Multilevel Analysis of Protest: Application for Small N Designs

Abstract
Protest is the result of complex multilevel processes. It is triggered by contextual factors such as political opportunities or events, depends on organizations’ mobilizing capacity as well as on the type of people who protest, and the characteristics of the populations they come from. In order to effectively study the antecedents that operate at various levels, social movement research needs to integrate data from multiple analytical levels and systematically examine the relationships across the various levels. While large N statistical techniques of multilevel modelling are well understood, less is known about applying multilevel analysis research examining small number of cases. The article develops conceptual and methodological tools for multilevel analysis of protests in studies with a small number of cases. First, it demonstrates the empirical requirements associated with analyzing three types of multilevel effects: contextual effects, composition effects and cross-level interactions. Next, specific multilevel small N designs that can be used to examine the three multilevel effects are presented. The last section uses the multilevel approach to examine the demobilization of anti-Iraq War protests in the United States.

Introduction
Protest is the result of complex multilevel processes. It is highly situational, triggered by contextual factors such as political opportunities or events. Its nature is also shaped by organizations’ mobilizing capacity and political actors’ recruitment efforts. Protest also depends on the type of people who protest, and the characteristics of the populations they come from. An effective study of the antecedents of protest that operate at various levels requires multilevel analysis. While large-N statistical techniques of multilevel modelling are well understood, less is known about applying multilevel analysis in research examining a small number of cases (or small N). Yet case study research of social movements often examines multilevel theories and considers data from various levels of analysis.

Drawing on insights from quantitative multilevel analysis and classical small N research, this article develops conceptual and methodological tools for multilevel analysis of protests in studies with a small number of cases. First, it discusses the main principles of multilevel analysis
and demonstrates the empirical requirements associated with analyzing three types of multilevel effects: contextual effects, composition effects and cross-level interactions. Next, the article introduces the principles of multilevel small N designs that can be used to examine observable implications derived from the three multilevel effects. The last section applies this approach to examine the demobilization of anti-Iraq War protests in the United States using multilevel data on anti-war protest events and protestors collected by Heaney and Rojas (2015).

**Multilevel character of social movements**

Determinants of protest and social movements are conceptually understood in a multilevel fashion to be a combination of the supply and demand sides of protest connected via various multilevel processes (Coleman 1990; Goldstone 2015; Klandermans 2004; McAdam, Tarrow, and Tilly 2001). The multilevel character of protest triggers various questions. Do countries show higher levels of mobilization for the women’s movement because of open gender-specific political opportunities, or because they have a greater share of people with feminist attitudes that protest? How do the experience and skills of social movement organization (SMO) members affect their strategies? Under what contextual conditions do micro-level economic grievances trigger protest? Do open political opportunities facilitate protest only for movements that have sufficient organizational capacity to mobilize?

In order to answer those questions, social movement research needs to integrate data from multiple analytical levels and systematically examine the relationships across the various levels. A quantitative method of multilevel modelling is well understood and has been employed to analyze protest using comparative survey data (Dalton, Van Sickle, and Weldon 2010; Quaranta 2015). Less is known about systematic multilevel analysis in the context of small-N research of social movements. To be sure, most case studies of social movements consider data from multiple levels and often examine multilevel theories. However, those multilevel analyses are generally conducted either in an informal and exploratory fashion or use a case- or process-oriented approach (della Porta 2008; George and Bennett 2005; Falleti and Lynch 2009; Klandermans and Staggenborg 2002; Della Porta 2014). Variable-oriented multilevel small N research of social movements as well as the methodological tools needed for such analysis have been largely overlooked.
Variable-oriented small N designs have been well-developed in the context of single-level analyses (Gerring 2010; George and Bennett 2005; Ragin and Rihoux 2008; Sekhon 2004), some methodological work on variable-oriented small-N designs has focused on their application in multilevel research (Kittel 2006; Anckar 2008; Przeworski and Teune 1970). For instance, Przeworski and Teune (1970, 72) originally introduced the most-different system design as a multilevel technique to control for macrolevel factors while examining microlevel effects. Developing this literature further and drawing on insights from quantitative multilevel analysis, this article develops multilevel approach in small-N research. It presents a conceptual understanding of the problem and uses simple empirical illustrations to make the arguments accessible to readers without a quantitative background. Other work explains the mathematical aspects of multilevel models (Gelman and Hill 2007).

The next section introduces three main types of multilevel effects – contextual effect, composition effect and cross-level interaction – and explains their observable implications. The subsequent section incorporates the empirical requirements of multilevel analysis into the logic of small-N comparative study. The last section applies this approach to examine the demobilization of anti-Iraq War protests in the United States using unique multilevel data on anti-war protest events and protestors collected by Heaney and Rojas (2015).

**Multilevel analytical framework**
The purpose of multilevel analysis is to examine sources originating from different levels of analysis. For simplicity, this article considers only two levels: macro (group) level and micro (lower-level sub-units that are nested within the group) level. The logic of multilevel analysis, however, extends to more complex settings, such as three levels or cross-classified structures with units represented by individuals who belong to several groups at the same time. Also, the labeling of levels as macro or micro depends on the context of the study. For instance, SMOs can be macrolevel groups in a study that analyzes the role of organizational structure in the level of protestors’ engagement, whereas they would be meso level in a three-level comparative study that examines human rights policies, mobilization of women’s rights groups and engagement of individual protestors.

The first step in multilevel analysis is to define the analytical levels at which the concepts of interest operate. Microlevel concepts are traits of microlevel units and can explain differences
among such units as well as across macrolevel groups because of their composition (this concept is discussed in more detail below). Macrolevel concepts capture characteristics at the macro level and potentially influence the sub-units that are members of specific groups. There are two basic types of macrolevel concepts: global and derived (Blakely and Woodward 2000; Schwartz 1994). Global macrolevel concepts (sometimes called analytical) are purely contextual phenomena that exist only at the macro level, such as political institutions, revolution, or income inequality. Derived macrolevel concepts (sometimes called aggregated) are contextual summaries of microlevel units, such as the proportion of ethnic minorities or the average budget of SMOs, which capture the contextual properties of the groups. Though derived factors are technically constructed from distributions of microlevel characteristics, they are conceptually different from the micro characteristics (Schwartz 1994). For instance, supportive public opinion, such as a high share of the population agreeing that women and men make equally good political leaders, indicates open discursive opportunities (McCammon et al. 2001). The discursive opportunity of a gender-equality-supportive culture is conceptually different from its microlevel counterpart of individual gender equality attitudes. Importantly, both microlevel and derived macrolevel factors have independent effects and they can condition each other. For instance, people with pro-gender equality attitudes might be more likely to join women’s movements in countries with open discursive opportunities than in those where opportunities are closed with a national political culture that is not supportive of gender equality.

Outcomes also have conceptual levels. While some outcomes in social movement research are systemic, such as revolutions or policy changes, many are, in fact, summaries of distributions of lower-level phenomena. For instance, protest events are aggregates of individual protestors across macrolevel groups, such as countries or time periods. The main advantage of multilevel analysis is that it enables empirical examination of outcomes as well as sources at their correct conceptual levels.

To explain the principles of multilevel analysis, this section uses an example of economic crisis and anti-austerity protest. An economic crisis is indicated by a high unemployment rate to demonstrate the theoretical and methodological differences between unemployment understood as a derived macrolevel factor and unemployment operating as a microlevel characteristic. Note that the principles of multilevel analysis are the same if analytical macrolevel factors (e.g. decentralized political institutions) are used instead of a derived macrolevel factor as applied in this example.
The second step in multilevel analysis is to theorize about the multilevel effects that connect factors operating at the different levels of analysis and specify their observable implications. Multilevel effects can be classified into three basic types: (1) composition mechanism, (2) contextual mechanism, and (3) cross-level interaction (Blakely and Woodward 2000; Greenland 2001). For example, suppose we wish to answer the following research question: Why do economic crises trigger anti-austerity protests in some countries but not in others? Three model theories represent the three different types of multilevel mechanisms: (1) theory of individual-level grievances as a composition mechanism; (2) economic crisis operating as a macrolevel grievance or threat that opens up national political opportunities as a contextual mechanism; and (3) theory suggesting that economic crises mobilize socio-economically deprived people because they reframe their personal grievances as a collectively shared mobilizing grievance as an example of cross-level interaction mechanism.

Table 1 and Figure 1 show a hypothetical simulation of how the three multilevel effects materialize empirically in the example of two countries (A and B) with a macrolevel positive relationship between anti-austerity protests (1,000 and 3,000 protestors) and economic crisis (10% and 30% unemployment rate). The right-hand columns in each panel in Table 1 show that both countries have 1 million people; country A has 1,000 anti-austerity protestors and country B has 3,000 anti-austerity protestors. The bottom rows of each panel in Table 1 display the number of unemployed people in the two countries in bold (100,000 vs. 300,000). The inner part of the tables shows protest participation among the employed and unemployed. In Figure 1, the x-axes in the individual graphs show the two countries with different unemployment rate (10% and 30%, respectively). The y-axes indicate protests: the square signs signify protests among the unemployed, and the circles denote protests among the employed; the size of the sign indicates the size of the group.

**Figure 1 about here**

**Table 1 about here**

*Composition mechanism*
The first type of multilevel effect is a composition mechanism. In general, composition theories are systemic derivatives of microlevel theories. Their macrolevel implication is that systemic differences result from a different share of microlevel processes. Our example of individual-level grievance theory suggests that people who suffer from socio-economic grievances will turn up at anti-austerity protests because they are personally affected by the economic crisis and thus have a reason to protest (della Porta 2014; Snow and Soule 2009). As Panel A in Figure 1 and Table 1 shows, the observable implication of this theory is that only unemployed people take part in anti-austerity protests in both countries, whereas no people with jobs protest (probability is 0).

Since it is a microlevel theory that does not consider the role of context, the implication is that individual grievances have a uniform effect across various situations and groups. Panel A in Figure 1 and Table 1 shows that the rate of protest among the unemployed is the same in both countries (1/100). The different levels of anti-austerity protest in the two countries must thus originate from a different share of microlevel relationships between unemployment and protest. As Panel A in Figure 1 and Table 1 shows, the only difference between country A and country B apart from the level of anti-austerity protest is a higher unemployment rate (30% vs. 10%), which directly affects the number of unemployed anti-austerity protestors (displayed by different sizes of squares across the two countries in Panel A of Figure 1 and by the number of unemployed protestors in Panel 1 of Table 1 (3,000 vs. 1,000)). All other parameters – protest rate of unemployed across the two countries, protest rate of employed across both countries, and the gap between employed and unemployed between the two – stay the same. As the empirical illustration of the composition effect shows, country B has more anti-austerity protestors than country A because it has a larger pool of potential protestors, i.e. more people with individual-level grievances who turn out to protest.

**Direct contextual mechanism**

A contextual explanation of how economic crises lead to anti-austerity protests represents the second type of multilevel effect shown in Panel B in Figure 1 and Table 1. Contextual effect theorizes that economic crises increase anti-austerity protests because a high unemployment rate is an exceptional threat that opens up opportunities for protest because it increases the salience of the economic crisis, disrupts existing political alignments and invites new actors to mobilize around this issue (Zamponi and Bosi 2016; Tarrow 2011). Unlike the composition mechanism, in
which the systemic effect of higher unemployment lies in a higher share of individual-level relationships between unemployment and protest, the cause has a contextual character in this scenario. A high unemployment level functions as a salient event that affects national politics and increases anti-austerity protests through its influence on opportunities that incentivize political mobilization. Since it is a contextual theory, it does not specify what types of people mobilize, or why some protest more than others in the same context. Contextual theory operates independently from (and on top of) microlevel processes that explain differences among people in the same situation. Specifically, in this example, the contextual theory assumes that open political opportunities increase a country’s central tendency, everyone’s probability, to protest. Independently from, and in addition to, the effect of people’s individual employment situations, all people in country B (which has a higher unemployment rate) are more likely to protest than those in country A (with lower unemployment).

Panel B in Figure 1 and Table 1 shows the contextual effect of the unemployment rate empirically. Country B, with a 30% unemployment rate and 3,000 protestors, has a much higher rate of protest among both employed and unemployed people (1/350 and 1/300 in Panel B Table 1) than country A with a low protest rate (1,000 protestors) and low unemployment rate (10%) among both employed and unemployed people (1/1,1050 and 1/700 in Panel B Table 1). Still, we can see that individual-level grievances matter, as unemployed people have a 0.0005 higher probability of protesting than employed individuals in both countries. However, the crucial difference from the composition effect shown in Panel A is that, in this scenario, the difference in levels of protest between employed and unemployed people does not explain the cross-country difference as it does in the composition explanation. As Panel B in Figure 1 shows, the main pattern explaining the difference between the two countries is the difference in the central tendency of both subgroups (employed and unemployed) to protest rather than the number of unemployed protestors.

**Cross-level interaction mechanism**

The third multilevel effect of how economic crises increase anti-austerity protests is cross-level interaction. Our example of cross-level effect theorizes that economic crises trigger anti-austerity protests because they motivate unemployed people to protest. As higher unemployment rate politicizes the issue of unemployment (Giugni and Grasso 2017; Jenkins and Perrow 1977),
unemployed people are, therefore, less likely to perceive their status as a personal failure than when the unemployment levels are low; they instead interpret their lack of a job as a collectively shared mobilizing grievance and turn out to protest (Snow and Soule 2009). Panel C in Figure 1 and Table 1 shows that the cross-level mechanism materializes with a much higher rate of protest among the unemployed in country B, which has a high unemployment rate (1/132) compared to employed people in the same country (1/968), and both the employed and unemployed in country A, which has a low unemployment rate (1/968 and 1/1,429).

The distinctive feature of the cross-level interaction effects is the contextual influence on the character of microlevel processes within the contexts. The contextual and composition effects discussed above do not contain contextual variation in the microlevel effects. As Panels A and B in Figure 1 and Table 1 illustrating the two previous mechanisms show, the gap in the protest rate between employed and unemployed individuals in the same context is uniform across both countries. By contrast, the cross-level interaction effect implies heterogeneity of the microlevel processes: the gap in protest participation by employed and unemployed persons in country A with a low unemployment rate is small and reversed (the probability of employed people to protest is by 0.0003 higher than of unemployed people in this scenario), whereas the gap between the employed and unemployed is much larger in country B (the probability of unemployed people to protest is by 0.007 higher than of employed people). The principle of the cross-level interaction mechanism is that the source of the different levels of anti-austerity protests in the two countries originates in this heterogeneity. Country B has a higher level of anti-austerity protests because the high unemployment rate boosts the protest rate of unemployed people.

The simulation showed how the three types of multilevel effects explain why a higher unemployment rate increases anti-austerity protests. As demonstrated, the multilevel analysis needs to simultaneously consider variation across and within the two groups in order to disentangle the three different theoretical mechanisms. Specifically, it requires information on the macrolevel distribution of causes and outcomes – the unemployment rate and the size of anti-austerity protests in the two countries – and information on the microlevel distributions of the explanatory factor and outcome within the groups – rates of protest among the employed and unemployed in both countries.
Obviously, single-level studies are not able to examine multilevel effects. Ecological studies, whether large or small N, cannot determine which of the three multilevel effects is at play. For instance, while an ecological study would have data on the unemployment rate and the number of protestors attending anti-austerity protest events (numbers in bold in Table 1), it does not have information on the within-systemic distributions on who the protestors are and the population they come from. In the simulation described above, all three multilevel theories were empirically possible using exactly the same macrolevel data on the unemployment rate and the size of anti-austerity protests. As demonstrated, different numbers can be plugged into the within-system distributions in the inner cells of Table 1, which can completely change the type of multilevel effect. This issue has been recognized as a problem of fundamental indeterminacy of ecological studies (Greenland 2001). Because ecological studies do not have within-group data, they cannot distinguish the various multilevel processes.

A similar problem applies to microlevel studies (Falleti and Lynch 2009; Schwartz 1994; Greenland 2001). Even if more microlevel studies from different groups are available, they cannot be used to study multilevel effects unless they are systematically integrated into a multilevel design that reflects cross-group and within-group variations. On their own, microlevel studies do not have the macrolevel data (contextual factors and composition of microlevel factors across groups) and information on central tendencies within groups (if they are not representative of the whole group) necessary to study the multilevel sources of differences among groups.

Furthermore, multilevel analyses are useful even if researchers do not primarily aim to disentangle the character of multilevel effects because unlike single-level studies they can effectively control for alternative explanations that originate at different analytical levels. The risk of bias due to confounders from other levels is an issue particularly in ecological studies. The reason is that ecological studies do not have information on joint distribution of microlevel factors that can confound the results of ecological studies (Blakely and Woodward 2000; Greenland 2001). For instance, the system-level relationship between unemployment and anti-austerity protest might be spurious. The higher levels of anti-austerity protest among the unemployed might result from a higher share of microlevel factors that jointly contribute to protest, such as a larger number of people who are simultaneously unemployed but with higher levels of education and involvement in SMOs. Ecological studies lack data on these microlevel joint distributions and thus cannot effectively incorporate them into the single-level analysis.
Multilevel small N designs
The purpose of sampling of cases in variable-oriented small N research is to isolate theoretically relevant variations among a smaller number of cases (Gerring 2010). The isolation of the variations is done via two sampling strategies, most similar and most different systems designs. The novel challenge in multilevel case study analysis is that it needs to coordinate the purposeful selection of cases at the macro level between groups and integrate it with a selection of sub-units nested within these macro groups.

Multilevel most-similar systems designs
The most-similar method involves choosing at least two cases that differ only in terms of the variables of interest – the explanatory factors and the outcome; all other characteristics that might affect the outcome must be similar. This approach is especially suitable for identifying and verifying the causal effects of interest (Gerring 2010). To identify the macrolevel element of the multilevel effect, we can use the standard most-similar method. It considers at least two groups that are similar in all relevant aspects but two factors – the macrolevel variation in the explanation of interest and the outcome. For instance, a study suggesting a contextual effect of tripartite consultative institutions among the state, workers and employers on the demobilization of the labor movement would examine one country with tripartite institutions and low labor protest and one country without tripartism and with a lot of labor protest while keeping other potential macrolevel explanatory factors – contextual as well as compositional – the same.

The novel element in multilevel design is the integration of a most-similar method at the level of microunits that are members of the macrolevel groups. The specific sampling method used differs based on the type of multilevel theory the study aims to examine. An analysis studying a contextual effect needs to include two most-similar microunits, each of which is in a different macrolevel group so that the only difference between the microunits is the macrolevel group to which they belong. Specifically, considering the example mentioned above, research examining the effect of tripartism on labor movement would sample one labor organization from a country with a tripartite institution and one organization that would be as similar as possible in theoretically relevant aspects from a country without a tripartism. Though sampling only one organization per country might seem counter-intuitive, the comparison including only one organization per country...
fully identifies the microlevel part of the contextual effect, i.e. the macrolevel contextual factor directly affecting the outcome of groups’ microunits. This design is comparable to identical twins being raised in different families. If all other factors except country membership are the same, nothing else can explain the variation in the number of protests between the two labor organizations.

In contrast to contextual theory, a multilevel study examining composition effects must combine the macrolevel most-similar comparison of two groups with the examination of four most-similar microunits that differ in the combination of values of microlevel explanatory factors and in their membership in macrolevel groups but are similar in all other ways. Consider an example of composition theory arguing, for instance, that countries have low levels of labor protest because they have more labor organizations with a centralized structure that decreases protest. At the macro level, such a study would sample one country with a high number of centralized labor organizations and a low level of labor strikes and another country with a high number of centralized organizations and many strikes; all other macrolevel factors, such as having tripartite institutions and other relevant macro factors, would be similar.

It is important to notice that the most-similar sampling of the two countries reflects the macrolevel implications of the composition theory, but there is no causal process taking place at the macro level. In composition effects, causal processes take place purely at the lower level; the macrolevel differences are only summaries of the microlevel effects. Therefore, a multilevel study of composition effect needs to focus the only difference available in the most-similar method on the microlevel causal effect underlying the composition theory. In the labor protest example, the microlevel causal effect is that decentralized labor organizations strike while centralized organizations do not, and this effect is uniform across countries. To identify this effect, the multilevel study examines two sets of two labor organizations in two countries that differ only in the centralization of their organizational structure; all other aspects are the same. To ensure that the only difference is organizations’ level of centralization, the sampling of microunits must use an identical “sameness” criteria in both countries; for instance, that labor organizations in each country have a similar number of members, similar budget, are involved in tripartite negotiations, have a similar age, rely on similar framing, etc. Also, notice that it is only the microlevel factor of de/centralization that is different among the microunits because the two countries that the
organizations are members of differ only in the number of such organizations, which is not a contextual factor in this scenario.

**Multilevel most-different systems designs**

The most-similar method can identify the causal effect of theoretical interest by simultaneously capturing the presence and absence of the explanatory factor while keeping other alternative characteristics the same. However, the weakness of this method is that it cannot tell whether the causal effect operates beyond the specific type of sampled cases, i.e., the values of factors at which the sampled cases are similar (Anckar 2008; Gerring 2010). For instance, imagine that the most-similar method applied the above-mentioned study of composition effect samples to two labor organizations with different degrees of centralization of their organizational structure, but both have a high budget, a large number of members, similar framing, are around 20 years old, etc. By examining only these two labor organizations, the study cannot determine whether the effect of centralization on protest operates only among organizations with, for instance, a large number of members or if it also demobilizes protests among low-membership organizations.

To tackle this issue, case study research often combines the most-similar design with a method that focuses on generating heterogeneity among cases, such as the most-different method or a diversity method (Gerring 2010). In contrast to the most-similar method, which aims to verify theories by empirically identifying the difference implied by the causal effect, the strength of the most-different and diversity methods is that they broaden the scope conditions of the causal effect as they examine whether the implications of the causal effect replicate in as diverse situations as possible. Specifically, the most-different method chooses cases that differ in a wide variety of respects but are the same in the key causal factor and the outcome. For instance, the multilevel study of the contextual effect of tripartite institutions mentioned above would sample most-different labor organizations in each country to generate heterogeneity within the macrolevel groups to show that the contextual effect, indicated by labor organization membership in the country, homogeneously increases the number of protests across various types of labor organizations. Notice that the only factors that the labor organizations in a given country have in common is membership in the same macrolevel group; all else is different. Therefore, the study can claim that the reason behind a similar level of protest among organizations in a country (relative to organizations from other countries) is due to the contextual effect.
A diversity method is another way of implementing heterogeneity into the case study design. Here the maximization of case variation is achieved by systematically selecting cases with individual values of various relevant factors (Gerring 2010). Specifically, this method identifies the most important alternative explanations and/or candidates for necessary conditions that are then used for sampling to show that the causal effect of interest operates even within these diverse cases. For instance, the multilevel study examining the contextual theory suggesting the effect of tripartite institutions would systematically sample, in both countries, labor organizations that are typical cases of, for instance, high and low membership and centralized and decentralized structure. As in the case of the most-different method, the purpose of this sampling would be to show that the contextual effect operates across diverse microlevel organizations with various characteristics.

To be sure, like in classical single-level small N studies, the cases with perfectly matching combinations of differences and similarities rarely exist. Because of that, the multilevel small N analysis can only try to approximate the isolation of most similar and most different characteristics of cases across multiple levels. Still, an approximate version of the isolated effects is valuable because it limits the range of possible competing explanations. Other approaches can then be used to further disentangle the those competing explanations. One strategy is a combination of the most similar and most different systems designs as suggested above. While most similar systems design can identify the causal effect of interest, the combination with most different systems design or a diversity method can compensate the imperfect similarity of compared cases by replicating the similarity patter across different most similar pairs.

The intimate knowledge of the studied cases is another source used in small N designs to disentangle alternative explanations (George and Bennett 2005). The fact that the study examines only a small number of cases allows researchers to have a rich understanding of the cases studied. Multilevel small N studies can utilize the rich knowledge of studied cases to theoretically evaluate the role of alternative explanations, to examine other types of observable implications or to trace the specific steps of the theorized causal processes.

A different strategy is to use mixed-methods designs and increase the number of studied cases at some of the levels and integrate the small N method with other types of research approaches. For instance, the ability of Qualitative Comparative Analysis (QCA) to detect causal
heterogeneity in middle-size samples can be well utilized to study variation of within system effects across a small number of macro systems (Denk 2010). For instance, a study examining the role of tripartite consultative institutions on protest of trade union organizations could use most similar systems design at the level of country (selecting one country with tripartite institutions and one country without tripartism while keeping other macro-level factors the same) and combine it with an all-cases design of a middle-size sample of trade union organizations within the two countries.

The data analysis would apply QCA to all individual trade union organizations analyzing various organizational-level characteristics and country membership (i.e. tripartism) included as an explicit element of the solution (Rohlfing 2012). Such analysis would be able to identify the three multilevel mechanisms explained above: direct contextual effect (a path including only the contextual condition and no other organizational-level characteristics), cross-level interaction effect (paths that include the contextual condition and organizational-level characteristics), and composition effects (paths that include only organizational-level characteristics without contextual condition). Note that the multilevel analysis is only possible thanks to the application of the most similar systems design at the macro-level. The isolation of competing macro-level explanations by purposeful sampling of the two countries allows identification of the macro-effect of tripartism. Without this step, the macrolevel element included in the QCA would indicate just organizations’ membership in country A and country B. An alternative strategy of including more macrolevel characteristics into the QCA would not work either as the analysis would not be able to distinguish the various macrolevel characteristics using just the two countries.¹

Following the same logic, we can combine a small-N study of macro groups with a quantitative analysis of a large number of microunits that are nested within the macro groups. Specifically, a study examining the effect of tripartite institutions would sample one country with and one without tripartism that are similar in all other relevant respects. In the next step, the study would analyze a nonbiased sample of microlevel units, such as a random sample or all-cases sample of large number of labor organizations, in both countries. Such a design would help identify the contextual effect of interest, in a similar way that the small-N case study combining the most-similar sampling and intimate knowledge of the studied cases did in the examples discussed above. The large N part with nested micro-units would allow statistical modeling of variation at the micro
level; this part of the design would strengthen the ability of this study to capture joint distribution of microlevel confounders.

Technically, this analysis would be an integration of a small N purposeful sampling of the two macrolevel groups with a Chow test (Chow 1960). It would use a regression analysis with surveyed organizations as the units of analysis, their reported level of protest as the dependent variable and a dummy variable indicating organizations’ country membership that, in this case, captures the factor of tripartite institutions because the two countries were selected as most-similar cases. Such analysis would identify the direct contextual effect of tripartism by a significant direct effect of country membership after controlling for organization-level factors; cross-level interaction effect would be indicated by a significant interaction effect between country membership (i.e. tripartism) and organization-level characteristics; composition (purely organization-level) mechanism would be indicated by a mediation effect of organization-level characteristic explaining the effect of country membership on organizations’ protest.

Application: demobilization of anti-Iraq War protest in the US
This section illustrates an application of a small N multilevel method to examine Heaney’s and Rojas’ (2015) theory that identity shift among Democrats explains the demobilization of American protests against the Iraq War. In their exceptional study on the relationship between party and protest politics, Heaney and Rojas collected data on protest events against the Iraq War in the United States from September 2001 to October 2012, data on the material indicators of the intensity of the war during that period and they surveyed a representative sample of protestors at anti-war demonstrations over the cycle of the protests and organizers and movement activists. Heaney and Rojas show detail information on the data used. The analysis presented in this article uses some of the data assembled to demonstrate the value of the multilevel small-N study. The purpose of this application is a heuristic exercise to illustrate the steps in a multilevel examination of theory about Democrats identity shift.

Heaney’s and Rojas’s (2015, 66) protest event analysis shows that the anti-Iraq protests peaked at between 20,000 and several hundred thousand protestors when Republicans controlled Congress and the presidency from 2001 to 2006. After Democrats gained a majority in Congress, antiwar protests slowly declined to the level of tens of thousands to a few thousand participants and declined even further to a few hundred to a few thousand participants after Obama became
president in 2009. The authors develop a theory of identity shift, suggesting that the movement demobilized because anti-war protestors who identified as Democrats prioritized their partisan loyalty over their identification with the movement and therefore stopped protesting after the Democrats came to power. Specifically, when protesting the Republican administration’s handling of the war in Iraq, Democratic protestors identified, according to the authors, with both political actors, the anti-war movement and the Democratic Party. However, the two political actors realigned after the Democrats’ electoral victories in 2006 and 2009 and Democratic protestors had to choose between their loyalties to the movement and to the party (Heaney and Rojas 2015, 91). Since for most people, partisan identities are stronger than movement identities and positions on specific policy issues, most Democratic protestors chose the party and left the movement. Importantly for our analysis, Heaney and Rojas assume that this process explains the aggregate result of the movement’s demobilization and is responsible for the decline in anti-war protests in the United States. They explained, “when Democrats stopped turning out, the movement could no longer achieve critical mass” (Heaney and Rojas 2015, 114).

To examine this theory, I use the multilevel framework introduced above. Heaney and Rojas’ theory of identity shift among Democrats implies a cross-level interaction effect (Panel C in Figure 1): the reason why we see the aggregate decline in anti-war protests is due to the change in context – Democrats coming to power – that alters the behavior among only one group of protestors – participants identifying as Democrats – who turn out in smaller numbers than they did when the Republicans were in power. As the identity shift mechanism happens only in a specific combination – people who identify as Democrats and only in the context of a Democratic government – the protests of people who do not identify as Democrats should be immune to the contextual change. Note that the reason why only Democrats are affected by the changing context while people not identifying as Democrats are not affected is an implication determined by the very nature of the micro-level mechanism of identity shift suggested by Heaney and Rojas. By principle, people who do not identify as Democrats cannot even develop the theorized causal mechanism of putting a priority emphasis on their Democratic identification.

We can empirically examine this theory with the multilevel small N research design outlined above. To identify the macrolevel part of the theory, the design samples two protest events that took place during a period with and without a Democratic government and with high and low levels of protest from the anti-war movement that are similar in other relevant factors and for which
protest survey data were collected. The first protest event is the “End of the War on Iraq!” protest that took place in Washington, DC on September 24, 2005. The analysis uses a middle category of 200,000 as the number of protestors that lies between organizers’ estimate of the turnout of 300,000 protestors and the media estimate of 100,000 protestors. The second protest is the “March on Wall Street” that was held in New York City (NYC) on April 4, 2009. The analysis uses 5,000 as the number of protestors based on Heaney and Rojas’ estimate of “thousands” from their participant observation; organizers’ estimate was 10,000 and the media estimate was hundreds.

While the two events differ in the contextual factor of Republicans in power in 2005 and Democrats in power in 2009, and the levels of protest were hundreds of thousands vs. thousands, they are similar in other relevant aspects. For instance, in both periods more than 160,000 US troops were deployed to the Middle East; both protests occurred in large cities on the East Coast of the US that show similar demographics and both have a similar share of Democrats in their populations (56%, Pew Research Center 2020). The two protests differ in the size of the population available to protest in the two different areas with Washington DC Metro Area having a population of around six million people and New York City Metro Area having a population of around 20 million. However, the difference has an opposite direction than predicted by identity-shift theory – larger NYC having a smaller protest in 2009 and smaller Washington DC having a bigger protest in 2005 – and thus probably does not confound the relation between Democrats-Republicans at power and the size of the two protests. Importantly, though neither of protests was primarily a partisan protest, protest in Washington DC might be more politically targeted at the party in power and increase protest among partisans while protests in NYC might be less attractive to people wanting to protest a particular party in power. The observable implication of this theory and the theory about identity shift among Democratic protestors is identical and this design would not be able to disentangle the two explanations. If the results of the analysis were supportive, we would need to use additional strategies outlined above to distinguish the two theories. In the microlevel part of the multilevel design, the study employs representative large N data on protestors at the two events collected by Heaney and Rojas (2015, 108–14) that provide information on protestors’ partisan identification. It also uses information from surveys conducted by Pew Research Center (2020) to get data on partisanship among the local populations (Washington DC Metro Area and NYC Metro Area).
To analyze the data, the study uses the procedure explained in Table 1 to trace the individual parameters we need. As shown in Table 2, first, we fill in the macrolevel factors of interest (shown in bold): Panel A indicates the period when Republicans were in power and 200,000 protestors; Panel B indicates the period when Democrats were in power and 5,000 protestors. The analysis then needs to fill in the rest of the information indicating the within-protest distributions. Heaney and Rojas (2015, 114) report the percentage of protestors by partisanship, which is used to calculate the absolute number of Democrats and others at each of the protests. As few self-identified Republicans took part in the anti-war protests (Heaney and Rojas 2015, 114), I leave them out of the analysis; I classify people with no party identification and those identifying with minor parties as “others.” Specifically, 50% of Democratic protestors at the event in 2005 that had 200,000 participants translates into 100,000 Democratic protestors and the same absolute number of other protestors. Likewise, 30% Democratic protestors correspond to 1,500 in absolute numbers in the second protest in 2009 and 70% of other non-Republican protestors makes 3,500 protestors.

Table 2 about here

Next, the analysis considers information on the rate of protest among Democrats and others to examine the expected cross-level interaction effect. In order to calculate those numbers, we need to know the absolute number of Democrats and others who do not identify as Republicans in the population. The Pew Research Center surveys report partisanship among people living in the two metro areas (Pew Research Center 2020). We use the size of the population in the two metro areas to calculate the absolute number of Democrats and others in the population in the two areas of protest. Since the population in Washington DC Metro Area is 6.2, using the 56% of Democrats indicated by the Pew Research Center survey leads to a calculation of 3.472 million Democrats in the population, as indicated in the bottom row of the first column in Panel A of Table 2. When 100,000 out of these 3.472 million Democrats took part in the “End the War in Iraq!” protest, the probability that a Democrat would protest was 0.02880, as shown in the top row of the first column in Panel A of Table 2. Replicating the same calculations for the other three groups of non-Republican “others” in 2005 Washington DC and 2009 NYC and Democrats in 2009 NYC, we can fill in the rest of the tables and calculate the protest rates that we need to test Heaney and
Rojas’ theory. The probabilities are shown in italics in the top row of the two panels in Table 2 and illustrated graphically in Figure 2.

**Figure 2 about here**

The theory of identity shift among Democratic protestors due to Democrats becoming incumbents implied a cross-level interaction effect predicting a decline in the protest rate among Democrats in 2009 Washington compared to 2005 NYC, while the protest rate of others who do not identify as Democrats should remain the same in both protests as they did not experience the theorized complex psychological process of identify shift. In sum, if the theory is correct, the only difference that we should see between the two protest events is a decline in the protest rate among Democrats.

As the probabilities of protest in Table 2 and Figure 2 show, the results do not support this theory. There is little difference between Democrats and others in how much their rate of protest decreased between the periods when Republicans and Democrats controlled the government. The rate of protest among Democrats declined by 99.5% (from 0.02880 to 0.00013) and the rate of protest among others declined by 99.1% (from 0.10080 to 0.00096) from 2004 Washington DC to 2009 NYC protest. In contrast to the theory that expects identity shift to be the main reason why the protests declined after the withdrawal of Democrats, we see that other protestors who did not experience an identity shift withdrew at a very similar level. The results show that the main difference between the two periods of protests is not a cross-level interaction effect, but a decline in the tendency of both Democrats and others to turn out to protest. This finding of a similar decline in protest rate of the two partisan groups suggests that the reason why the anti-war movement demobilized was not because of a process that was unique to Democrats, as suggested by the theory of identity shift among Democrats, but due to some contextual effect that was common to Democratic and other protestors.

For instance, the reason behind declined participation rate among Democrats as well as others might be reduced mobilization efforts by Democrat-affiliated organizations and activists, which is another explanation that Heaney and Rojas develop (2015, chap. 5). This causal mechanism is, however, different from the theory expecting identity shift at the level of ordinary protestors. In the contextual scenario, the process of shifting loyalties operates at the level of
mobilizing actors and decreases their recruitment activity that then affects ordinary protestors via lower chance of being asked to protest regardless their individual partisanship.

**Conclusion**

Socio-political phenomena are results of complex multilevel processes. The collection of representative large-N data from multiple levels is often not possible or suitable. Next to social movement research there are other fields in social-science research that work with a small or middle size number of cases, such as comparative research of political parties, regional differences in policy-making, implementation of foreign aid in different communities or comparative analyses of higher education. Small N research is well-suited for multilevel analysis. A study examining a few cases at one level of analysis can be easily expanded with a collection of data from different analytical levels. In fact, most existing small N studies already study data from various analytical levels. The goal of this article was to provide small N researchers with analytical tools to systematically integrate the data from multiple levels into a variable-oriented small N multilevel analysis. While this strategy will not provide an answer to all issues, these analyses can strengthen the arguments based on deep knowledge of cases. It can also help to analytically approach micro-level confounders in ecological small N studies, and inspire new multilevel theorizing.
Bibliography


Table 1: Empirical Representation of Multilevel Effects

<table>
<thead>
<tr>
<th></th>
<th>Country A</th>
<th></th>
<th>Country B</th>
<th></th>
<th>Country B</th>
<th></th>
<th>Country B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protest</td>
<td>0</td>
<td>1/100</td>
<td>1000</td>
<td>0</td>
<td>1/100</td>
<td>3,000</td>
<td></td>
</tr>
<tr>
<td>No protest</td>
<td>900,000</td>
<td>99,000</td>
<td>1,000,000</td>
<td>700,000</td>
<td>297,000</td>
<td>1,000,000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>900,000</td>
<td>100,000</td>
<td>1,000,000</td>
<td>700,000</td>
<td>300,000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>Unemployed</td>
<td></td>
<td></td>
<td>Employed</td>
<td>Unemployed</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**PANEL A: COMPOSITION EFFECT**

<table>
<thead>
<tr>
<th></th>
<th>Country A</th>
<th></th>
<th>Country B</th>
<th></th>
<th>Country B</th>
<th></th>
<th>Country B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protest</td>
<td>1/1,050</td>
<td>857</td>
<td>1000</td>
<td>1/1700</td>
<td>2,000</td>
<td>1/350</td>
<td></td>
</tr>
<tr>
<td>No protest</td>
<td>899,143</td>
<td>99,857</td>
<td>1,000,000</td>
<td>698,000</td>
<td>299,000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>900,000</td>
<td>100,000</td>
<td>1,000,000</td>
<td>700,000</td>
<td>300,000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>Unemployed</td>
<td></td>
<td></td>
<td>Employed</td>
<td>Unemployed</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**PANEL B: DIRECT CONTEXTUAL EFFECT**

<table>
<thead>
<tr>
<th></th>
<th>Country A</th>
<th></th>
<th>Country B</th>
<th></th>
<th>Country B</th>
<th></th>
<th>Country B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protest</td>
<td>1/968</td>
<td>930</td>
<td>1000</td>
<td>1/968</td>
<td>723</td>
<td>1/132</td>
<td></td>
</tr>
<tr>
<td>No protest</td>
<td>899,100</td>
<td>99,900</td>
<td>1,000,000</td>
<td>697,900</td>
<td>299,100</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>900,000</td>
<td>100,000</td>
<td>1,000,000</td>
<td>700,000</td>
<td>300,000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>Unemployed</td>
<td></td>
<td></td>
<td>Employed</td>
<td>Unemployed</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**PANEL C: CROSS-LEVEL INTERACTION EFFECT**

Note: The data show a simulation example illustrating observable implications of the three multilevel effects.

* The number shows protest rate among the group of employed or unemployed.
Table 2: Analysis of cross-level interaction effect in anti-war protest

<table>
<thead>
<tr>
<th></th>
<th>Republicans at Power</th>
<th>Democrats at Power</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(September 24, 2005 “End the War on Iraq”, Washington DC)</td>
<td>(April 4, 2009 “March on the Wall Street”, NYC)</td>
</tr>
<tr>
<td>P Protest</td>
<td>.02880</td>
<td>.00001</td>
</tr>
<tr>
<td>Protest</td>
<td>100,000</td>
<td>1,500</td>
</tr>
<tr>
<td>No protest</td>
<td>3,372,000</td>
<td>11,366,500</td>
</tr>
<tr>
<td>Protest</td>
<td>200,000</td>
<td>3,500</td>
</tr>
<tr>
<td>No protest</td>
<td>3,472,000</td>
<td>11,368,000</td>
</tr>
<tr>
<td>Democrats</td>
<td>992,000</td>
<td>3,654,000</td>
</tr>
<tr>
<td>Other</td>
<td>4,464,000</td>
<td>15,022,000</td>
</tr>
</tbody>
</table>

Figure 1: Three Multilevel Effects

Panel A: Composition Effect
- Employed
- Unemployed

Panel B: Contextual Effect
- Employed
- Unemployed

Panel C: Cross-level Interaction
- Employed
- Unemployed

Note: The Figures show a simulation example illustrating observable implications of the three multilevel effects based on data presented in Table 1.
Figure 2: Anti-War Protest and Partisanship during Republican and Democratic Administration

Note: The Figures is based on data presented in Table 2.
Endnotes

1 The multilevel QCA analysis lies in the combination of Denk’s (2010) and Rohlfing’s (2012) suggestions. While Denk’s strategy recognizes the multilevel character of the data, it captures only cross-level interaction effects and fails to detect direct contextual and compositional (purely lower-level) effects. Rohlfing’s article recognizes this issue but it does not solve the identification problem of collinear macro-level characteristics as the application shown in the article is conducted on single-level data that, unlike units nested in macro groups, do not share all characteristics of the context they are embedded in.

2 The results are substantively similar if we would take the whole US as the source population of the two protests (specifically 99% decline in protest rate among Democrats and 97% decline in protest rate among others) as well as if Washington DC Metropolitan Area had the same absolute size of its population like NYC Metropolitan Area (specifically 99 % decline for Democrats and 97 % decline for others).