Published in Policy Sciences 49, 335-348, 2016,
which should be used for any reference to this work

Identifying context and cause in small-N settings: a comparative multilevel analysis

Eva Thomann[1] · **Anita Manatschal**[2]

Abstract Qualitative small-N comparisons face the challenge to detect context-bound causality under conditions of limited empirical diversity. Rather than treating context as a causal factor, we test the usefulness of the novel method of comparative multilevel analysis (CMA) to identify and understand the role of context as a contingent necessary condition that enables a causal relationship to unfold. Combining CMA with pairwise comparisons, we assess how organ donation policies in Switzerland and Spain affect relatives' refusal rates in a small-N setting exhibiting multiple contextual levels. To tackle limited diversity systematically, we suggest to refine the CMA methodology by accounting for several contexts and referring to higher-order constructs. Applying CMA with these refinements, we find voluntary information measures only affect refusal rates in contexts of a credible state *explicitly* supporting organ donation. The fact that CMA can easily be combined with other analytical and conceptual approaches makes it an effective technique to identify contextual effects in small-N research.

Keywords Comparative multilevel analysis · Contextual effects · Small-N comparison · Limited empirical diversity · Galton's problem

Introduction

Recent years have witnessed unprecedented methodological progress in qualitative research (e.g., Blatter and Haverland 2012; Levi-Faur 2006; Ragin 2000). One challenge arising from nonlinear, complex causality and interdependencies between cases ("Galton's

✉ Eva Thomann
eva.thomann@mzes.uni-mannheim.de

Anita Manatschal
anita.manatschal@ipw.unibe.ch

[1] Mannheim Centre for European Social Research, University of Mannheim, 68131 Mannheim, Germany

[2] Institute of Political Science, University of Bern, Fabrikstrasse 8, 3012 Bern, Switzerland

problem") is the need to be "attentive to the interaction between causal mechanisms and the context in which they operate" (Falletti and Lynch 2009, p. 1144). Policy sciences and realist evaluation approaches acknowledge that contextual embedding is required to understand how governance concepts turn into effective policy configurations (Pawson 2002; Voß et al. 2009). The assumption of context-bound causality implies that context sometimes affects only the operation of a causal mechanism, without being a cause for the outcome itself (Blatter and Haverland 2012; Goertz 1994). A novel analytical framework adopting this perspective is comparative multilevel analysis (CMA) for comparing subsystems from different contexts in small-N research (Denk 2010).

This paper uses CMA to identify the role of different policy instruments for relatives' low refusal rates to organ donation. A public policy instrument is a set of techniques by which public actors "wield their power in attempting to ensure support and effect or prevent social change" (Vedung 1998, p. 21). Existing studies suggest that a plethora of policy instruments matter for relatives' refusal rates (Manatschal and Thomann 2011), depending on the analytical level or context they focus on. The organ donation process constitutes a prototypical example of how causation is characterized by "many variables, few cases" and context-bound (Byrne 2009). Similar perceptions of the organ donation issue may diffuse within cultural and political contexts, which alter how instruments affect people's decisions (Braun and Gilardi 2006).

Contrary to the basic setting imagined by Denk (2010), our contextual setting is itself multilevelled. We compare small hospitals and large hospitals in two contrasting national contexts: Spain as the international example of best practice, with very low refusal rates, and Switzerland, which exhibits comparatively high refusal rates despite having adopted elements of the Spanish model, with a pronounced regional variation between German-speaking and French- and Italian-speaking cantons.

Based on the premise that CMA assesses a specific effect of context as changing meaning (Goertz 1994), our application suggests that CMA helps researchers to identify whether causal relationships in genuine small-N settings are linear, i.e., hold across contexts, or nonlinear, i.e., contingent upon the context which implies interdependencies between cases (Goertz and Starr 2003; Ross and Homer 1976). Simultaneously, our analysis identifies clear limitations of CMA in handling equifinality and limited diversity at the context level. One of CMA's main advantages is that it can easily be complemented with other techniques and methods (Denk and Lehtinen 2014). This allows us to propose practical refinements of the CMA-approach. These involve accounting for several contexts and reducing the number of variables by resorting to higher-level theoretical constructs. Applying CMA with these refinements, we find that voluntary information measures only unfold an effect on refusal rates in the context of a credible state *explicitly* supporting organ donation.

We now discuss relevant aspects of causal complexity and explain how the CMA methodology can address them. Section three presents the outcome, case selection, explanatory and contextual factors and our analytical strategy. Section four continues with the empirical analysis. We discuss the major findings in section five, while chapter 6 wraps up our concluding remarks.

Complex causation, contextual effects and comparative multilevel analysis

Causal complexity and Galton's problem

Case-based research typically asserts that, while causality is real, it is also complex. Complexity arises from the fact that systems are made up of many layered components that

interact in non-simple ways (Byrne 2009; Glennan 2010). Complexity implies that causal relationships as middle-range theories may vary over time and space (Goertz 1994; Pawson 2002). Cases may not constitute causally independent observations: temporal or spatial interdependencies produce diffusion processes, thus creating a specific context for functional intra-system explanations. The latter might not appear linear when cases are compared across, e.g., temporal, cultural, regional or national contexts in which these diffusion mechanisms are absent ("Galton's problem"; Braun and Gilardi 2006). As a consequence, scholars often dismiss functional in favor of contextual explanations (Ross and Homer 1976). However, these interdependencies may also indicate the nonlinearity of some political interactions: the same outcome might be generated by different causal factors, depending on the context (*equifinality*; Ragin 2000). In this vein, critical realism, realist evaluation and the policy sciences approach assert that causality is inherently *context-bound* (Byrne 2009; Voß et al. 2009). Context-bound causality implies to distinguish generative mechanisms of a policy and its contiguous context, where the latter triggers the former (Pawson 2002). A comprehensive solution to Galton's problem accounts for both diffusion and function, by comparing cases within and across contexts and including the relevant external variables (Ross and Homer 1976). CMA enables researchers to do both, by addressing simultaneously the possibility of context-bound causality (comparing across contexts) and the "too many variables, too few cases" issue (*limited empirical diversity*) which often arises in small-N research when more variables are added to an explanation.

Comparative multilevel analysis as an approach to context-bound causality

Comparative multilevel analysis (Denk 2010) proposes a fairly simple set of four completions to conventional qualitative comparative methodology which facilitate the analysis of context-bound causality in small-N research.

The first step of a CMA study consists in grouping cases in relation to their similarities at the system level, thereby creating different subsystems within a multilevel structure. Second, intra-system analysis compares cases *within* each group. This produces as many comparative expressions of a causal relationship between explanatory factors and outcomes as there are groups. Third, intersystem analyses compare the comparative expressions of the groups themselves *between* groups. Fourth and finally, expressions for those grouped comparisons are formulated. These expressions describe whatever differences (D) or similarities (S) exist between the grouped cases regarding the relationship between explanatory factors and the outcome, and the context (Denk 2010, p. 33). Differences in these expressions indicate that the context impacts on the relationship between the explanatory factors and the outcome. Similarities signify that the context does not matter.

By accounting for contextual factors and their (ir)relevance for causal relationships, CMA enables researchers to identify whether causal relations hold across contexts (linear) or depend upon the context that implies interdependencies between cases (nonlinear). Given that Denk (2010) does not discuss these aspects systematically, we now specify how CMA can address causal complexity in terms of context-bound causality and limited diversity.

First, Denk (2010) only rudimentarily specifies the type of contextual effect that CMA assumes. We suggest that CMA does not assess context as a *cause of* or a *barrier to* an effect, but as a part of the environment *changing meaning*, affecting how cause and effect interact (Goertz 1994). It refers to context as "the relevant aspects of a setting (analytical, temporal, spatial, or institutional) in which a set of initial conditions leads (…) to an outcome (…) that is, those [aspects] that allow the mechanism to produce the outcome"

(Falletti and Lynch 2009, p. 1152). Context C is a *contingent necessary condition* for a relationship between a cause X and an effect Y to unfold (Goertz and Starr 2003, p. 72): only if C, then X → Y (Denk and Lehtinen 2014, p. 3477). Context as changing meaning can be any social, historical, cultural, political or economic factor that influences the relationship between two (or more) variables (Falletti and Lynch 2009, p. 1159; Glennan 2010; Goertz 1994, pp. 14, 26, 28). In this study, policy instruments form the relevant contextual and causal conditions.

Second, limited diversity makes it difficult to identify the (ir)relevance of the numerous variables for the observed differences. The number of variables can be reduced empirically or theoretically (Peters 1998, p. 70ff). The empirical strategy of intra-system comparison is, however, limited, as it cannot detect whether certain explanatory factors emerge as relevant for an outcome only in one context. By including multiple contexts in the analysis, CMA surpasses this limitation (Denk 2010). As another strategy to refine CMA, we apply a more parsimonious theory by subsuming policy instruments using higher-order theoretical constructs (Cioffi-Revilla and Starr 2003; Goertz 2006).

Materials and methods

Case selection and data

We use CMA to discover which policy instruments impact on the number of refusals by deceased patients' relatives as a share of total requests for organ donation in divergent contexts. Swiss refusal rates exceeded Spanish rates by 2.5 times in 2009 (Council of Europe 2011). Table 1 also reveals striking differences in refusal rates among the German-speaking and French- and Italian-speaking (=Latin) linguistic regions within Switzerland. Finally, refusal rates are higher in large than in small hospitals in Switzerland, but not in Spain. The contexts shared by geographically proximate hospitals may be linked to the diffusion of similar perceptions of organ donation issues (Braun and Gilardi 2006). Applying Galton's problem here means that this can foster similar ways in which instruments affect people's response to the organ donation request. Our research design must include the relevant contexts that could induce such processes (Goertz 1994).

We compare large with small hospitals as the *smallest units of analysis between which both refusal rates and policy instruments vary systematically*. Lower levels of aggregation would not present new insights. Big hospitals have a division of neurosurgery, whereas small hospitals have none (Council of Europe 2011). We thus compare six cases in five different contexts, three regional and two national—the relevant contextual levels at which the organ donation system is organized. Contrary to Switzerland, there is no regional variation in policy instruments in Spain. Spain's Model of organ donation was established before the decentralization of its health care system. It was implemented in a highly uniform manner in all Spanish regions (Manatschal and Thomann 2011). This is reflected in the low variance of refusal rates across the 17 Spanish regions ranging from 7.7 to 30 % in 2008 (ONT 2008). We expect hospitals embedded in the same regional and/or national context to be interdependent. Conversely, apart from comparable organizational backgrounds at the hospital, regional and national level resulting from the worldwide diffusion of elements of the Spanish Model (Manatschal and Thomann 2011, pp. 40, 43), Swiss and Spanish hospitals should be more independent from each other.

Our analysis bases on data collected through literature and document analysis and 28 expert interviews in both countries (Manatschal and Thomann 2011). The Spanish sample

Table 1 Research setting, policy instruments and refusal rates

Level of analysis	Policy instrument	CH Context 4				ESP Context 5	
System II (national)	Active public awareness raising	ar	ar	ar	ar	AR	AR
	Quality monitoring	qm	qm	qm	qm	QM	QM
	Cost coverage	cc	cc	cc	cc	CC	CC
	Presumed consent	pc	pc	pc	pc	PC	PC
		CH-GE Context 1		CH-L Context 2		ESP Context 3	
System I (CH & ESP regions)	Comprehensive education of hospital staff	edu	edu	EDU	EDU	EDU	EDU
		Large hospitals	Small hospitals	Large hospitals	Small hospitals	Large hospitals	Small hospitals
Subsystem (cases)	Active family approach	fam	fam	fam	fam	FAM	FAM
	Donor coordinator reimbursement	RC	rc	RC	RC	RC	RC
	Donor transfer for retrieval	t	T	t	T	t	t
	Refusal rate	49.0%	78.0%	28.4%	75.0%	19.8%	16.2%

Capital letters = policy instrument is applied, lower case letters = policy instrument is not applied

CH: Switzerland, CH-GE: German part of Switzerland, CH-L: Latin (i.e., French and Italian) part of Switzerland, ESP: Spain. Refusal rate = number of refusals by next of kin as share of total requests (in %).

Data sources: Swiss Donor Action (mean of 2007, 2008, 2009), Organicazión Nacional de Trasplantes (ONT 2009)

comprises all hospitals in 2009 (71 big, 68 small hospitals), whereas data are only available for 50 % of the Swiss hospitals,[1] which are, however, considered representative (Manatschal and Thomann 2011, pp. 33, 64ff).

Policy instruments and their classification

Our analysis focuses on policy instruments, which are crucial for refusal rates. Policy instruments in morality policies such as organ donation reflect the relationship between the governing and the governed, the broader political and cultural context, and a (de)validation of particular sets of norms (Engeli and Varone 2011; Howlett 2009; Lascoumes and Le Galès 2007). We hence classify our policy instruments according to different degrees of state authority exercised (*coerciveness*), using Vedung's (1998) tripartite classification into sticks (regulations; most coercive), carrots (economic or non-monetary [dis-]incentives)

[1] French part: 2009, 2008: 4 big and 12 small donor action participant hospitals; 2007: 3/13. German part: 2009: 4 big and 11 small participant hospitals; 2008: 4/15, 2007: 3/16.

and sermons (information; least coercive) (see case studies in Manatschal and Thomann 2011). Table 1 summarizes the research setting and the policy instruments prevailing at different analytical levels. Capital letters indicate an instrument's systematic and encompassing application; otherwise, lower case letters are used.

Policy instruments may figure as causally productive or relevant properties, or as contextual background conditions (Glennan 2010) that act as a "part of the environment" which allows the causal mechanism to produce a certain outcome (Denk and Lehtinen 2014, p. 3476; Falletti and Lynch 2009, pp. 1153–1154). We argue below that only those factors causally affect refusal rates that directly influence either the motivation of staff or the decision of individual relatives in the situation of the donor request. Conversely, other policy instruments may matter as a context for how these causal factors work, if they are neither direct properties of the cause nor of the effect (Glennan 2010). This distinction enables us to address both the question of context-bound causality and limited diversity, by reducing the number of possible causal factors. Factors that do not vary at least one level of analysis are not discussed.

Causal factors at the hospital level

As an important *sermon*, a personal and temporal decoupling of the request for organ donation (*family approach*) from the notification of death decreases the likeliness of refusal. Repeating the request several times ("reapproach") can lead initially undecided relatives to reconsider their decision. The adequate *reimbursement of donor coordinators* is a crucial economic incentive (*carrot*) for organ procurement. Economic resources available determine the efforts and expertise at disposal for family requests. *Donor transfer* represents a negative incentive for relatives: Many refusals are due to concerns about what will happen with the deceased's body.

Regional context

Educational programs (*sermon*) for intensive care nurses, doctors and donor coordinators coincide with low refusal rates. They provide the hospital staff involved in the core organ recruitment processes with the skills for optimal care and communication needed for the family approach to positively affect relatives' decision.

National context

A first national *sermon* regards *public awareness raising* (information and education) aimed at influencing the population's knowledge of and attitude toward organ donation. While awareness raising does not directly influence donor rates, policy instruments are more likely to work as intended in a context of removal of taboos from, and sensitization toward the organ donation topic. *Quality monitoring programs* provide hospitals with information to adapt and improve donation processes.

Adequate *cost coverage (carrot)* refers to hospital funding of organ donation processes preceding organ retrieval. While this does not affect the relatives' decision, it is an important aspect of the broader setting, which, when absent or incomplete, represents a disincentive for the commitment of hospital staff.

The *legal model of consent* obliges relatives to take a decision (*stick*). Under presumed consent as practiced in Spain, relatives must express their opposition to donation. By contrast, under the explicit/informed consent model used in Switzerland, relatives must explicitly express their agreement with donation. There is no unambiguous evidence that presumed consent affects refusal rates, but it influences how the precise request is formulated.

Classification

The national context itself is characterized by numerous contextual variables. We aim at a more precise statement about what the relevant characteristic of a given context is. Is it a subset of these contextual conditions, or rather a first-order characteristic (Cioffi-Revilla and Starr 2003), that makes the difference? CMA itself offers no tool to deal with limited diversity at the contextual level. Reference to higher-order constructs offers a promising strategy to tackle this issue. The numerous contextual conditions can be conceived of as exchangeable indicators or secondary-level dimensions of an overarching, basic-level concept (Goertz 2006). This multilevel conception of policy instruments allows us to identify the relevance of different contextual factors, or their underlying characteristics, as contingent necessary conditions.

We identify three conceptual levels for our policy instruments (Table 2): At the indicator level, we have differing policy instruments. At a secondary level, these policy instruments can be classified as instrument types depending on their *coerciveness* (sermons, carrots or sticks). At the basic level, we identify the criterion *explicitness* as a higher-order theoretical construct that expresses whether governments take an explicit position regarding the values underlying a morality policy (Engeli and Varone 2011). For example, a sermon is more explicit if it represents state action aimed at influencing the decision with regard to organ donation, compared to a sermon representing a neutral or absent official position in this matter. The logical operator "OR" attributes the single instruments as substitutable second-order factors to types and varieties of policy instruments. For instance, either edu or QM is empirical manifestations of a sermon. In turn, either PC (stick) or AR (sermon) indicates the explicitness of the policy instruments in question.

Analytical strategy

CMA requires researchers to choose a method for analyzing causal relationships between cases inside contexts (Denk 2010). Based on our case study insights, we apply Levi-Faur's

Table 2 Types and varieties of policy instruments

Basic level	Extent to which state influences policy goal (varieties)	Explicit			Non-explicit		
Secondary level	Coerciveness (types)	Sermons	Sticks	Carrots	Sermons	Sticks	Carrots
Indicator level	Single policy instruments	EDU QM AR FAM	PC	RC t CC	edu qm ar fam	pc	T rc cc

Own illustration based on Vedung (1998) and Goertz (2006)

(2006) techniques of paired comparison, as the number of cases is too low for alternative approaches such as qualitative comparative analysis (QCA; Ragin 2000).[2]

While using a covariational logic, pairwise comparisons integrate a high degree of detail to ensure that the causal connections drawn are real (Peters 1998, pp. 66–67; Tarrow 2010). Table 3 summarizes the inferential strategies (methods of agreement or difference; Mill 1843) applied to the most similar system analyses at the subsystem level, which help control for other explanatory factors (Levi-Faur 2006).

As our analysis starts with paired comparisons at the hospital level, we code high (R) or low (r) refusal rates comparing large with small hospitals individually for each region. Although dichotomization implies a loss of information (Goertz 2006), it captures the essential differences between small and big hospitals, is a widespread practice for pairwise comparisons (Tarrow 2010), and seems appropriate for the use of CMA. The outcomes we seek to explain are low refusal rates (r) in each region.

Furthermore, Table 3 clarifies the research questions that emerge at each analytical level: following the bottom–up, stepwise procedure of CMA, we first seek to explain low refusal rates in interdependent hospitals at the subsystem level. At the system I (regions) and II (nations) levels, we account for possible nonlinearity by asking whether context

Table 3 Analytical questions and strategy

3. How to address limited diversity at the context level?						
Comparing all formulizations 1) using single policy instruments, and 2) higher order constructs						
2. Does context matter?						
Analytical step	**INTER4**					
System II (nations)	Context 4				Context 5	
Analytical step					**INTER2 & 3**	
	INTER1					
System I (regions)	Context 1		Context 2		Context 3	
1. What explains the Outcome (r)?						
Analytical step	**INTRA1**		**INTRA2**		**INTRA3**	
Inferential strategy	MSSD + MMD		MSSD + MMD		MSSD + MMA	
Subsystem (cases)	LH	SH	LH	SH	LH	SH
Outcome	r	R	r	R	r	r

R high refusal rate, *r* low refusal rate, *LH* large hospitals, *SH* small hospitals, *MSSD* most similar systems design, *MMD* Mill's method of difference, *MMA* Mill's method of agreement

[2] Rohlfing (2012) highlights that CMA becomes superfluous when combined with QCA.

matters for differences in the ways in which policy instruments affect relatives' responses between independent cases. Lastly we tackle the question, how to address the limited diversity that followed from accounting for numerous possibly relevant contextual factors, by referring to the overarching first-order characteristic of explicitness.

Results

Intra-system analyses

A first intra-system analysis compares large and small hospitals in German-speaking Switzerland. Applying the most similar systems strategy and Mill's method of difference, we can infer from Table 1 that no active family approach (fam), which prevails both with low and high refusal rates, is no critical variable (Levi-Faur 2006, p. 58). Instead, two variables account for the comparatively lower refusal rates in large German-speaking hospitals: donors are not transferred to a different retrieval hospital (t), and the donor coordinators are adequately reimbursed for their activities (RC). We can formulate a first comparative expression:

$$\textbf{CH-GE:} \quad \mathrm{RC},\ \mathrm{t} \rightarrow \mathrm{r} \qquad \text{(expression1)}$$

In the Latin part of Switzerland, only the lack of donor transfer (t) accounts for lower refusal rates in large hospitals—neither the family approach nor the fact that donor coordinators are reimbursed can explain the different refusal rates. The second comparative expression reads as follows:

$$\textbf{CH-L:} \quad \mathrm{t} \rightarrow \mathrm{r} \qquad \text{(expression2)}$$

In our third intra-system analysis, small and large hospitals in Spain are compared using a most similar systems strategy and Mill's method of agreement (Levi-Faur 2006). Yet, as the policy instruments do not vary, none of them can be singled out as irrelevant. Hence, low refusal rates in this context can be explained by an adequate coordinator reimbursement (RC), an active family approach (FAM), and no donor transfer (t):

$$\textbf{ESP:} \quad \mathrm{FAM}, \mathrm{RC}, \mathrm{t} \rightarrow \mathrm{r} \qquad \text{(expression3)}$$

Intersystem analyses

To assess whether context matters, we formulate expressions for the grouped comparisons using square brackets (Denk 2010, p. 33f). The regional contextual factor "education of hospital staff" (systemic level I) precedes the causal relationship between explanatory factors and outcome represented by the comparative expression inside the brackets:

$$\begin{array}{lll} \textbf{CH-GE:} & \mathrm{edu} & [\mathrm{RC}, \mathrm{t} \rightarrow \mathrm{r}] \\ \textbf{CH-L:} & \mathrm{EDU} & [\mathrm{t} \rightarrow \mathrm{r}] \\ \textbf{ESP regions:} & \mathrm{EDU} & [\mathrm{FAM}, \mathrm{RC}, \mathrm{t} \rightarrow \mathrm{r}] \end{array} \qquad \text{(formalization1)}$$

Based on formalization 1, we perform a first regional intersystem analysis between the two Swiss regions. We find that context matters, as in different contexts (D) different conditions (D) have the same effect (S), indicating equifinality:

Intersystem analysis 1: DDS = DD.

The combination DD denotes the result of this comparison. The first letter of the combination expresses differences or similarities in context; the second indicates differences or similarities in the causal expressions. Since we only explain low refusal rates (r) the difference derives from the cause, not from the effect.

As Denk and Lehtinen (2014, p. 3481) point out, "the context may be a conditional condition for relationships between some conditions and the outcome, while other conditions have relationships with the same outcome independent of the context". While the abandonment of donor transfer (t) is relevant in both contexts, an adequate reimbursement of the donor coordinators performing the donation request (RC) only contributes to lower refusal rates when the overall sensitization of the hospital staff for organ donation is low (edu). The decisive contextual condition for the effect of RC is the less comprehensive education of hospital staff in German-speaking Switzerland (edu), suggesting that low awareness might make staff more amenable to financial incentives.

To cross-validate our findings, we compare German-speaking Switzerland with Spanish regions in a second regional intersystem analysis. Again, the summary of the two expressions suggests that context matters:

Intersystem analysis 2: DDS = DD.

Only in the context of a comprehensive staff education (EDU) in Spanish regions does the family approach (FAM) unfold its causal role. In contrast to the findings of the first intersystem analysis, both adequate reimbursement (RC) and no donor transfer (t) appear causally relevant for low refusal rates in these two contexts. The causal role of coordinator reimbursement (RC) persists independently of the comprehensiveness of staff education (EDU or edu). Hence, rather than validating the results of the first intersystem analysis, the second intersystem analysis yields a different conclusion.

This result puts into question the role of edu as a contingent necessary condition for RC to impact on refusal rates. Different layers of context can interact with each other (Falletti and Lynch 2009, p. 14): the contextual effect of education for the causal relationship between RC and r detected above may persist only in the Swiss national context.

Note that the CMA analysis would have ended without shedding light on this nonlinearity if we had not included the Spanish context for cross-validation. This experience suggests that, while CMA can detect contextual effects, issues of limited diversity may still limit the robustness of the findings. Depending on the research setting, it might therefore be advisable to compare more than just two contexts. Denk and Lehtinen (2014) do so, but they do not explain how to deal with inconclusive contextual effects.

We then conduct a third regional intersystem analysis by comparing Latin-speaking Switzerland with Spanish regions. We find that different factors yield the same outcome in a similar context—that is, equifinality within similar contexts. Denk (2010) did not foresee this scenario (Rohlfing 2012):

Intersystem analysis 3: SDS = SD.

This SD scenario is another indicator that a *different* context than the one examined might matter. Hence, we explore the possibility that the relevant contextual conditions are situated at the national level. We thus take CMA one step further by analyzing contexts at two levels. Another pair of square brackets designates the second contextual system level with the national policy factors preceding these brackets. The logical operator "OR",

designated by the Boolean + operator (cf. Goertz 2006), summarizes the situation in the Swiss context.

CH :	ar, qm, cc, pc	[edu [RC, t → r]] + [EDU [t → r]]
ESP :	AR, QM, CC, PC	[EDU [FAM, RC, t → r]]

(formalization2)

Based on formalization 2, we perform a fourth intersystem analysis to examine the role of the national context:

Intersystem analysis 4: DDS = DD.

We find that different conditions lead to a similar outcome in different national contexts, which, hence, matter. While the absence of donor transfer (t) leads to low refusal rates in both contexts, only in the Spanish context an active family approach (FAM) contributes to low refusal rates. The Spanish context comprises active public awareness raising (AR), comprehensive quality monitoring (QM), consistent cost coverage (CC), and the legal model of presumed consent (PC). Conversely, the absence of an active family approach in Switzerland does not appear causally relevant for differences in refusal rates.

Addressing limited diversity

Without knowing which (combination) of these national policy instruments is decisive, it is difficult to interpret the above contextual statement substantially. CMA itself offers no tool to identify the relative relevance of contextual conditions. We resort to higher-order constructs (Goertz 2006) to answer our last question: How to address limited diversity at the context level?

We thus rewrite formalization 2 in terms of *types* and *varieties* of policy instruments. For example, since all policy instruments representing sermons (type) are explicit (variety) in Spain and non-explicit in Switzerland, we denote them with SERM and serm, respectively.

CH :	serm, carr, sticks	[serm + SERM [CARR → r]]
ESP :	SERM, CARR, STICKS	[SERM [SERM, CARR → r]]

(formalization3)

The resulting formalization 3 is more insightful and easier to interpret than formalization 2. At the hospital level, positive incentives (CARR) contribute to low refusal rates, independent of the context (linear effect). In addition, in Spain, the fact that hospitals provide explicit, but non-compulsory information (SERM) when performing the family approach accounts for low refusal rates. The national context comprising explicit instruments acts as a contingent necessary condition for this effect, which does not exist in Switzerland.

Formalization 3 thus strongly suggests that it is not different types (sermons, sticks or carrots), but the *explicitness* of policy instruments at the national level which forms the relevant context. The types of policy instruments we observed as second-order factors are themselves substitutable instances of the underlying first-order context "explicitness" (Cioffi-Revilla and Starr 2003). In substantial terms, the comparison of Switzerland and Spain indicates, first, that incentives may lower refusal rates. Second, the context of a state explicitly and comprehensively supporting the goal of organ donation is needed for voluntary information measures to, first, be explicit and, second, unfold an effect on refusal

rates. This context affects whether people perceive the provided information as credible (Spain) or irrelevant (Switzerland). This conclusion is straightforward to grasp and may be of a high relevance for policy makers.

Discussion

Denk's (2010) CMA is a recent innovation in small-N research that helps researchers to compare cases across contexts. This study shows how CMA identifies whether causal patterns are linear, i.e., hold across contexts, or nonlinear, i.e., depend upon relevant contextual variables arising from interdependencies between cases (Ross and Homer 1976). We clarified that CMA assesses whether context changes meaning as a contingent necessary condition for a causal relationship at the subsystem level, but not whether context is itself a cause of the outcome. We used Levi-Faur's (2006) inferential strategies for pairwise comparisons to explain low refusal rates to organ donation in different contexts, and tested CMA's usefulness to discover whether context matters. Despite its potential to account for context-bound causality, we also demonstrated several limits of the CMA methodology and propose three practical refinements.

First, we encountered equifinality, which should be accounted for more systematically by CMA (refinement 1). Second, contextual effects uncovered by CMA can be inconclusive because of limited diversity. Hence, to reach more conclusive results and to enhance their robustness, we propose to validate the analysis by extending it to several contexts (refinement 2). We have shown one way in which CMA can be applied to multistage contexts. However, third, CMA reached its limits in handling a high number of contextual factors at the national level. We propose a complementary strategy (refinement 3): The number of contextual (and causal) conditions can be reduced, and their role as first- or second-order causal factors clarified, by creating higher-order constructs (Cioffi-Revilla and Starr 2003; Goertz 2006; Ragin 2000).

Thereby, we detected a more general pattern underlying the relationship between policy instruments and refusal rates to organ donation. Our findings illustrate how policy design needs to be contextually embedded to show effect. Specific incentives may lower refusal rates independently of the context. However, information measures only influence refusal rates in a context of a "credible" state that explicitly and comprehensively supports the goal of organ donation (Engeli and Varone 2011; Howlett 2009). CMA has helped us to reveal such patterns, which are difficult to anticipate by general policy models (Voß et al. 2009). To capitalize on the well-documented effect of the family approach on refusal rates, regional and national policymakers must hence consistently communicate the desirability of the goal of organ donation, rather than cherry-picking single elements of the Spanish model.

Conclusion

The application of CMA to comparisons of a small number of cases may help researchers to tackle Galton's problem and context-bound causality by systematizing the complex qualitative data in a way that uncovers contextual effects on causal relationships within subsystems. To address limited diversity in complex contexts, one of CMA's main advantages is that it can easily be complemented with other techniques (Denk and Lehtinen

2014, p. 3481). Limited empirical diversity can be addressed by increasing the number of cases, reducing the number of variables, or resorting to higher-level constructs. This paper has suggested and illustrated ways in which this can be fruitfully done to isolate complex causal and multilevel contextual patterns.

References

Blatter, J., & Haverland, M. (2012). *Designing case studies: Explanatory approaches in small-N research*. Basingstoke: Palgrave Macmillan.

Braun, D., & Gilardi, F. (2006). Taking 'Galton's problem' seriously. Towards a theory of policy diffusion. *Journal of Theoretical Politics, 18*(3), 298–322.

Byrne, D. (2009). Complex realist and configurational approaches to cases: A radical synthesis. In D. Byrne & C. C. Ragin (Eds.), *The SAGE handbook of case-based methods* (pp. 101–111). London: Sage.

Cioffi-Revilla, C., & Starr, H. (2003). Opportunity, willingness, and political uncertainty: Theoretical foundations of politics. In G. Goertz & H. Starr (Eds.), *Necessary conditions: Theory, methodology and applications* (pp. 225–248). New York: Rowman & Littlefield.

Council of Europe. (2011). *Newsletter transplant 2011. International figures on donation and transplantation—2010.* Retrieved July 20, 2014, from http://www.ont.es/publicaciones/Documents/Newsletter2011.pdf

Denk, T. (2010). Comparative multilevel analysis: Proposal for a methodology. *International Journal of Social Research Methodology, 13*(1), 29–39.

Denk, T., & Lehtinen, S. (2014). Contextual analyses with QCA-methods. *Quality & Quantity,* 48(6), 3475–3487.

Engeli, I., & Varone, F. (2011). Governing morality issues through procedural policies. *Swiss Political Science Review, 17*(3), 239–258.

Falletti, T., & Lynch, J. (2009). Context and causal mechanisms in political analysis. *Comparative Political Studies, 42*(9), 1143–1166.

Glennan, S. (2010). Mechanisms, causes, and the layered model of the world. *Philosophy and Phenomenological Research, 81*(2), 362–381.

Goertz, G. (1994). *Contexts of international politics.* Cambridge: Cambridge University Press.

Goertz, G. (2006). *Social science concepts. A user's guide*. Princeton, NJ: Princeton University Press.

Goertz, G., & Starr, H. (2003). *Necessary conditions: Theory, methodology, and applications.* New York: Rowman & Littlefield.

Howlett, M. (2009). Governance modes, policy regimes and operational plans: A multi-level nested model of policy instrument choice and policy design. *Policy Sciences, 42*(1), 73–89.

Lascoumes, P., & Le Galès, P. (2007). Introduction: Understanding public policy through its instruments—From the nature of instruments to the sociology of public policy instrumentation. *Governance, 20*(1), 1–21.

Levi-Faur, D. (2006). A question of size? A heuristics for stepwise comparative research design. In B. Rihoux & H. Grimm (Eds.), *Innovative comparative methods for policy analysis* (pp. 43–66). New York: Springer.

Manatschal, A., & Thomann, E. (2011). *Vergleich des Organspendewesens in der Schweiz und Spanien.* Study mandated by the Swiss Federal Office of Public Health. Bern: Institute of Political Science and Büro Vatter AG.

Mill, S. J. (1843). *A system of logic*. London: Harrison and Co. Printers.

ONT. (2008). *Actividád de Donación 2008*. Madrid: Organización Nacional de Trasplantes ONT.

ONT. (2009). *Programa de garantía de calidad del proceso de donación. Memoria de resultados de la autoevaluación año 2009*. Madrid: Organización Nacional de Trasplantes ONT.

Pawson, R. (2002). Evidence-based policy: The promise of realist synthesis. *Evaluation, 8*(3), 340–358.

Peters, B. G. (1998). *Comparative politics, theory and methods*. Basingstoke: Palgrave.

Ragin, C. (2000). *Fuzzy-set social science*. Chicago: University of Chicago Press.

Rohlfing, I. (2012). Analyzing multilevel data with QCA: A straightforward procedure. *International Journal of Social Research Methodology, 15*(6), 497–506.

Ross, M. C., & Homer, E. (1976). Galton's problem in cross-national research. *World Politics, 29*(1), 1–28.

Swisstransplant. (2009). *Jahresbericht 2009*. Retrieved July 20, 2014, from http://www.swisstransplant.org/11/organspende-organ-transplantation-zuteilung-koordination-warteliste-statistiken.php

Tarrow, S. (2010). The strategy of paired comparison: Toward a theory of practice. *Comparative Political Studies, 43*(2), 230–259.

Vedung, E. (1998). Policy instruments: Typologies and theories. In M.-L. Bemelmans-Videc, R. C. Rist, & E. Vedung (Eds.), *Carrots, sticks and sermons: Policy instruments and their evaluation* (pp. 21–58). New Brunswick: Transaction Publishers.

Voß, J., Smith, A., & Grin, J. (2009). Designing long-term policy: Rethinking transition management. *Policy Sciences, 42*(4), 275–302.