The case survey method and applications in political science

Newig, Jens; Fritsch, Oliver

THE CASE SURVEY METHOD AND APPLICATIONS IN POLITICAL SCIENCE

Paper presented at the APSA 2009 meeting,

3-6 September 2009, Toronto

Panel on "Case study meta-analysis:
Methodological challenges and applications in the political sciences"

Jens Newig[1] / Oliver Fritsch[2]

Abstract

The ever-growing body of empirical insights in political science constitutes a major challenge for scholars in almost all areas of research. Meta-analytic techniques, the "case survey" method in particular, allow to systematically and rigorously synthesize previous case-based research by drawing on the richness of the case material, on different researchers and research designs, and at the same time allowing for a much wider generalization than from single cases. We review existing applications of case-based meta-analytic methods in political science and related fields and critically assesses strengths and limitations of the approach for political science. Placing case study meta-analysis in the broader context of methods of secondary analysis that synthesize previous research, we develop a typology of methods according to the respective source of data as well as the method of integration. We review previous applications of the case survey method in political research and related fields and discuss the relevance of these findings as well as methodological innovations and flaws. We portray the case survey method (case-based meta-analysis) in more detail, drawing on experiences from our own research. From these, we draw conclusions for the applicability of case meta-analysis, discussing strengths and weaknesses of the method specifically related to political research and conclude with an outlook for further methodological improvements in the field of meta-analytical and other evidence-based methods.

Key words

Research integration and synthesis, case survey method, systematic review, meta-synthesis, qualitative and quantitative analysis, environmental effectiveness, participatory governance.

---

[1] Prof. Dr. Jens Newig, Workgroup Sustainable Development and Participation, Institute for Environmental and Sustainability Communication, Leuphana University Lüneburg, Scharnhorststr. 1, 21335 Lüneburg, Germany, Phone: +49 (0) 4131 677-1726, E-mail: newig@uni.leuphana.de

[2] Oliver Fritsch, MA, Department of Policy Analysis, National Environmental Research Institute, Aarhus University, Kalø, Grenåvej 14, 8410 Rønde, Denmark, E-mail: olf@dmu.dk

1 Introduction

The ever-growing body of empirical insights in political science constitutes a major challenge for scholars in almost all areas of research. Both in established fields of inquiry and in allegedly more exotic or remote disciplines keeping track of all original (primary) studies is hardly possible. Therefore, researchers increasingly rely on integrative research reports, and efforts of reviewing certain strands of literature become a common feature in professional journals spotlighting the synthesis of primary research as a key issue of today's political science. While research methodology is highly developed for conducting statistical analyses, single case studies or comparison, surprisingly little effort has been devoted to synthesizing, aggregating and integrating primary research.

This is of particular relevance for areas of study that are, due to their inherent complexity, dominated by research designs focusing on case studies. The increasing attention paid to holistic or extremely context-dependent phenomena such as 'governance' are cases in point. However, single case studies, in particular, suffer from the familiar problems of non-generalizability. Small-N comparative case studies are doing better in this respect, but are still very much context-dependent. Large-N comparative case studies would be an excellent choice, but are usually extremely costly and therefore very seldom done. The problem is not so much that we are lacking knowledge but that this knowledge lies distributed over a myriad of single and small-*N* case studies, available in almost any given area of political science research. These present "an intellectual goldmine awaiting discovery" (Jensen and Rodgers 2001).

For more than three decades, methods to aggregate empirical case-study data have been available (Lucas 1974; Yin and Heald 1975). Yet surprisingly little use has been made of these methods, and little effort has been devoted to enhancing this methodology. These meta-analytic techniques, the "case survey" method in particular, allow to systematically and rigorously synthesize previous case-based research by drawing on the richness of the case material, on different researchers and research designs, and at the same time allowing for a much wider generalization than from single cases. However, a number of methodological challenges remain, not least due to the limited experiences within political science research.

This paper reviews existing applications of case-based meta-analytic methods in political science and related fields and critically assesses strengths and limitations of the approach for political science. To this end, the paper proceeds as follows: The following chapter (2) places case study meta-analysis in the broader context of methods of secondary analysis that synthesize previous research. We develop a typology of methods according to the respective source of data as well as the method of integration. Subsequently, we review previous applications of the case survey method in political research and related fields and discuss the relevance of these findings as well as methodological innovations and flaws (3). We then portray the case survey method (case-based meta-analysis) in more detail, drawing on experiences from our own research (4). From these, we draw conclusions for the applicability of case meta-analysis, discussing strengths and weaknesses of the method specifically related to political research and conclude with an outlook for further methodological improvements in the field of meta-analytical and other evidence-based methods (5).

2 Synthesizing original research: Methods in comparison

Synthesis of previous research has a long tradition. We focus here on methods of *secondary* analysis (i.e. analysis of analyses that have already been conducted and typically been published) that synthesize *multiple* studies. Accordingly, we do not cover secondary analysis of single studies that solely aim at analyzing them under a perspective different from that of the original authors.

In order to shed light on the somewhat ambiguous and contradictory terminology in the field, we developed a typology to classify different approaches. This typology is based on two dimensions of research synthesis: the source of data and the method of integration. Both dimensions can be conceptualized either in a qualitative or in a quantitative way.

- *Source of data:* Approaches to synthesize research differ quite substantially as to whether the original research consists of qualitative case studies or of quantitative studies. While case studies typically provide in-depth insights into processes or organizations as a whole, quantitative studies usually refer to a larger number of much more restricted 'cases' (such as standardized interviews), allowing for statistically significant results. Meta-analyzing qualitative case studies implies that the case is the unit of analysis (Lucas 1974: 9). A 'case' can be identical with a particular publication, but it can also be either a part of a publication (one that analyses multiple cases) or consist of data from multiple publications (Bullock and Tubbs 1987). By contrast, quantitative studies are almost exclusively identical with a particular publication. A third type of original study are comparative collective case studies, which will be discussed towards the end of this section.
- *Method of integration:* Original data from different sources can be analyzed in different ways, ranging from a narrative, ad-hoc manner to systematic or quantitative or otherwise highly structured methods.

Table 1 gives an overview of research synthesis approaches according to the used source of data and the method of integration. These are generally applicable to various social sciences and are not limited to political science in particular.

<table>
<tr><th>Source of data
Method of integration</th><th>Qualitative case studies
(unit = case)</th><th>Quantitative studies
(unit = article)</th></tr>
<tr><td>Narrative / ad hoc</td><td colspan="2">Traditional review</td></tr>
<tr><td>Qualitative, interpretive</td><td>Meta-synthesis</td><td>---</td></tr>
<tr><td>Systematic, but not quantitative</td><td colspan="2">Systematic review</td></tr>
<tr><td rowspan="2">Quantitative or otherwise highly structured (statistical or QCA)</td><td colspan="2">Meta-analysis (in a broader sense)</td></tr>
<tr><td>Case survey
(case meta-analysis)</td><td>Meta-analysis
(in the narrowest sense)</td></tr>
</table>

Table 1: Typology of research synthesis approaches according to the used source of data and the method of integration.

The simplest, and possibly the most widely used way to aggregate original data is the *traditional review*. It is used for reviewing every kind of conceptual and empirical literature, including case studies and quantitative studies. Relying primarily on the subjective insight and knowledge of the examiner, traditional reviews lend themselves mainly to exploratory reviews aiming to summarize a certain research literature without applying a strict research question. The advantage lies in the 'holistic' approach in which the examiner can put his or her own judgments of particular studies and compare them in a flexible manner. They present an informed, albeit largely subjective assessment of a particular researcher. The backside of the approach is the lack of transparency and replicability (King et al. 2004). As traditional reviews typically do not develop clear criteria as to which studies are to be included, their results can hardly be replicated by other researchers. "Research confirmed that traditional reviews, in which researchers make relative judgments about what works by using some unknown and inexplicit process of reasoning, were fraught with potential for bias" (Petrosino et al. 2001: 19). When, however, a precise research question is to be answered, there are better, more stringent approaches at hand.

Largely depending on the research goal and overall research approach, the synthesis of qualitative case studies can be either qualitative-interpretive or positivist. The interpretive approach is called *meta-synthesis* (Walsh and Downe 2005). Here, the original author's intentions and categorizations – his or her "understanding of key metaphors, phrases, ideas, concepts, and relations in each study" (Walsh and Downe 2005: 208) – are identified, interpreted and contrasted with those of other case studies. Importantly, the original meaning of concepts by the original author is to be preserved. Sticking so closely to the original authors' framings, the interpretive approach is somewhat at odds with evidence-based research. Positivist approaches, by contrast, reanalyze original data with the aim of answering particular research questions that need not coincide with the intentions of the original authors. This implies that terminologies of the original authors may be adapted to that of the secondary research interest. The simplest form of positivist research synthesis of case study data is the *systematic review*, also referred to as "integrative research review" (Cooper 1982). While some authors equate this with meta-analysis (e.g. Walsh and Downe 2005; Möser 2006), we reserve this term for a systematic, but not quantitative (or otherwise highly structured) method of integrating case study data. "Systematic reviews will include detail about each stage of the decision process, including the question that guided the review, the criteria for studies to be included, and the methods used to search for and screen evaluation reports. It will also detail how analyses were done and how conclusions were reached" (Petrosino et al. 2001: 20). Moreover, quantitative studies can likewise be integrated be means of systematic reviews.

A yet more sophisticated method to integrate qualitative case studies is the *case survey* method (Lucas 1974; Yin and Heald 1975; Larsson 1993), also called *case meta-analysis* (Bullock and Tubbs 1987) or simply meta-analysis (Jensen and Rodgers 2001; Rodgers and Hunter 1992).[3] The case survey method is a particular form of large-*N* meta-analysis. Case surveys integrate qualitative studies, transforming qualitative data into (semi-) quantitative data, using a coding scheme and expert judgments by multiple coders (Lucas 1974; Yin and Heald 1975; Larsson 1993). The results can be analyzed with available analytical methods such as standard probabilistic statistics or set-theoretic approaches such as Qualitative Comparative Analysis (QCA; Ragin

[3] We will discuss below in what way this method, which integrates qualitative case studies, differs from what is commonly known as meta-analysis (integrating quantitative studies).

2004). Thus, case surveys draw on the richness of the case material, on different researchers and research designs, and allow for a much wider generalization than from single cases.

Meta-analysis in the stricter sense refers to the synthesis of *quantitative studies* with quantitative means (see Hunter and Schmidt 2004 for a recent overview). Glass (1977) introduced the term meta-analysis, arguing that the integration of many "weak" research findings (e.g. no strong correlations, or biases) can lead to "strong" findings. The central issue is to relate all original studies to a common effect-size such as correlation coefficients. This, of course, is only possible when the studies to be integrated measure the same kind of empirical objects such as highly standardized polls. It thus does not come as a surprise that meta-analyses are most widely used in the natural sciences, in medicine, in economics, and in other settings where a specific, replicable sort of intervention is being tested. Nelson and Kennedy (2009), for instance, in a "meta-meta-analysis" assessed 140 statistical meta-analyses in environmental and resource economics, concluding that "the incomplete or incorrect methods found in many studies, even if applied to consistently measured effect-sizes, will still restrict the application of meta-analytical results to policy problems".

One important source of data this typology does not cover explicitly is the *comparative collective case study* (Yin 2004; Shkedi 2005), typically with small to medium *N*. In terms of research approach and generalizability, they range somewhere between single case studies and quantitative studies. If performed in a stringent manner, applying the same concepts to all cases, comparative case studies can provide superior insights compared to single case studies. As of now, no established method – beyond (systematic) reviews – is at hand to rigorously integrate findings from different original multiple case studies on a related subject. The difficulty lies in the fact that, not being quantitative, no effect-sizes can be compared as in standard meta-analysis. Nevertheless, comparative case studies can of course be treated as collections of single case studies. These may then – possibly together with other single or multiple case studies – be analyzed through meta-synthesis, systematic review, or case survey. A methodological desiratum would be to develop a methodology to systematically integrate findings from different comparative case studies that explicitly take into account the benefits of comparative research designs of the original studies.

3 The case survey method in political research: Previous applications

Of all research synthesis approaches sketched above, the *case survey* method is unique in that it analyses original qualitative case studies in a rigorous, highly structured (and mostly quantitative) way (cf. table 1). In that respect, it is more challenging, but also more promising in terms of scientific knowledge gains than any of the aforementioned approaches.

The case survey method has originally been developed by researchers at the Rand Corporation for purposes of public policy analysis (Lucas 1974; Yin & Heald 1975). Almost identical approaches have been put forward under the terms of "structured content analysis of cases" (Jauch et al. 1980) or "case meta-analysis" in the field of organizational research and development (Bullock and Tubbs 1987). The methodology has been further developed by scholars from management science (Larsson 1993) and public administration (Jensen and Rodgers 2001, who simply call it "meta-analysis").

Despite its high potential benefits for rigorous integration of case-based data, the application of the case survey method has remained scarce in political research. The most common method of inquiry continues to be the single case study. Several comparative collective case studies are available in different subfields such as negotiated rulemaking (e.g. Coglianese 1997) or public governance (e.g. Bovens et al. 2001). Although typically, due to limited resources, comparative collective case studies are restricted to small sample sizes, some even analyze close to a 1000 cases (e.g. Lubell et al. 2002 study the creation of 958 watershed partnerships in the United States).

Study	*Research question*	*Case selection criteria*	*No of cases*	*No of variables*	*No of raters per case*	*Author participation*	*Scales used*	*Analysis of data*
Yin et al. 1973	Value of forms of citizen participation	Completeness, reliability of data, variety, universe of cases not determined	51	72	1.0	No	Nominal	Bivariate correlation
Yin and Heald 1975	Effects of decentralization in agencies	Time period: 1960-73	215	118	1.1	No	Nominal, 4-5-point	Bivariate correlation
Mintzberg et al. 1976	Structure of unstructured decision processes in organizations	Time period 1965-70	25	21	2.0	n/a	Sequence	Frequency count
Herek et al. 1987	Quality of decision process during crises related to outcome?	Most severe international disputes since WW II with US involvement	19	13	1.2 + 2 external	No	5-point collapsed into binary	Bivariate correlation
Wolf 1993	Bureaucratic effectiveness of administrations	US federal agencies, not authored by employee, at least 15 pages, with effectiveness evaluation	85	15	1.0	No	Mostly 5-point	Bivariate correlation, maximum likelihood
Geist and Lambin 2001	Land-use change: What drives tropical deforestation?	ISI journal articles; loss of forest cover related to causal factors	152	8	?	No	Nominal, nested	Correlation and regression
Beierle and Cayford 2002	Social goals of public participation	Participation, US, environment, at least 5 pages	239	> 100	1.1	No	Likert, 3-point	Bivariate correlation, multivariate analysis
Exadaktylos and Radaelli 2009	Research design in Europeanization studies	SSCI articles on *Europeani?ation* and *Politics,* cited at least 5 times	32 + 32	7	2	No	Binary	Bivariate correlation
Newig and Fritsch 2009	Environmental effectiveness of participation in environmental governance	North America and Europe, sufficient data, at least 10 pages	47	90	1.3	No	Likert, 5-point	Bivariate correlation

Table 2: Overview of case surveys in political science in chronological order of publication.

Although on the whole, the case survey methodology has been used sparsely, a considerable number can be found in the management sciences (e.g. Miller et al. 1991 on the relationship between technology and organizational structure of enterprises; Manimala 1992 on 164 cases on the relation of entrepreneurial heuristics and innovativeness; Larsson and Finkelstein 1999 on 61 cases of merger & acquisition performance, and Winch 2008 on 59 cases of business-to-business services in international markets). Interestingly, most case surveys found in different fields of study concern the assessment of outcomes of certain interventions (medicine, health, education, management, public administration), focusing on "whether it works" (Lipsey 1992).

A comprehensive literature review[4] has revealed the small number of nine published case surveys conducted in the wider field of political science (see table 2). The earliest studies were performed in the context of the Rand Corporation: Both concerned public administration issue: Yin et al. (1973) studied experiences with public involvement in order to guide models for citizen participation in programs by the US department of Health, Education and Welfare in order to help improve these programs. While this study, as the authors themselves acknowledge, in some respects lacked the methodological rigor of the case survey, the second study, exploring effects of decentralization in agencies (Yin and Heald 1975) may be considered "the first fully-developed application of the case survey method" (ibid.: 372). Somewhat different from the typical methodology, Mintzberg et al. (1976) performed a case survey without using any methodological label for their approach. They aggregated data from 25 student reports on decision processes in public and private organizations by applying a simple coding procedure. Apparently, the authors were the first to consistently use two coders for each case. Of the subsequent case surveys (see table 2), the study by Exadaktylos and Radaelli (2009) is noteworthy in that it does not examine real-world political phenomena but research design employed in Europeanization studies. The study by Geist and Lambin (2001) is also particular in that it presents an interdisciplinary undertaking, linking variables from fields such as natural sciences (land-cover change), economics, demography, and political science.

Comparing the nine case surveys identified in the broad area of political science, we find a great variety of specific approaches. A common case selection criterion was completeness of information (at least five studies reported this explicitly); some, but not all studies (e.g. Yin et al. 1973) defined the universe of cases. Some only used cases from particular time periods. At least two studies explicitly restricted their survey to cases published in ISI journals, although this is clearly not recommended in case survey methodology (see below). Numbers of surveyed case studies range from 19 to 239; the number of variables from 7 to 118. Surprisingly, only two of the nine studies consistently used two raters and none of them more than two, although parts of the literature recommend three or even more raters. Also, none of these surveys has drawn on author participation as recommended by Larsson (1993). The by far most popular method of analysis was bivariate correlation.

[4] Starting from the classical methodological literature (Lucas 1974; Yin and Heald 1975; Bullock and Tubbs 1987; Larsson 1993), we searched the ISI databases for articles citing these. Moreover, we searched the ISI Social Science Citation Index for articles on "case survey", yielding 16 results, most of which belong to the domain of business administration.

4 How to conduct a case survey?

Summarizing the most important steps of a case survey, we largely draw on the contributions by Bullock and Tubbs (1987) and Larsson (1993), who provide the most detailed accounts on how to conduct a case survey in practice. While we basically follow the logic provided by these authors, we amend the suggested research process owing to our own experiences with the case survey methodology (cf. the overview in table 3).

Bullock & Tubbs (1987)	*Larsson (1992)*	*Our suggestion*
1 Develop research questions	1 Develop research questions	1 Develop research questions
		2 Decide on the methodology
2 Set criteria for study collection	2 Set case selection criteria	3 Define case selection criteria
3 Collect studies	3 Collect case sample	4 Collect case sample
4 Develop coding scheme	4 Design coding scheme	5 Design initial coding scheme
		6 Pretest and iterative revision of coding scheme
5 Code studies	5 Code cases through multiple raters	7 Final coding of cases through multiple raters
	6 Code cases by original authors	
6 Compute reliability, estimates of coding quality	7 Measure interrater reliability	8 Measure interrater reliability
7 Resolve rating discrepancies	8 Resolve coding discrepancies	9 Resolve important, but not all, coding discrepancies
	9 Statistical analysis of coding validity	
	10 Statistical analysis of the impact of specific case study characteristics	10 Statistical analysis of biases
8 Analyze consensus data matrix	11 Statistical analysis of the created case data set	11 Analysis of the created case data set (statistical or other)
9 Report results	12 Report the study	12 Report the study

Table 3: Process schemes for conducting a case survey.

1. Develop research questions. As a highly structured approach to integrating case-based knowledge, the case survey methodology evidently profits from a clear research question or questions. These need not be specific hypotheses to be tested but can also be of a more exploratory nature. But even then, a thorough reflection on the kind of knowledge one seeks to gain from the comparison of case study experience appears essential. Typically, research questions will already have been developed before making the decision to study these by means of a case survey.

2. Decide on the methodology. Both the nature of the research question(s) and the availability of case-based or other previous research will determine whether or not a case-survey will be conducted at all and if so, how precisely this will be done. Case surveys are particularly useful when case studies dominate the area of research, when a broad range of conditions is of interest and when an experimental design is impossible (Larsson 1993). Before conducting a case survey, decisions on important methodological issues should be taken. The more cases are to be in-

cluded, the more important it will be to decide early e.g. on the number of coders and whether or not original authors will be surveyed in addition to written material, for this has considerable consequences on the resources needed (time, expertise and personnel).

3. Define case selection criteria. In order to be sure about the generalizability of the study, it is most important to define the population or 'universe' of empirical phenomena to be studied (Lucas 1974; Ragin 2006), answering the question: What could possibly be a case? This includes considerations about possible "control groups": If a researcher is interested, say, in studying the different effects of participation in environmental governance, he or she ought to include both participatory and non-participatory cases (or more and less participatory ones) of a 'universe' defined as environmental decision procedures.

As opposed to more traditional reviews, it is advised not to restrict the universe of cases by publication status or methodological rigor of the studies. This is because the case survey allows to control for the impact of these aspects as part of the analysis. Excluding studies because of their publication status or supposed 'lack' of methodological rigor could introduce unwanted biases. "Any presumption by the reviewer that rigor [and publication status] and results are associated is converted to coded variables and tested statistically" (Bullock and Tubbs 1987: 184). However, it is worth considering carefully whether – particularly – older studies could not provide important insights as well. As with the above criteria, the time of the case study can and should always be controlled for as third variable. In terms of thematic scope of the study, it seems useful to define the universe of cases quite tightly and study a significant sample of it rather than defining a broad universe and studying a comparatively small sample of it.

4. Collect case sample. In order to construct a representative sample of case studies from the above defined "universe of cases", as many case studies as possible should be identified, covering as many sources as possible (Larsson 1993: 1530). Then either this totality or a random subset of cases can be used for coding. A criterion to exclude studies could be lacking data. Especially with more complex research questions and resulting number of variables as was the case in our study (Newig and Fritsch 2009), a minimum level of detail regarding e.g. context, process and results was needed (we considered only cases study accounts with a minimum of ten pages).

Two particular issues may arise concerning the question of what constitutes a "case". First, there may be more than one publication on a particular subject (e.g. organization or decision process). Then either only one of these publications may be used or, as we recommend, all available publications should be merged during coding, together constituting one case (see Bullock and Tubbs 1987). Second, a greater challenge is posed by contexts in which it is difficult to delimit a case either spatially or temporally. For instance, in our case survey on the effectiveness of participatory processes, we came across "cases" in which a long series of events, including a number of different and more or less 'successful' participation and decision processes were carried out on a particular issue and in the same community over periods of more than a decade. Integrating all these subsequent procedures with all their dynamics as a single "case" would have been very difficult, if not impossible in comparison with more temporally limited processes. We therefore decided to select one or more of these procedures and code these as separate cases. Prior experience in the same community was then coded as context variables such as the existence of trust, conflicts, or social networks; and subsequent events were coded as outcome variables.

5. Design initial coding scheme. Based on the theoretical considerations in step 1, a coding scheme is to be developed. "A coding scheme is a set of decision rules. These decision rules ex-

plicitly define for a rater how to convert some [qualitative] information to a variable code" (Bullock and Tubbs 1987: 189). While most coding rules (variables or survey questions) derive directly from formulated hypotheses, room may be given for other variables that have little conceptual foundation but might nevertheless turn out – during a pre-coding phase – to be important in explaining the consequences of participatory governance, thus enabling an inductive analysis to complement the mainly deductive, hypothesis-testing approach. The coding scheme constitutes the 'core' of a case survey, "documenting and guiding the conversion of qualitative case study data into quantified variables" (Larsson 1993). Thus, care should be taken to develop a good coding scheme.

We recommend a detailed and rather comprehensive scheme for reasons of research quality, efficiency and precaution: (1) Collecting much information from each case enables a thorough and meaningful analysis. (2) Selecting, reading and coding each case presents a considerable effort, no matter how simple the coding scheme. Thus, coding, e.g. twice as many variables, implies considerably less than twice the effort. (3) Often, research designs are open for 'surprises', e.g. the possibility of unpredicted correlations between variables should be considered right from the start and allowed for in the coding scheme. Therefore, more rather than fewer variables should be included. It is always possible to later omit or aggregate data, but not vice versa. Finally, separate fields can be used to document the reliability of data for each variable, allowing for the separate statistical analysis of subsets of data that meet certain reliability standards.

6. Pretest and iterative revision of coding scheme. From our experience, we recommend to perform a pretest of the coding scheme. A small number of cases (e.g. three or five) may be test-coded independently by at least two different raters. This allows to check whether the coding rules are clearly understandable, whether all important information from the cases regarding the research questions is actually covered by the coding scheme, and whether coding rules are clear enough to yield a high interrater reliability. The latter should be computed according to step 8. One way is to test-code as many cases as needed and iteratively revise the coding scheme until a certain, desired interrater reliability is reached. From this point on, the coding scheme should remain absolutely unaltered.

7. Final coding of cases through multiple raters. The most laborious, time- and resource-intensive step is typically the final coding of all sample cases. In order to be able to compute interrater reliability, all cases should be read and coded by at least two trained persons. Training ought to be performed on test cases (such as the ones used in step 6), which will not be included in the final data set of coded cases in order to prevent the learning effect to bias the results. There is some disagreement in the literature regarding whether to use two, three or more raters. Bullock and Tubbs (1987) found two well-trained raters adequate, whereas Larsson (1993) argues for three or more raters. We discuss this when turning to the resolving of coding discrepancies (step 9), which is actually part of the coding procedure. Whatever the procedure, there will typically be one single code for each variable and case. For reasons of transparency and in order to minimize coding errors, it is important to code unavailable data with a special code rather than leaving cells in the database blank. From our experience (Newig and Fritsch 2009), reading and coding (90 variables) a written case study (20 to 60 pages) takes a trained coder typically between 5 and 10 hours, not counting the time needed for resolving coding discrepancies

8. Measure interrater reliability. Several measures are at hand to measure interrater reliability. The simplest one is percent agreement, calculating for one variable or case the fraction of identi-

cal codes of all codes. This measure is, however, largely dependent on the used scales and the number of raters, for it is more likely that percent agreement is higher with few raters and with binary variables than with more differentiated scales. For a thorough discussion of possible measures cf. Bullock and Tubbs (1987: 195-200); Larsson (1993: 1533-4). It is important to note that not only the scales used for coding and the number of coders, but also the used measures greatly influence the reported interrater reliability of different case surveys, making these figures difficult to be compared directly.

9. Resolve important, but not all, coding discrepancies. Using at least two independent coders brings up the issue of how to resolve discrepant codes. A recommended method is consensus rating, meaning that discrepant codes are discussed among all coders and subsequently resolved. Working with three raters allows to resolve persisting disagreement by "majority vote" and also to detect simple coding errors more easily. When using more detailed scales, such as 5-point Likert scales, it does not seem necessary to us to resolve all coding discrepancies by consensus. The reason is that coding qualitative cases ultimately involves a subjective, interpretive element such that the "expert judgment" of one person need not fully equal that of another. We therefore suggest that only major discrepancies (such as those with a difference of 2 or more on a 5-point scale) be discussed. After discussion, and for all minor discrepancies, the arithmetic mean should be computed in order to obtain one single coding for each variable and case.

10. Statistical analysis of biases. A number of possible biases should be examined and, if possible, controlled for statistically. For an in-depth discussion of author, publication, selection, coding and other biases cf. Beierle and Cayford (2002, Annex D). Generally, all factors possibly introducing bias (such as publication status or reliability of data) should be coded as variables and then controlled for statistically.

11. Analysis of the created case data set (statistical or other). Manifold the methods to analyze large amounts of codes as generated through a case survey. Next to bivariate correlations, regression analysis, path analysis and other established techniques may be performed, depending on the respective analytic intentions. Prior to this kind of analysis, aggregative methods on variables (such as factor analysis) or on cases (such as cluster analysis) may be performed. As an alternative to statistical methods, which generally assume linear independence of variables, set-theoretic approaches such as QCA (Qualitative Comparative Analysis, Ragin 2004) may be suitable as well.

12. Report the study. Case surveys, although a relatively simple undertaking compared to the efforts of conducting a larger number of original case studies, typically produces a large amount of data and methodological accounts. These include the documentation of research steps (which cases to be included and why), the coding scheme (which may well spread over a dozen of pages), and certainly the codes for all cases and variables. Add to this statistical analyses and graphical representations. While for journal articles and even books it is often impossible to document all this material, it is desirable to make it publicly available in order to allow full replication of the conducted case survey (Bullock and Tubbs 1987). Nowadays, cheap web space makes it possible to store this material on publicly accessible internet pages.

5 Conclusions and outlook for further methodological developments

Previous applications have demonstrated that the case survey can be a powerful methodology for rigorously synthesizing and integrating case-based knowledge in political research. It is applicable across a wide spectrum of research topics, ranging from organizational studies to decision processes, outcome evaluations and even research strategies of other authors. This last point (see Exadaktylos and Radaelli 2009) shows that case survey not only can be used to integrate empirical real-world findings, but also to synthesize research on other levels of analysis as well.

Placing the case survey methodology in the context of other strategies to synthesize and integrate primary research shows two important implications: First, the type of original data (qualitative case studies or quantitative studies) restricts the choice of available strategies of integration. Meta-analysis, for instance, can only be performed on quantitative studies, while the case survey presupposes qualitative cases. Second, within this limited scope, researchers can choose among a range of different strategies of integration, ranging from narrative to quantitative approaches. Given the prevailing ambiguites regarding the concepts of "case survey", "meta-analysis" and the like, our typology of research synthesis approaches seeks to contribute to conceptual clarification.

The review of completed case surveys in political research reveals a number of insights: First and foremost, the methodology – under whatever label – is still rarely being employed, with only three studies in the 1970s, one each in the 1980s and 1990s, and four in the 2000s. Other social sciences, notably management research, have developed a more lively tradition of case survey studies. Second, hardly any of the case surveys in political science meet the basic quality criteria put forward by Bullock and Tubbs (1987) and Larsson (1993), mostly because they do not consistently draw on at least two raters, or because they exclude studies due to their publication status. Surprisingly, none of these studies has used author participation.

Regarding the analytical steps necessary to perform a case survey, we have introduced a number of small amendments to the previous state of the art. We propose to perform a pre-test of the coding-scheme, which allows to test and iteratively revise the scheme as possible ambiguities in the coding instructions may occur. Second, we have developed a way how to define what is a 'case' in rather long and complex histories.

The case survey method is generally applicable to a wide range of research subjects and not restricted to political science. However, those areas of political science that can profit most from the methodology entail particular challenges. For instance, complex decision-making processes in the public realm, involving a large number of state and non-state actors and spanning larger time-frames can make it difficult to clearly define what is a 'case': Not only does the social context play an important role, but also path dependencies in a complex series of events have to be taken into account. Moreover, the temporal delimitation is difficult. We propose to single out a partiularly interesting part of a larger series of historical events and processes, and code prior events as context variables, and subsequent events as results of the focal process.

Will the case survey methodology turn out to be a panacea for gaining reliable knowledge on complex issues of political research? We do believe that the method is strongly under-utilized and bears much potential to bring research integration great steps forward. However, we also see other methods of inquiry appearing on the horizon that can complement case-survey re-

search. Next to case studies, mass surveys and research syntheses we picture field experiments to re-gain importance in political research. Experimental methods are becoming increasingly important in social science (Oakley et al. 2003). Their obvious advantage over classical observations (single and comparative case studies) lies in the possibility for unbiased inference about causal relations. Whereas laboratory settings allow for precisely controlled contexts (such as in experimental economics), real world (field) experiments combine the advantage of natural political contexts with methodological benefits of random assignment (Druckman et al. 2006). In political science, experimental research is increasingly used, but still on a very low level and far from being established as state-of-the art (Green and Gerber 2003). Applications mostly involve mass political behaviour such as in political psychology, electoral politics and legislative politics (Druckman et al. 2006: 627). In contrast, few if any studies have been conducted in complex areas such as governance research. Given the growing experience with experimental methods, the time seems now ripe to attempt experimental research in such complex areas of political and administrative science (Stoker and John 2009).

Whereas case surveys are paricularly useful to aggregate and integrate qualitative reserach from very different sources and in very different contexts, these could be complemented by field experiments for particular (sub-)types of cases, drawing on a subset of variables of those used in a case survey's coding scheme. The combination of these methods opens up promising perspectives for future evidence-based research.

Acknowledgements

Parts of this research have been conducted within the ongoing projects "Environmental Consequences of Participatory Governance – A comparative meta-analysis of case studies in environmental decision-making (ECOPAG)", funded by the German Research Foundation (2009–2012), and "Multi-level Governance of Natural Resources: Tools and Processes for Water and Biodiversity (GOVERNAT)", a Marie Curie Research Training Network funded by the European Commission under the 6th Framework Programme, Contract No. 0035536 (2007-2010).

References

Beierle, T.C. and J. Cayford. 2002. *Democracy in Practice. Public Participation in Environmental Decisions*. Washington, DC: Resources for the Future.

Bovens, M., P. 'T Hart and B.G. Peters. 2001. *Success and Failure in Public Governance: A Comparative Analysis*. Cheltenham, UK: Edward Elgar.

Bullock, R.J. and M.E. Tubbs. 1987. 'The Case Meta-Analysis Method for OD.' *Research in Organizational Change and Development* 1: 171-228.

Coglianese, C. 1997. 'Assessing Consensus: The Promise And Performance Of Negotiated Rule-making.' *Duke Law Journal* 46: 1255-346.

Cooper, H.M. 1982. 'Scientific Guidelines for Conducting Integrative Research Reviews.' *Review of Educational Research* 52 (2): 291-302.

Druckman, J.N., D.P. Green, J.H. Kuklinsky and A. Lupia. 2006. 'The Growth and Development of Experimental Research in Political Science.' *American Political Science Review* 100 (4): 627-35.

Exadaktylos, T. and C.M. Radaelli. 2009. 'Research Design in European Studies: The Case of Europeanization.' *Jcms-Journal of Common Market Studies* 47 (3): 507-30.

Geist, H.J. and E.F. Lambin. 2001. *What Drives Tropical Deforestation? A meta-analysis of proximate and underlying causes of deforestation based on subnational case study evidence. LUCC Report Series No. 4*. Louvain-la-Neuve.

Glass, G.V. 1977. 'Integrating Findings: The Meta-Analysis of Research.' *Review of Research In Education* 5: 351-79.

Green, D.P. and A.S. Gerber. 2003. 'The Underprovision of Experiments in Political Science.' *The Annals of the American Academy of Political and Social Science* 589: 94-112.

Herek, G.M., I.L. Janis and P. Huth. 1987. 'Decision-Making During International Crises - Is Quality Of Process Related To Outcome?' *Journal of Conflict Resolution* 31 (2): 203-26.

Hunter, J.E. and F.L. Schmidt. 2004. *Methods of Meta-analysis. Correcting Error and Bias in Research Findings*. Thousand Oaks, London, New Delhi: Sage Publications.

Jauch, L.R., R.N. Osborn and T.N. Martin. 1980. 'Structured Content Analysis of Cases: A Complementary Method for Organizational Research.' *The Academy of Management Review* 5 (4): 517-25.

Jensen, J.L. and R. Rodgers. 2001. 'Cumulating the Intellectual Gold of Case Study Research.' *Public Administration Review* 61 (2): 235-46.

King, G., R.O. Keohane and S. Verba. 2004. The Importance of Research Design. In *Rethinking Social Inquiry. Diverse Tools, Shared Standards*, eds. H.E. Brady and D. Collier. Lanham, Boulder, New York, Toronto, Oxford: Rowman & Littlefield Publishers: 181-92.

Larsson, R. 1993. 'Case Survey Methodology: Quantitative Analysis of Patterns across Case Studies.' *The Academy of Management Journal* 36 (6): 1515-46.

Larsson, R. and S. Finkelstein. 1999. 'Integrating Strategic, Organizational, and Human Resource Perspectives on Mergers and Acquisitions: A Case Survey of Synergy Realization.' *Organization Science* 10 (1): 1-26.

Lipsey, M.W. 1992. Meta-analysis in Evaluation Research: Moving from Description to Explanation. In *Using theory to improve program and policy evaluations*, eds. H.-t. Chen and P.H. Rossi. New York et al.: Greenwood Press.

Lubell, M., M. Schneider, J.T. Scholz and M. Mete. 2002. 'Watershed Partnerships and the Emergence of Collective Action Institutions.' *American Journal of Political Science* 46 (1): 148-63.

Lucas, W.A. 1974. The case survey method: Aggregating case experience. Santa Monica.

Manimala, M.J. 1992. 'Entrepreneurial heuristics - a comparison between high PI (pioneering-innovative) and low PI ventures.' *Journal of Business Venturing* 7 (6): 477-504.

Miller, C.C., W.H. Glick, Y.-D. Wang and G.P. Huber. 1991. 'Understanding technology-structure relationships: theory development and meta-analytic theory testing.' *Academy of Management Journal* 34 (2): 370-99.

Mintzberg, H., D. Raisinghani and A. Théoret. 1976. 'The structure of "unstructured" decision processes.' *Administrative science quarterly* 21 (2): 246-75.

Möser, G. 2006. *Systematic review and quantitative meta-analysis: Methodological foundations and practical applications in the domain of transport policy*. Gießen.

Nelson, J. and P. Kennedy. 2009. 'The Use (and Abuse) of Meta-Analysis in Environmental and Natural Resource Economics: An Assessment.' *Environmental & Resource Economics* 42 (3): 345-77.

Newig, J. and O. Fritsch. 2009. 'Environmental Governance: Participatory, Multi-Level – And Effective?' *Environmental Policy and Governance* 19 (3): 197-214.

Oakley, A., V. Strange, T. Toroyan, M. Wiggins, I. Roberts and J. Stephenson. 2003. 'Using Random Allocation to Evaluate Social Interventions: Three Recent U.K. Examples.' *The Annals of the American Academy of Political and Social Science* 589: 170-89.

Petrosino, A., R.F. Boruch, H. Soydan, L. Duggan and J. Sanchez-Meca. 2001. Meeting the challenges of evidence-based policy: The Campbell Collaboration. Paper read at 2001 Jerry Lee Crime Prevention Symposium, Apr, at College Pk, Maryland.

Ragin, C.C. 2004. Turning the Tables: How Case-Oriented Research Challenges Variable-Oriented Research. In *Rethinking Social Inquiry. Diverse Tools, Shared Standards*, eds. H.E. Brady and D. Collier. Lanham, Boulder, New York, Toronto, Oxford: Rowman & Littlefield Publishers: 123-38.

Ragin, C.C. 2006. 'How to Lure Analytic Social Science Out of the Doldrums. Some Lessons from Comparative Research.' *International Sociology* 21 (5): 633-46.

Rodgers, R. and J.E. Hunter. 1992. 'Foundation of Good Management Practice in Government: Management by Objectives.' *Public Administration Review* 52 (1): 27-39.

Shkedi, A. 2005. *Multiple case narrative: a qualitative approach to studying multiple populations*. Amsterdam: Benjamins.

Stoker, G. and P. John. 2009. 'Design Experiments: Engaging Policy Makers in the Search for Evidence about What Works.' *Political Studies* 57 (2): 356-73.

Walsh, D. and S. Downe. 2005. 'Meta-synthesis method for qualitative research: a literature review.' *Journal of Advanced Nursing* 50 (2): 204–11.

Winch, G.M. 2008. 'Internationalisation strategies in business-to-business services: the case of architectural practice.' *Service Industries Journal* 28 (1): 1-13.

Wolf, P.J. 1993. 'A Case Survey of Bureaucratic Effectiveness in U.S. Cabinet Agencies: Preliminary Results.' *Journal of Public Administration Research and Theory* 3 (2): 161-81.

Yin, R.K. 2004. *Case study research : design and methods*. 3. ed, *Applied social research methods series ; 5*. Thousand Oaks, Calif. [u.a.]: Sage.

Yin, R.K. and K.A. Heald. 1975. 'Using the Case Survey Method to Analyze Policy Studies.' *Administrative Science Quarterly* 20: 371-81.

Yin, R.K., W.A. Lucas, P.L. Szanton and J.A. Spindler. 1973. *Citizen Organizations: Increasing Client Control over Service.* . Santa Monica, CA: Rand Corporation.