Understanding and Researching Complexity with Qualitative Comparative Analysis: Evaluating Transportation Infrastructure Projects

Stefan Verweij & Lasse Gerrits

This is the version of the article that was accepted for publication by the journal. The final, definitive article has been published in Evaluation by SAGE Publications Ltd, all rights reserved.


Abstract

This article proposes a complexity-informed framework for evaluating transportation infrastructure projects. This is done through four steps. First, the properties of infrastructure development projects are discussed. This leads to the conclusion that the specific locality or contextuality of a given project is important for explaining the outcome. Hence, there is a need for an ontology and epistemology that addresses the importance of this contextuality. The second step concerns the development of the prerequisites for a methodological framework that follows from this epistemology and ontology. The third step is the assessment of common infrastructure evaluation methods against these prerequisites. This leads to the conclusion that a comparative case-based approach is the most suitable way to study the relationship between context and outcomes in projects. A framework based on qualitative comparative analysis (QCA) is presented in the fourth step. The article concludes with a discussion of the further development of QCA.

Keywords

Complexity, Evaluation, Qualitative Comparative Analysis (QCA), Transportation Infrastructure Projects
Understanding and Researching Complexity with Qualitative Comparative Analysis: Evaluating Transportation Infrastructure Projects

Introduction

Transportation infrastructure projects are characterized by budget overruns, time delays and public resistance. For example, the decision-making process for infrastructure development in the Netherlands takes about 11 years on average (Advisory Committee VBI, 2008). Despite this, cost overruns above 40 percent occur frequently, and sometimes exceed 80 percent (Flyvbjerg et al., 2003a). This is not a new finding: a European study from a decade before had similar results (WRR, 1994), and the effectiveness, efficiency and legitimacy of infrastructure development have been debated in the Netherlands since the 1970s (De Hoo, 1982). The long-running discussion highlights the persistence of the complexity underlying infrastructure development and the lack of any quick fixes.

The poor performance of such projects has led to various attempts to better understand infrastructure development, and the causes and solutions of the problems the sector faces. For instance, the Netherlands Institute for Transport Policy Analysis (KiM) strongly advocates the need for more and better infrastructure development evaluation, emphasizing the need to improve our understanding of why projects perform the way they do (KiM, 2009). However, it is more difficult to assess ‘what’ causes than ‘that’ causes (Flyvbjerg et al., 2003a). That is, it is relatively easier to identify patterns of cost overruns and time delays, for example, than the underlying reasons why they occurred. Project evaluation has tended to focus on the comparison of ‘before and after’ situations, and has not adequately incorporated the influence of contextual local conditions on infrastructure project development (Sanderson, 2000). The dominant evaluation methodologies impede policy learning since they do not account sufficiently for the complex nature of policy systems (Sanderson (2000). KiM also similarly commented that in addition to ex-post evaluations being carried out on a marginal basis, current evaluations also suffered from methodological deficiencies (KiM, 2009; PBL and KiM, 2010), which are related to the misfit between the way infrastructure development projects are understood and the methodologies used to evaluate them. Hence, there is a need for evaluation methodologies that do justice to the complex nature of infrastructure projects.

Although evaluation methodologies should pay attention to the contextual nature of projects to account for the influence of idiosyncratic events, this does not imply an exclusive focus on single-N in-depth evaluations. After all, evaluation aims to improve infrastructure development practice, which requires some degree of generalization (cf. Sanderson, 2000), and solely focusing on one particular context of one particular case does not allow recurring patterns to be analyzed. At this point, as will be shown in the following section, the requirements of contextuality and generalization seem to be in conflict with each other. On the one hand, the dominant
methods used today are often inadequate because they discount local conditions in cross-case comparative studies. On the other hand, generalization is difficult from in-depth case studies.

This article aims to reconcile this dilemma. It presents a complexity-informed evaluation framework based on qualitative comparative analysis (QCA) (Ragin, 1987; 2000; 2008a) for evaluating complex infrastructure development projects. This method integrates the generic patterns of variable-oriented studies with the idiosyncratic events of case-based studies. Before presenting this method, we clarify what complex infrastructure projects are so as to establish their fit with the method(s) used to evaluate them (cf. Sanderson, 2000). This article aims to answer the following question: what are the ontological, epistemological and methodological components of a method for assessing the influence of idiosyncratic events and recurring patterns on infrastructure development projects and what should a coherent evaluation method based on these components consist of?

Generalization versus contextuality

Flyvbjerg et al. (2003a; see also Flyvbjerg et al., 2002; 2003b; 2004; 2005; Flyvbjerg, 2007b; 2009; Næss et al., 2006) compared 258 large infrastructure projects in 20 different countries, and found a number of patterns. For instance, actual costs are on average 45 percent higher than estimates, cost overruns are a global phenomenon, and cost performance has not improved over the past 70 years. Although this study and its findings are of indisputable importance in accounting for the ‘that’ of infrastructure development performance, it cannot explain the why of cost overrun (cf. Flyvbjerg et al., 2003a). That is, while their study finds some differences between rail, road and fixed link projects, and between project types within these categories, this large-sample study, by its very nature, cannot explain the influence of local conditions on cost performance. For example, Flyvbjerg et al. (2003a: 19) state that:

‘For the Channel Tunnel, changed safety requirements were a main cause of overrun. For the Great Belt link, environmental concerns and accidents with flooding and a devastating fire made the budget balloon. For the Øresund link, it proved more costly than estimated to carve major new transport infrastructure into densely populated Copenhagen, and so on.’

In general, the study points to many similarities amongst projects in terms of their performance. For instance, the Øresund link and Channel Tunnel are reasonably similar regarding the percentage of cost overrun (i.e. respectively 70 and 80 percent) (Flyvbjerg et al., 2003a). However, the causal paths leading to those results are different. An ex-post evaluation into the cost performance of the Channel Tunnel revealed that, in addition to the findings of Flyvbjerg and colleagues, other factors also led to cost overruns (Anguera, 2006). These included the absence of a clear project owner from the outset, the unforeseen advent of low cost airlines leading to reduced train ridership, political events involving the French and British
governments, difficult ground conditions, and transport-related incidents such as the PanAm crash at Lockerbie. Although the Øresund link project is similar in some of these aspects, these specific conditions do not account for its cost overruns.

This brief example points to the tension between generic patterns and factors that are specific to a certain situation. On the one hand, large-N quantitative studies such as those performed by Flyvbjerg and colleagues provide lasting insights into generic patterns such as cost overruns and their causes. However, these studies do not allow a detailed analysis of the idiosyncratic nature of such projects, even though specific events may have significantly contributed to the project outcomes. On the other hand, studies such as Anguera (2006) provide important insights into the unique and idiosyncratic nature of individual projects. However, such case-oriented studies have little to contribute to the development of patterns from different cases. By their very nature, both types have their advantages and disadvantages. Table 1 provides an illustrative list of both types of infrastructure development studies (excluding the studies by Flyvbjerg and colleagues mentioned above).


Table 1: variable- and case-oriented studies in transport infrastructure development

Matching ontology and epistemology with methodology

Understanding complex infrastructure projects

If an infrastructure project is said to be complex, it usually means that it is perceived to be difficult. However, complexity is not just a generic statement about the effort necessary to complete a project, nor is it a truism. Instead, it is a multi-layered concept. In abstract terms, developing infrastructure means modifying an existing system. An example of a system is a built area, which consists of interacting three-dimensional units (e.g. rooms, buildings, assemblages of buildings), two dimensional units (i.e. the layout or distribution of the three-dimensional units across a given space), and linear units or transport networks linking the three-dimensional units, thereby largely determining the area’s layout (Marshall, 2009). Together, these form a built syntax specific to a given area. While some patterns of activity can be found in various examples of infrastructure development, e.g. suburbanization and subsequent commuter travel patterns, other properties are specific to a local situation, e.g. particular difficult ground conditions that would make the construction
of a road to a future suburb prohibitively expensive. Therefore, an infrastructure
developer wishing to change something in a specific situation deals with a specific
built syntax that is a mixture of generic elements and local conditions.

Over time, the interaction of generic and specific elements leads to a local built
environment that is unique in its particularity, even though it still has recognizable
elements (Byrne, 1998; 2001; 2003; 2005; Marshall, 2009). A built area becomes
even more complex if its social fabric is taken into account – individuals and social
groups living, working, traveling and recreating in any given area – since it directly
influences the existing and future infrastructure requirements. This syntax of built
and social components is nested in the way that properties of a subunit (e.g. a street)
reflect the properties of its whole (e.g. a district), but not to the extent that both levels
are exact copies (Marshall, 2009). Thus, the local built order emerges from the
interaction between generic and specific physical and social elements (Allen, 1998).
Hence, infrastructure developers have to deal with a unique local area that
nonetheless exhibits similarities with other areas. Developing infrastructure in built
areas is, therefore, not just a matter of applying generic planning, building or
managerial rules to a certain area; these need to be adapted to fit the local conditions.
It is pivotal for the development of infrastructure in built areas that this specific
pattern of local conditions and generic developments is researched to understand ex-
ante how a project should be executed, and to understand ex-post what leads to
certain outcomes.

This perspective focuses on a number of dimensions. First, infrastructure
development takes place within a specific interacting mix of local conditions and
generic patterns that occurs in any given location. Secondly, this points to the fact
that the causal relationships between site-specific conditions and generic
developments are poorly known and, if known, only for that specific time and place.
Thus, known causal relationships specific to a certain area are by definition case-
specific. Indeed, the unique nature of built systems implies that other systems are
constituted differently, although the emerging order can be quite similar. Thirdly,
this emergent nature of any built area implies that it is the result of longitudinal
development. That is, it is the result of past changes and events that are to some
extent path-dependent. Taking these three points together, this article understands
built areas as complex systems (e.g. Batty, 2010). The next question then is: how
should this complexity as such be named, understood and researched? This three-
part question is answered in the following section, resulting in the formulation of
requisites for complexity-informed evaluation.

Foundations for understanding and researching situated complexity

The complexity described above can be characterized in two ways: simplistic or
generic complexity and complex or situated complexity (Byrne, 2005; Cilliers, 2001).
Generic complexity focuses on the emergence of complex processes and structures
from a limited set of variables. It assumes a general set or rules from which emergent
complexity flows (Buijs et al., 2009). Although elegant, this approach is only part of
the issue, since, as we noted in the previous section, the emergent nature of infrastructure projects is partly determined by local conditions. Built areas are open systems (cf. De Roo, 2010), meaning that their composition and behavior is constituted through interaction with their environment, resulting, as stated above, in a specific locally-situated mix of generic and specific elements. Considering built areas as open systems assumes that an explanation for the development (or lack thereof) of a project can be found in its contextuality, i.e. that local conditions contain explanatory variables (cf. Mjøset, 2009).

Buijs et al. (2009) use ‘situated complexity’ to focus on the explanatory value of the contextuality of a phenomenon. Although some argue that a research methodology should start from either the generic or situated approach (e.g. Bar-Yam, 1997), Buijs et al. (2009) state that a case can and should be made for systematic in-depth comparison across systems. They argue that while open systems ‘do not operate according to general rules applied in all contexts’ (2009: 37), a systematic comparison can reveal differences and similarities between the operations of different systems. This approach to situated complexity focuses both on recurring patterns over multiple systems and the idiosyncratic events in particular systems, since both determine how systems develop over time. The research methodology presented in this article starts from that premise.

The second part of the question concerns the way complexity is understood, which is basically a question of how reality can be understood. The classical divide is between positivism and postpositivism. To some extent, this divide coincides with the difference between generic and situated complexity. Positivism is primarily concerned with determining general rules by taking reality apart in discrete components, which matches the aims of studies of generic complexity. However, postpositivism has many different sub-strands that range from the extreme relativism of social constructivism to the more realist thesis of negotiated subjectivism (Byrne, 2003; Haynes, 2001; Uprichard and Byrne, 2006) or critical realism (Guba and Lincoln, 1989). The common theme within those strands is that the contextuality is explanatory for what is being observed, which coincides with situated complexity.

The fact-value dichotomy that underlies the positivist stance has been thoroughly undermined (Bateson, 1984; Byrne, 2002; Fischer and Forester, 1987; 1993). Complex and systemic causality is always subject to interpretation and consequently debatable as every interpretation carries with it normative judgments (Williams, 2009). If systems are said to be open, than it follows that their boundaries do not exist a priori, and any individual will develop a particular demarcation or set of boundary judgments about the system which includes and excludes variables (i.e. a reduction of real complexity) that may be connected but not perceived as such by the observer (Checkland, 1981; Cilliers, 2001). Thus, there is no unambiguous separation between systems and their context, and the observer is as much part of the complexity as the system or agents that are observed. Situated complexity is therefore not confined to the presupposed demarcations of systems but intersects all system representations by respondents.
This reduction of real complexity both compromises the research, as well as keeps it manageable (Cilliers, 2001; 2005). It requires that multiple perspectives on a particular system are taken into account. If a multitude of observers can develop a multitude of boundary judgments about what is taken into account or not with a multitude of perspectives about how something is being perceived, chances are that a larger part of the complexity of the system is captured (Cilliers, 1998; 2005). This type of thinking implies a convergence of the fact-value dichotomy, but not a postmodern stance where subjective storytelling is all that remains. It means that, although temporal in time and place (i.e. specific to a given locality), cause and effect relations do exist and can be known through respondents’ perceptions (i.e. it is agent-bound) (Byrne, 2001; 2003; 2005; Morçöl, 2001). Causality can still be determined in terms of change and response relationships that lead to certain observable effects (cf. Bryman, 2004; Hammersley, 2008; 2009). The ontological point of departure in this article is therefore complex realism (cf. Byrne, 2002; Harvey, 2009; Reed and Harvey, 1992).

Researching situated complexity from the perspective of complex realism requires a methodology that focusses both on recurring patterns over multiple systems as cases and on systemic peculiarities that acknowledges that these patterns and peculiarities are interpretations but point to real causalities nevertheless. Taking these points together, the first criterion for a complexity-informed methodological framework is [1] that it balances between in-depth understanding and reductionist generalization. Second, since studying situated complexity requires an in-depth understanding of cases, [2] the method has to be case-based. Moreover, since single case studies cannot be employed for explanatory purposes in other cases and, therefore, do not allow statements about patterns across systems, a comparative case study approach is required. Third, as stated above, situated complexity focuses on the explanatory value of contextuality. This contextuality emerges from the interaction of generic patterns and specific events and therefore it can be inferred that, in methodological terms, [3] the method should allow the observation and analysis of complex interaction between the variables. This is discussed in the following sections. Finally, since complex systems such as built areas are the result of longitudinal development, [4] the method has to consider how situated complexity came into being over time (i.e. complex dynamics).

Towards a complexity informed case-comparative framework

These four requirements are only partly fulfilled by the two general evaluation approaches presented above. Variable-oriented studies, such as those conducted by Flyvbjerg and colleagues (2003a), examine relationships between general features of infrastructure projects. These features are conceived as variables (e.g. project type, topography and cost overrun) and are then correlated with each other. This makes it is possible to deduce empirical generalizations about structural processes relevant to a larger number of cases (Ragin, 1987). For instance, cost underestimation and
overruns occur more often in rail projects and, within rail projects, overruns are higher in developing nations than in North America and Europe (Flyvbjerg, 2007a).

However, ‘the simplifying assumptions that make this approach possible often violate commonsense notions of causation and sometimes pose serious obstacles to making interpretative statements about specific cases or even about categories of cases’ (Ragin, 1987: xiii). In essence, Flyvbjerg’s studies cannot account for the idiosyncratic nature of specific cases as the study of Anguera (2006) can, because variable-oriented studies are by their nature not case-oriented. Furthermore, correlational methods cannot deal with contextuality; they do not allow for complex causality (see next section). For example, in the work of Flyvbjerg and colleagues, the importance of the context is recognized by pointing to the fact that cost overruns are due to different circumstances. However, this is not reflected in the generic patterns that appear in the research. Finally, variable-oriented studies can account for time. For instance, Flyvbjerg et al. (2003b) conclude that infrastructure development performance has not improved in the past 70 years. However, such studies have a hard time including complex dynamics, such as the influence of particular events (e.g. the crash at Lockerbie) on the course of the development of a specific case.

Case-based methods are by their nature sensitive to the complexity, diversity and (historical) uniqueness of cases (Ragin, 1987). Projects are treated holistically and not as collections of parts; case studies are sensitive to contextuality and temporality. For instance, Anguera (2006) is able to discuss in detail the Channel Tunnel project, its performance, and the key factors that influenced the latter. However, in the words of Aus (2009: 175), ‘most case studies (…) could maliciously be qualified as theoretical ‘data dumps’. One of the methodological reasons for this rather unfortunate state-of-the-art is that single case studies can hardly be employed for explanatory purposes.’ Hence, the methodology needs to be case-comparative to allow for causal inference (Aus, 2009) – for studying patterns across cases.

However, when case study material is analyzed and compared in case-comparative studies, this often happens rather loosely and in a non-formalized manner (Rihoux, 2006). Often, qualitative comparative studies, necessarily limited in the number of cases, result in an overview of the most important similarities and differences (Rihoux and De Meur, 2009). This comparative process is often not formalized in the sense that little insight is and can be provided into the way it was performed, including identifying the decisions that influenced the outcomes of the comparative process. In addition, the rich data represents many possible causal conditions that are often hard to disentangle. Consequently, the scientific value of these studies is often questioned (Ragin, 1987). Table 2 summarizes the strengths and weaknesses of both approaches and compares them with QCA as a hybrid alternative research approach that integrates the strengths of both approaches, thereby mitigating their weaknesses.
Qualitative comparative analysis

A hybrid alternative

Qualitative comparative analysis (QCA) is an umbrella term for three comparative methods: crisp set QCA (csQCA), multi-value QCA (mvQCA) and fuzzy set QCA (fsQCA) (Rihoux and Ragin, 2009). It aims to integrate the case-oriented and variable-oriented approaches so that scientists do not have to choose between an ‘understanding of complexity and knowledge of generality’ (Ragin et al., 2003: 324). QCA can be used to ‘achieve a systematic comparison across a smaller number of individual cases (e.g. a sample of between 10 and 30 cases) in order to preserve complexity, and yet being as parsimonious as possible and illuminating otherwise often hidden causal paths on a micro level’ (Rihoux and Lobe, 2009: 228).

It is a comparative case-based approach that allows the examination of multiple causal configurations (Byrne, 2009). This approach aims to uncover the most frequent combinations of causal conditions (i.e. variables) that produce a certain outcome. For instance, Anguera (2006) presents several factors that affected cost overruns in the Channel Tunnel project. However, different configurations may produce the outcome. For instance, the Øresund link and the Channel Tunnel have similar outcomes, but their configurational paths towards that outcome are different. In addition, factors can have different effects in different contexts. For instance, the advent of low-cost airlines had different effects on the Channel Tunnel and the Øresund link because of their different contexts and natures. Grofman and Schneider (2009) and Schneider and Wagemann (2010) refer to these characteristics of complex causality as conjunctural causation, equiparnility and multifinality respectively. They add to this the notion of asymmetric causality, i.e. that the presence and absence of outcomes require different explanations. In sum, the approach ticks all of the boxes in Table 2, excluding, as discussed later, the last one concerning temporality.

<table>
<thead>
<tr>
<th>Approach</th>
<th>Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable-oriented</td>
<td>In-depth vs. generalization</td>
</tr>
<tr>
<td></td>
<td>Case-based vs. comparative</td>
</tr>
<tr>
<td></td>
<td>Attention to context</td>
</tr>
<tr>
<td></td>
<td>Attention to time</td>
</tr>
<tr>
<td>Case-oriented</td>
<td>Generic patterns</td>
</tr>
<tr>
<td></td>
<td>Comparative, not case-based</td>
</tr>
<tr>
<td></td>
<td>Limited</td>
</tr>
<tr>
<td></td>
<td>Limited</td>
</tr>
<tr>
<td>Qualitative</td>
<td>In-depth, focus on idiosyncrasies</td>
</tr>
<tr>
<td>Comparative</td>
<td>Not comparative, case-based</td>
</tr>
<tr>
<td>Analysis</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Qualitative</td>
<td>Iterations between generic patterns and idiosyncrasies</td>
</tr>
<tr>
<td>Comparative Analysis</td>
<td>Case-based comparison as main feature</td>
</tr>
<tr>
<td></td>
<td>Yes, strong</td>
</tr>
<tr>
<td></td>
<td>Yes, but weak</td>
</tr>
</tbody>
</table>

Table 2: variable-oriented versus case-oriented approach
QCA as a complexity-informed research approach

It is important to clarify that QCA is not just a method; it is first of all a research approach (Rihoux, 2003). Central to this approach is the dialogue between theoretical ideas and empirical evidence (Berg-Schlosser et al., 2009; Ragin, 1987; 2000; 2008a; Yamasaki and Rihoux, 2009), which is especially important in the selection and construction of cases and variables. In QCA, variables are conceptualized as causal conditions or sets. For instance, the Channel Tunnel, Øresund link and Great Belt link cases are members of the ‘transportation infrastructure projects’ set.

What makes set theory interesting for case comparative studies is that sets can be intersected (i.e. the set operator ‘logical and’ – referring to conjunctural causation) and unified (i.e. the set operator ‘logical or’ – referring to equifinality and multifinality). QCA is able to systematically compare and analyze these set conjunctions (i.e. configurations or causal recipes) (Ragin, 1987; 2000; 2008a; Smithson and Verkuilen, 2006).

However, social phenomena such as infrastructure development are often difficult to grasp in terms of sets. For instance, compare the relative ease of defining the set boundaries of concepts such as ‘democracy’ and ‘legitimacy’ with ‘transport modes’. Producing a list of projects that fit a certain type of transport mode is probably easier than listing the legitimate public participation processes involved in infrastructure development projects (cf. Smithson and Verkuilen, 2006). Therefore, theoretical and substantive knowledge should be used to substantiate the construction (and membership) of sets.

Illustrating QCA for the evaluation of situated complexity

Using set theory implies a focus on set relations instead of correlations. Instead of studying the net-additive effects of variables, QCA studies the necessity and sufficiency of relations between (configurations of) sets and the outcome (Ragin, 1987; 1999; 2000; 2008a; Schneider and Wagemann, 2010). A condition is necessary if it has to be present for the outcome to occur, indicated by the outcome being a subset of the causal condition. Suppose condition B is the set ‘cost overrun’ and condition A is the set ‘construction delay’. Then, Figure 1 shows that every case that exhibits cost overrun also exhibits construction delay. If a case does not exhibit A, then it cannot be in set B. This would point to construction delay is a necessary condition for the outcome to occur.
A condition is *sufficient* if it can produce the outcome by itself, indicated by the condition being a subset of the outcome. Now suppose that B is the set ‘construction delay’ and A is ‘cost overrun’. Then, Figure 1 shows that every case that exhibits construction delay also exhibits cost overrun. If construction delay is the only condition at play, then this figure indicates that it alone is sufficient to produce cost overrun.

However, necessary and sufficient conditions can and most often are combined within a causal recipe since causation is complex (Ragin 2000). This means that there are usually no purely necessary or sufficient conditions for an outcome to occur (i.e. overlapping sets). Such conditions are called INUS conditions, which can be defined as an ‘*insufficient but non-redundant* part of an *unnecessary but sufficient* condition’ (Mackie, 1980: 62, original italics). For example, imagine that the Channel Tunnel, Great Belt link and Øresund link have different ‘scores’ on three causal conditions A, B and C. Suppose that the Channel Tunnel exhibits conditions A and B, the Great Belt link conditions B and C and the Øresund link conditions A and C. Thus, there are three configurations: A*B, B*C and A*C. The * sign indicates ‘logical and’. These three different paths produce the outcome cost overrun. This implies ‘logical or’, indicated by a + sign. This can be written as a Boolean expression, namely:

\[ A*B + B*C + A*C \rightarrow \text{cost overrun} \]

This means first that none of the three conditions is individually sufficient since in none of the three cases the outcome is produced by a single condition. Second, it means that none of the three conditions is necessary. For instance, condition A is not necessary for cost overrun to appear, since the outcome can also appear with the combination B*C. It does mean, however, that A is an INUS condition: it is an insufficient (i.e. it cannot produce the outcome by itself) but non-redundant (i.e. it is a necessary condition in both the combinations A*B and A*C) part of an unnecessary (i.e. A*B and A*C are not necessary since cost overrun also appears in B*C) but sufficient (i.e. A*B and A*C are sufficient for cost overrun to occur) condition. Finally, it is important to note that ‘neither necessity nor sufficiency exists independently of theories that propose causes’ (Ragin, 2008b: 42), because this distinction is only

![Figure 1: set relations](image-url)
meaningful in the context of theoretical perspectives. In other words, the contextuality of the research method is confined to the included sets, whose construction is substantiated with theoretical and substantive knowledge.

When each case is assigned a set membership score for each of the conditions, the researcher can move from mere theoretical (i.e. set construction) and empirical (i.e. case scoring) description to comparative analysis. The first step is the construction of the truth table. The fundamental unit of analysis is the truth table row (Ragin, 1999). A truth table lists all of the logically possible configurations – expressed by the exponential formula $2^k$ (k being the number of conditions) since a condition can be both present and absent – and sorts (i.e. assess the empirical presence of) the cases over to these configurations in the ‘distribution of cases’ column. Next, for each causal recipe the outcome value is defined. For illustrative purposes, a hypothetical truth table with three conditions is depicted below.

<table>
<thead>
<tr>
<th>Condition A</th>
<th>Condition B</th>
<th>Condition C</th>
<th>Outcome</th>
<th>Distribution of cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

*Table 3: hypothetical truth table*

Note that the process described so far has taken the researcher from qualitative thick-case descriptions of cases – which are preferably based upon multiple perspectives by (multiple) researchers – via the construction of sets and the assignment of set membership scores towards the construction of the truth table. One might argue that this discounts the unique and complex nature of cases. However, the method actually facilitates the interpretive and iterative research process by formalizing the comparative procedures from an in-depth understanding of complex systems to the identification of recurring patterns over multiple systems, without undermining the essence of complex systems: their contextuality and complex causation.

Next, the truth table can be minimized to produce a so-called solution (i.e. a statement about patterns across the cases). Maximum complexity was assumed a priori, and this complexity is now brought back to its core. This minimization process is structured by using Boolean algebra. Its basic procedure can be summarized as follows: ‘if two Boolean expressions differ in only one causal condition yet produce the same outcome, then the causal condition that distinguishes the two expressions can be considered irrelevant and can be removed to create a simpler, combined expression’ (Ragin, 1987: 93). For example, the previous hypothetical truth table gives the following Boolean expressions:
\[ A*B*C + A*B*~C + A*~B*~C + A*~B*~C + ~A*B*C \rightarrow \text{cost overrun} \]

The tilde sign (\(\sim\)) indicates the absence of a condition. This formula shows five empirically observed paths towards cost overrun. As a next step it can be evaluated what the sufficient, necessary and/or INUS conditions are in this expression.

The first two configurations can be minimized to \(A*B\): whether condition \(C\) is present or not, cost overrun appears nonetheless. This pairwise minimization procedure is further displayed in Figure 2 and results in the following solution formula:

\[ A + B*C \rightarrow \text{cost overrun} \]

It is pivotal that the resulting formula is not applied mechanically in concluding the comparative analysis. As stated above, QCA is first of all a case-based qualitative approach. It involves the researcher interpreting the formula and critically assessing it in light of the individual cases: does it make sense? As a consequence, the comparative research process should consist of several iterations between data and concepts, generating increased understanding/interpretation of the cases. Indeed, the researcher will often be forced to do so since contradictory configurations (i.e. a configuration that produces opposing outcomes across cases) require adapting the selection and conceptualization of conditions. As such, QCA allows a continuous search into the conditions under which infrastructure projects are being realized, and enables the researcher to find causal patterns that are deeply hidden in the appearances of the projects.

**Discussion and conclusions**

Rooted in the logic of situated complexity, we made the case for the use of QCA to evaluate infrastructure projects, thereby contributing to the discussion of complexity
in evaluation (cf. Callaghan, 2008) and of QCA in evaluation (cf. Befani et al., 2007; Marx, 2005). We provided a brief overview of the logic and main procedures of QCA. Inevitably, some in-depth aspects were left out, such as limited diversity / logical remainders, contradictory rows, counterfactual analysis and the consistency and coverage measures; unfortunately, it is not possible to discuss these due to the limited space available. However, it is necessary here to elaborate on a further development of QCA, namely fuzzy set QCA or fsQCA.

fsQCA was first introduced in Ragin (2000). Although QCA and fsQCA are equal in their set-theoretic and configurational rationale, a dichotomous distinction is made in QCA (Ragin, 1987) between the absence and presence of causal conditions in a case (i.e. crisp sets). fsQCA, on the other hand, allows for finer gradients in the degree of set membership (i.e. a variable does not need to be fully present or absent in a case) (cf. Ragin, 2000; 2008a). For instance, with crisp set QCA, an infrastructure project can be fully in or fully out of the set ‘cost overrun’ (i.e. it can score 0 or 1). However, one might argue that a cost overrun of 70 percent in the Øresund link project is substantially different from a cost overrun of 110 percent in the Great Belt project (see Flyvbjerg et al., 2003a). Therefore, the latter may have a stronger presence in the ‘cost overrun’ set than the former. This can be formalized by assigning different set memberships to these cases (e.g. 1.0 for the Great Belt project and 0.75 for the Øresund project). In this manner, fsQCA can be said to more accurately describe the complexity of infrastructure projects compared to QCA based on crisp sets.

We formulated four requirements for an evaluation approach that would match the complexity of infrastructure projects. While fsQCA fulfills three of the four requisites, it does less well on the fourth dimension (time). In essence, QCA is a static method (Rihoux, 2003) and does not fully capture the dynamics of complex systems (Gerrits, 2011). Some provisional workarounds suggested by Rihoux (2003), De Meur et al. (2009) and others capture the time dimension by: using multiple iterations of the method (i.e. before, during and after a certain intervention), interpreting the time dimension, conceptualizing time as (part of) a set, or complementing QCA with other methods. Attempts are also being made to develop a distinct time-inclusive type of QCA (Caren and Panofsky, 2005).

However, all options are compromises and researchers should be aware that while QCA is useful for mapping the systemic complexity of infrastructure development, it is less useful for incorporating time dynamics. One could argue that this is not really a problem in the case of ex-post evaluation, which measures the static outcome of a project, but may be more of a concern in ex-durante evaluations, where time is an important part of the evaluation. However, we argue that one needs the time-dimension to understand the emergence of outcomes, i.e. whether the observations at t_{n+1} are the result of the configurational changes that took place at t_{n-1}. Time matters in evaluation and is an unresolved issue with the QCA approach.

Another inherent limit of (fs)QCA is the number of conditions that can be taken into account, since the logically possible number of combinations increases exponentially. At the same time, the addition of a new case to the set of cases being
compared can lead to a different solution formula. Although both issues are not unique to QCA – one could argue that such limitations are an innate part of research into social reality – they should not be ignored. Regarding the limited number of conditions being considered, Rihoux (2003) and De Meur et al. (2009) suggest that a possible remedy is to carry out multiple routines, i.e. building an increasingly clearer set of conditions by going through the QCA process multiple times. This way, the researcher is able to find out, in a very transparent way (Rihoux et al., 2009), which conditions do not matter or yield the same or similar results. These conditions can then be excluded from the analysis or grouped together as macro variables. Schneider and Wagemann (2006) provide an alternative staged approach: first, remote (e.g. contextual) and proximate factors are analyzed separately, and in the second stage, the remote and proximate factors that have been found to be influential are analyzed together.

The second issue, the possibility of arriving at different conclusions after adding new cases, is actually part of the philosophy behind QCA and its roots in systemic thinking. With QCA, the researcher does not strive to identify a single central tendency that reflects reality as more cases are added (Rihoux 2003). Rather, it helps researchers to examine the different causal pathways that lead to a particular outcome and how such pathways are linked to individual cases. Adding a new case can lead to the discovery of a new pathway. More commonly deployed comparisons aim to find the variable that controls for differences and similarities in multiple cases. Following the discussion on generic and situated complexity, such a search is beside the point. Case can have unique pathways and comparison should be used to highlight the particularities of the pathways. It is also through this perspective that thinking in terms of dependent or independent variables is replaced by thinking in configurations (Aus, 2009), which resembles social reality more closely.

In short, our argument is that (fs)QCA is a promising approach that largely meets the evaluation requirements set out at the start. Some issues will have to be dealt with, most prominently the issue of time. The next step is to use (fs)QCA in an evaluation study to evaluate its viability as a tool for analyzing complex infrastructure development projects.

Acknowledgements

We would like to thank Christopher Koliba, Asim Zia, Frank Boons, Nanny Bressers, Jasper Eshuis, Mike Duijn and Göktuğ Morçöl for their valuable feedback on earlier versions of this work. We thank the two anonymous reviewers for their enthusiastic words and final comments to finish the article.

Funding

This work was supported by Next Generation Infrastructures [grant 03.24EUR]; and the Netherlands Organisation for Scientific Research Veni-program [grant 451-10-022].
References


Cantarelli CC, Flyvbjerg B, Van Wee B and Molin EJE (2010) Lock-in and its influence on the project performance of large-scale transportation infrastructure projects: Investigating the way in which lock-in can emerge and


