</table>

Table 2: Outcome (y) for presence/absence of necessary and sufficient conditions (x)

The two canonical methods elaborated by J.S. Mill in the 19th century are often seen as the basis for systematic comparison in social science. ‘The method of agreement’ and ‘the method of difference’ deal with exactly necessary and sufficient respectively. The method of agreement seeks to eliminate potentially necessary causes. With a common outcome across cases, a hypothetical cause that is not shared by the cases cannot logically be said to be ‘necessary’ to produce the outcome (Mahoney 2000). In the figure below x2 is the cause of y according to the method of agreement.

Figure 1: Method of agreement – an illustration

The method of difference on the other hand, sets out to eliminate potentially sufficient causes. The possible independent variables are examined in order to find a pattern where only one independent variable varies systematically along with the dependent variable. In a situation with differences in outcome across cases, a hypothetical cause common across cases cannot
logically be said to be sufficient (Mahoney 2000). Only x4 in the figure below cannot be eliminated as a potential sufficient cause according to the method of difference.

The methods aim at eliminating false causes of a phenomenon. This will only work if all relevant aspects of the case are included at the outset (Hammersley, Gomm & Foster 2000): This “omitted variable bias” (King, Verba & Keohane 1994:168) arises first because of the fact that all the features of cases are not always straightforward observable. Second, necessary and sufficient causes tend to be trivial and tautological, respectively (Mahoney 2000). There will always exist a lot of causes that can be said to be necessary in a broad sense for an outcome. For instance, the act of driving a car is strictly speaking a necessary cause for causing a car accident. Likewise, willingly hitting another car will alone be a sufficient cause for a car accident. Since the elimination or inclusion of variables tend to alter the result of the analysis, the researcher faces a “specification problem” (Mahoney 2000).

Another frequent cited problem with causation in case study research arise from multiple and conjunctural causation. When causation is multiple, different variables or different configurations of variables cause the same outcome. Conjunctural causation is when “a phenomenon or a change emerges from the intersection of appropriate preconditions.” (Ragin 1987:25). It is when a specific configuration of x’s causes y: without one of them, then no y.
When causation is multiple and conjunctural – the effect of one variable is changed by the context it occurs in. This implies that a relationship could be causally significant in one setting but not in another (illusory commonality) or opposite that different features have the same effect when associated with different other features (illusory difference). The methods from Mill, assume that no such interaction effect is present. One frequently used example of these deficits is the attempt to explain car accidents. In a situation where two drivers experience accidents; one of them have been drinking, but both were running a red light, the method of agreement would rule out alcohol consumption as having anything to do with the accidents. Likewise, in another situation where one driver collides and one does not and where both have been drinking, the indirect method of difference would render alcohol consumption as irrelevant. Lieberson (1992, 1994) shows in greatest detail how these methods can produce strange and counter-intuitive conclusions. The interaction problem cannot be resolved by combining variables according to Lieberson. The reason is that there is no way to assess which interaction variables should be included (1992:213, note 6). The only strategy for resolving the single cause logic of Mill’s methods is a true experimental situation, where one group is subject to a stimuli and another group is not, while everything else is held constant, assuring that the *cetera* are indeed *paria* (Lijphart 1971:683). Mill himself was aware of this and rendered the application of the method of difference to political science “completely out of the question” (J.S. Mill, in Lijphart 1971:688). When looking for single causes in multivariate data using statistical techniques, the control for variables is very “degrees of freedom consuming”, thus requiring large samples. In case study research one possible strategy is to relax the assumptions slightly, instead aiming at maximizing “the ratio between the amount of variance of the operative variables and the amount of variance of the control variables” (Lijphart 1975:163). Lieberson himself offers another possible strategy: in
an experimental situation it makes sense to look for single causes (1994:1234). Attempting to design case studies as (quasi) experiments is therefore an escape path. Good design is one necessary but not sufficient condition, as the discussion of Mill’s methods should have demonstrated. The use of theory is another key issue. An a-theoretical comparative case study is probably not the most appropriate approach when studying the causes of car accidents. Basing it on theory would allow including for instance what we know about the effects of alcohol on reaction time etc.

4 How are the challenges dealt with?

The principal challenge in all kind of studies is to control for unwanted variation. All forms of research strategies therefore can be said to work under the same “logic of inference” (King, Verba & Keohane 1994). The central idea is to maximize the leverage over research problem. One way of approaching this is to find as many observable implications of the theory as possible – to explain as much as possible with as little as possible (King, Verba & Keohane 1994:29).

Theory is pivotal to the idea of achieving general insight from case studies. Even single cases can be seen as belonging to a larger class of events in a theoretical universe. This requires a notion of cases as “cases of something” (Ragin & Becker 1992) and not merely as research sites (Eisenhardt 1989). The important objective is to be conscious about how cases fit in an empirical or theoretical universe. The following paragraphs will give a relative brief summary of different types of case studies. More attention will be given to the matters of selecting cases and different analytical designs.

Types of case studies

Based on Lijphart (1971), and Andersen (1990) different types of case studies can be delineated. The most important distinction made is that between studies that are a-theoretical, motivated by an interest in the phenomenon studied in itself, and studies that aim more explicitly at general knowledge. The ‘etnomethodological’ tradition values a-theoretical ‘intrinsic’ case studies that can give insight into actors’ subjective experience. The studies are undertaken for descriptive rather than generalisation purposes. The population under study is given and empirical. Lijphart (1971:691) argue that strictly speaking a-theoretical studies are impossible because any analysis will be guided by at least some vague notions of theory and some knowledge on other cases.
Theoretical studies covers ‘interpretative studies’ and studies aimed at generating or testing hypotheses and theories. Interpretative studies like a-theoretical studies have a largely descriptive purpose but existing theory is explicitly applied to the cases. Studies attempting to generate or test and develop, both obviously have an analytical purpose, but whereas the former typically studies an empirical given population, the latter is conducted in a theoretical universe\(^2\). It is methodologically important to separate these two kinds of studies. When a case study generates hypotheses, the hypotheses should obviously not be tested on the same material. The hypotheses ought to predict “new facts” (Lakatos 1970), that is facts formerly unobserved, which apparently requires additional data.

<table>
<thead>
<tr>
<th>Population / Universe</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Descriptive</td>
</tr>
<tr>
<td></td>
<td>Empirical</td>
</tr>
<tr>
<td></td>
<td>Theoretical</td>
</tr>
</tbody>
</table>

Table 3: Types of case studies

The other key dividing line is the one between \(N=1\) studies and \(N>1\) studies. It is argued that well chosen single cases can overthrow theory with only \textit{one observation} (Eckstein 1975:127). A well-chosen case would be a case that is somehow crucial for the theory (Eckstein 1975). Crucial cases can test theory in a “most likely manner” (a case where the theory should apply but it turns out that it doesn’t) or a “least likely manner” (the theory is not meant to apply here, but still it does) (Eckstein 1975). But still “we always do better (or, in the extreme, no worse) with more observations as the basis of our generalization” (King, Verba & Keohane 1994:212). Particularly testing of the kind Eckstein suggest is unlikely with strictly one observation. There are several reasons for this. One is that it is very rare that one single variable represents an explanation. Further, there is always the possibility of measurement error causing the rejection of a true hypothesis. Finally, there is always possible that an unknown variable produces the outcome (King, Verba & Keohane 1994).

This division between single case and comparative case studies is potentially confusing, because it as indicated above is not necessarily the number of \textit{cases} but rather the number of \textit{observations} that makes a difference. An observation is defined as “a measure of

\(^2\) Lijphart (1971) also include “deviant case analyses” as a separate category. It seems reasonable however to view this as a special case of a theory generating study, namely one where the established theory is know to not hold and therefore new theory is sought generated.
one dependent variable on one unit (and for as many explanatory variable measures as are available on that same unit)” (King, Verba & Keohane 1994:217). A case study of one organisation can therefore yield a great number of observations. Still, comparative studies of multiple case (each with potentially many observations), can increase the possibilities of generality. It is important to emphasise that comparative case studies are not necessarily undertaken in order to achieve a greater statistical representativeness. More to the point is the possibility of getting more available observations and more design variables. The comparison of similarities and differences across cases can be a good approach for testing hypotheses, given careful selection and appropriate design.

Selection of cases

Cases can be seen as cases of something even though the actual empirical material comes from i.e. a unique phenomenon. Hence the term ‘case’ works grammatically in the same sense as ‘instance’ or ‘example’ (Gomm, Hammersley & Foster 2000:102 note5). Within this perspective, cases are identified in terms of broader categories. Even a ‘unique’ case, then, is regarded unique relative to what is common or representative of some categorical group. It could be a case of something from a theoretical as well as from an empirical universe. When case study research is regarded as a research strategy it is different from the actual gathering of data through e.g. fieldwork. Even case studies geared towards an understanding of the unique will benefit from an analytical design. In theoretical interpretative studies, the case is an example of a class of phenomena. When it comes to studies using theory for the purpose of generality, selection based on theory can make even single case studies into implicitly comparative studies – suggesting that they belong to and have validity in a theoretical universe.

A case is therefore not necessarily something that is readily observable. Rather it could be seen as a continuous process of constructing during the research. For instance, the empirical phenomenon of competitive tendering and contracting out in the Oslo care sector, can be studied as a case of changing the role of politicians (From & Kolberg 2003c), as a case of the role discretion in contract negotiations (From & Kolberg 2003b) or as a case of changing industrial relations (From, Heløe & Kolberg 2003) as well as overall being a case of moderate successful public sector reforms (From & Kolberg 2003a). Selection of cases therefore, is not necessarily to be understood in the statistical sense as drawing something from a hat. Rather it may involve constructing based theory and prior knowledge.

Design
The experimental method is often regarded as the ultimate scientific method. Even the most sophisticated statistical methods can only approximate an experiment. The ideal experiment offers control over both the selection of observation (instead of having them assigned) and the assignment of values (e.g. treatment / non treatment) of the explanatory variables (King, Verba & Keohane 1994:125). An experiment gives the possibility of definite linking causes to effects. Control with unwanted variation is the key feature. The concept of conducting experiments is often associated with laboratory work. This is not necessarily the case. For instance, the study of twins employs an implicitly comparative design, where the researchers are given the possibility to control for all of the variation in DNA between two individuals. Differences in outcome (e.g. attitudes, behaviour) must therefore be caused by environmental factors.

Case studies can be regarded as (quasi)experiments. Analytical control, through design of the analysis is achievable. To Lijphart (1971) comparative methods is an inferior method to the experimental method. Still, admitting that it can often be the only usable method, his discussion of strategies for resolving the “many variables, small N” problem is very instructive: 1) Increase the number of cases; 2) Reduce the “property space” (e.g. factor analysis); 3) Focus on comparable cases; 4) Focus on key variables. It is the third and the fourth suggestions that are of greatest interest here. ‘Comparable’ means: “similar in a large number of important characteristics (variables) which one want to treat as constants, but dissimilar as far as those variables are concerned which one wants to relate to each other” (Lijphart 1971:687). When cases are dissimilar, it is necessary to control the association under interest for the variation on other variables, i.e. the number of estimated parameters increases and the df problem arises unless further observations are included. Obviously, when cases are matched on central variables, the number of free parameters decrease because it is not necessary to control for different values on these variables: it is not necessary to estimate the effect of x on y both when z is high and low. Proper design therefore can help reducing the number of ‘operational variables’: either through not allowing them to vary (variables are turned into parameters) or through consider the variation to be irrelevant (based on theory or prior knowledge).

In comparative designs, the key feature is the comparison of similarities and differences. The positive comparative method is about finding similarities in independent variables associated with a common outcome (what could have been the features the countries of north-western Europe and North America had in common since they developed a
The negative comparative method attempts to find independent variables associated with divergent outcomes (Smelser 1973). The negative comparative method is often referred to as “most similar systems”. Here the logical foundation is Mill’s method of difference. A matching of cases on as many potential explanatory variables as possible reduces the number of operational variables that can cause an outcome different across cases. “Deviant case” studies are often more or less explicitly applying a most similar systems logic. The deviant case is a case with a dependent variable score different from the expectations based on the independent variables (similar with other cases in an empirical or theoretical universe).

The positive comparative method is often called “most different systems”. This method is based on the logic from Mill’s method of agreement. The central objective is a variation on as many potential explanatory variables as possible in order to isolate the invariant variables associated with a common outcome. This feature is strength of the most different system approach. The matching of variables associated with the most similar system approach can only be done for a finite number of known variables. But the analytical randomisation of the “most different systems” logic, in principle holds for all variables (Andersen 1990). Since differences in explanatory variables cannot explain an outcome common across cases the variables with a common value across cases are taken to be the real explanatory variables. The method is useful for additional testing of hypothesis. If the assumed association between variables still hold under conditions different from the original ones (i.e. in a most different system), our belief in the association is strengthened.

5 Conclusion – useful applications

Design and modelling helps dealing with the inherent complexity of social phenomena. A selection of cases with an aim to maximize the variation of the relationship under interest while minimizing other variation reduces the number of ‘operational variables’. Thus, Ragin (1987:55) is not necessarily correct when he states that: “Permanent causes [i.e. structural features] cannot be removed for purpose of experimentation because they are linked to constituent aspects of the unit”. While not removed, careful design can control for them (via selection or with the aid of theory) either turning them into parameters or rendering them non significant.

Within the research area of public sector reforms, the most common approach is to focus upon either the ‘whys’ of the reforms or the results of the reforms. Studies within the research stream of political economy typically attempt to explain the scope and pace of
Case studies and generalisation

reforms through analyses of politicians’ and the population’s attitudes towards reform (Sørensen & Bay 2002). The output side is also well covered through numerous studies measuring economical effects (see for instance Hodge 2000 for a meta analysis of these). Less effort seems to have been put into gaining an understanding of how the reform processes occur. How the various facets of change are managed is obviously of great importance for the chances of achieving success or fiasco in the implementation of the reforms, and also eventually for the results. Well-designed case studies can give insight into such processes. This is in line with the wide-spread perception of case studies. Importantly however, well-designed case studies of public sector change processes might also produce general knowledge on how to conduct successful public sector reforms.


**Literature**


Campbell, Donald (1975). "Degrees of Freedom" and the Case Study. *Comparative Political Studies, 8*(2), 178-193.


