LIVES WORKING PAPER 2012 / 19

TITLE

Spatially weighted context data and their application to collective war experiences

Research paper

Authors

Guy Elcheroth,
Sandra Penic, Rachel Fasel,
Francesco Giudici,
Stephanie Glaeser,
Dominique Joye,
Jean-Marie Le Goff,
Davide Morselli, Dario Spini

http://dx.doi.org/10.12682/lives.2296-1658.2012.19
ISSN 2296-1658

FNSNF
SWISS NATIONAL SCIENCE FOUNDATION

The National Centres of Competence in Research (NCCR) are a research instrument of the Swiss National Science Foundation.

LIVES
Swiss National Centre of Competence in Research

Authors

Elcheroth, G.; Penic, S.; Fasel, R.; Giudici, F.; Glaeser, S.; Joye, D.; Le Goff, J.-M., Morselli, D.; Spini, D.

Abstract

In this paper, we introduce spatially weighted context data as a new approach for studying the contextual dimension of factors that shape social behaviour and collective worldviews. First, we briefly discuss the current contribution of multilevel regression to the study of contextual effects. We subsequently provide a formal definition of spatially weighted context data, as a complement to and extension of the existing multilevel analyses, which allows studying contextual influences that decrease with increasing distance, rather than contextual influences that are bound within discrete contexts. To show how spatially weighted context data can be generated and used in practice, we present a research application about the impact of the collective experiences of war across the former Yugoslavia. Using geographically stratified survey data from the TRACES project, we illustrate how empirical conclusions about the collective impact of war events vary as a function of the scale at which context effects are being modelled. Furthermore, we show how observed geographic patterns can be explained by underlying patterns of social proximity between the concerned populations, and propose a procedure to estimate the part of spatial dependency explained by models applying specific definitions of social proximity. In the final section, we discuss the boundary conditions for the use of spatially weighted context data and summarise the contribution of the proposed approach to existing methods for the study of context effects in the social sciences.

Keywords

Multilevel analyses | Spatial analyses | Context effects | Scale effects | Collective experiences | War exposure | Formative years | Collective guilt | Former Yugoslavia

Authors' affiliation

University of Lausanne, Switzerland

Correspondence to

Guy.Elcheroth@unil.ch

** *The TRACES project was financed by the Swiss National Science Foundation (SNF, grant number No100012-109623). Our current work on context data benefits from financial support by the SNF within the framework of the NCCR LIVES. The authors are grateful to Mathieu Cossutta and Till Junge for their work on the R-package Spacom, which accompanies this paper, and for their competent advice on mathematical equations.*

LIVES
Swiss National Centre of Competence in Research

LIVES Working Papers
University of Lausanne
Bâtiment Vidy ▪ 1015 Lausanne ▪ Switzerland
T. + 41 21 692 38 38
F. +41 21 692 32 35
contact@lives-nccr.ch

*** Guy Elcheroth drafted the paper and coordinated the working group, which developed the spatially weighted context data approach. Sandra Penic performed most of the final data analyses presented in the paper. All other co-authors, who appear in alphabetical order, made a substantial contribution to the conceptualisation of spatially weighted context data and to the implementation of the approach within the TRACES project.

Swiss National Centre of Competence in Research

LIVES Working Papers
University of Lausanne
Bâtiment Vidy ▪ 1015 Lausanne ▪ Switzerland
T. + 41 21 692 38 38
F. +41 21 692 32 35
contact@lives-nccr.ch

1. Background and motives for the development of spatially weighted context data

In the present paper, we introduce descriptive analyses and multilevel modelling using spatially weighted context data. This tool for contextual analysis complements classic multilevel regression in conditions where critical assumptions regarding the underlying structure of contextual data do not fully apply and, more specifically, when contextual units lack clear-cut boundaries. In these circumstances, spatial weighting functions allow to study open-ended contextual influences that unfold as a decreasing function of geographic and/or social distances.

As we will explain in the following paragraphs, the idea of spatially weighted context data emerged from the particular framework of a research project about the psychosocial impact of political transition and war on a single generation in the former Yugoslavia. However, although the path that led us to spatially weighted context data was based on a particular substantive agenda and research site, there are good reasons to believe that the methodological challenges that were *revealed* by these particular contingencies are not completely *bounded* to these contingencies. The specific research background has definitively made it infeasible to ignore certain limitations of the existing methods. This does not mean that these limitations do not exist in other contexts. Furthermore, if the same problems exist elsewhere, the solutions to these problems that we have developed might also be valuable elsewhere. Thus, the main goals of the present paper are to facilitate the transfer of a partially novel approach to contextual data analysis to other research fields and sites and to stimulate a wider debate on the current methods of contextual analysis and their possible alternatives or extensions.

Against the backdrop of comparative survey research on the impact of collective war experiences across the former Yugoslavia, three different types of factors combined to make the requirement for new context analytic tools highly salient. The first factor was our theoretical agenda: we aimed to contribute to the development of models of the way *s*ocial behaviour is rooted in collective experiences, particularly how the social behaviour relevant to the reconstruction of postwar communities is rooted in people's collective exposure to war events (Elcheroth, 2006; Spini, Elcheroth & Fasel, 2008). The basic premise that drives our thinking in this field is that the way in which people act depends on the way that they interpret their social worlds. Moreover, how they act and think is shaped not only by previous events that they experienced firsthand but also by events that have happened to other people to whom they are connected in some way, including those that they have never met personally (see Elcheroth, Doise & Reicher, 2011). If this simple insight holds some truth, it potentially complicates the work of empirical social scientists. Taking the contextual dimension of the impact of collective experiences into account introduces a series of unresolved methodological challenges related to the operationalisation of 'context' in quantitative data analyses.

The second factor that motivated our interest in alternative analytic tools stemmed from a situation in which new context data had to be created using finite resources. Once we had chosen the former Yugoslavia as the site to deepen the analyses of the social impact of collective war experiences, the tension between the need to remedy the shortage of social scientific data documenting the exposure of people in the former Yugoslavia to the war events and the limited resources that are generally available to generate such data about this part of Europe became dramatically apparent. This situation provided a pragmatic incentive to take a second look at the assumption that comparative contextual data can only be reliably obtained when representative, and thus sufficiently large, probability samples are available for *each* contextual unit and to creatively explore alternative strategies.

The third factor is also related to the studied site and phenomena. While the social scientists analysing, and possibly comparing, Western societies might often comfortably rely on nation-states with stable institutions and boundaries as embedding contexts for the social dynamics that they aim to explain, no similarly consensual definitions of contexts are available for the former Yugoslavian postwar societies. In contrast, research on nationalism in contexts torn by competing definitions of nationhood makes it very difficult to rely on the taken-for-granted definitions of nations as unproblematic contextual units. These conditions heighten the need to rely on non-politicised definitions of contextual units, which do not depend on the phenomena under investigation, e.g., the fact that to some people, at some point in time, Croatia, the Serb Republic of Bosnia, or Kosovo constituted separate nations, whereas for others and/or at other points in time, the same territories and populations were viewed as inseparable parts of Yugoslavia, Bosnia-Herzegovina, or Serbia.

To define the potential gaps in the social scientists' methodological toolbox that can be filled by analyses using spatially weighted context data, we herein attempt to abstract more general underlying assumptions, boundary conditions, and comparative advantages of this new approach to contextual analyses beyond their particular applications to collective war experiences in the former Yugoslavia. Thus, in the remainder of this paper, we will begin by briefly reviewing the contributions of multilevel analysis to the study of context effects and then explain why multilevel analysis was not sufficient to describe and explain the impact of collective war experiences in the former Yugoslavia. In the third section, we outline formal definitions of spatially weighted context data and the multilevel regression models that incorporate these data. We will then present in more detail the application of spatially weighted context data to the collective war experiences in the former Yugoslavia for descriptive (section 4) and modelling (section 5) purposes. In section 6, we attempt to locate spatially weighted context data within the broader social scientific movement toward the 'rediscovery of context'. We also discuss possible avenues for transferring the method to other fields and spell out the boundary conditions that must be respected when doing so. Section 7 concludes with summarising statements about the current and potential contributions of spatially weighted context data to comparative survey methodology.

2. Why was classic multilevel regression insufficient?

Over the last decade, multilevel regression has become a standard method of quantitative data analysis among social scientists. Essentially, multilevel analysis allows the variance of a micro-level outcome variable to be decomposed across two or more levels of analysis, and to explain this variance using specific predictor variables at each level (Hox, 2002; Courgeau & Baccaini, 1998). A widely used terminology distinguishes between so-called 'random effects' and 'fixed effects' in multilevel models. In the simple and frequent case of a two-level model where individual survey responses are hierarchically nested into clusters, groups or, as we will call them hereafter, contextual units (i.e., several individuals are affiliated to each context), 'random effects' refer to the fact that the intercept and/or slope values of the individual-level regression coefficients can be different in each context and are assumed to vary according to a distribution that can be summarised by a central tendency and variance component. Provided that the variance component is significant and non-trivial (i.e., regression coefficients do vary across contexts), relevant context data can be entered into the multilevel model to estimate the extent to which the observed variations between contexts can be explained by 'fixed effects' of specific contextual variables.

Three qualities of multilevel analysis are of particular interest for the present discussion. The first is that multilevel analyses represent the tool of choice for disentangling composition and context effects. As long as the data are measured and analysed at a single level, individual and ecological relationships are inevitably confounded. As a concrete example, unemployed people might have different political behaviour than employed people. The sum of these differences would constitute the *composition effect* component of the correlation between unemployment rates and, for example, the electoral outcomes at an aggregate level. However, high levels of unemployment also change the context within which *all* citizens (i.e., including the employed) make political judgements (see Lewis-Beck, 1990). Part of the observed correlation at the aggregate level is caused by a *true context effect* and cannot be reduced to the sum of the individual-level correlates of unemployment. According to Bakke, O'Loughlin, and Ward (2009, p. 1016), in the sole field of electoral geography, a "growing recognition of the value of multilevel modeling approaches for separating out the individual (first-level) and community (second-level) effects has generated dozens of studies that indicate modest (5 to 15 percent) but important contributions of contextual effects".

If one suspects that both composition and context effects contribute to an observed correlation, multilevel regression provides an appealing solution, provided appropriate micro-level data are available and individuals can be related to the relevant social contexts (see Diez-Roux, 1998). The same micro-level measure can then be used in a multilevel model as an individual-level predictor to control for composition effects and as an aggregate indicator to test for the presence of a true context effect in the variations between contexts that remain unexplained by the composition effect. Elcheroth (2006) applied this procedure to the collective war exposure and showed that composition and context effects actually went in two opposite directions. While living in an environment where war trauma is a frequent experience appeared to facilitate collective stigmatisation of war crimes at a community level, personal experiences of war trauma paradoxically increased the chances of passive acceptance of war crimes. Interestingly, Bakke, O'Loughlin, and Ward (2009) reported a similar pattern of opposite directions of the individual and contextual effects of war exposure in North Caucasus, although these authors were interested in different outcome variables and used a different contextual scale. In their study, personal experiences of violence were associated with an increased likelihood of never forgiving members of other nationalities, but contextual exposure to violent events occurring within a radius of 50 kilometres was associated with a *lower* likelihood of never forgiving.

It is important to stress though that the use of aggregate measures as context data appears to presuppose that the sample of individual values measured within each context is actually representative of the context. In the previously mentioned example, aggregate community values could be used because a sufficiently large random sample of respondents was available for each of the studied communities. However, these survey circumstances might be the exception rather than the rule. If representative samples are required *within* each context, then the prospects of generating context data at a lower scale, e.g., treating small geographic areas rather than entire countries as contextual units and calculating different aggregate values for each of them, are seriously limited.

The second notable quality of current multilevel modelling is that it allows for the adjustment of parameter estimates in particular contexts, when their reliability is weak, in particular due to a low number of observations within the context and/or high variability across the observations. A procedure generally referred to as *shrinkage* and, more specifically, as 'partial pooling' (Gelman & Hill 2007) or 'empirical Bayes estimation' (Hox, 2003) consists of computing a *weighted average* of the mean of the observations within a particular context and the mean of the observations across all contexts. A higher weight is given to the overall mean relative to the specific context mean when the latter is imbued with high uncertainty, i.e., when it relies on a small number of observations and/or observations that vary strongly among themselves. An application of this approach outside of the social sciences nicely exposes its relationship with the more general Bayesian approach of adjusting statistical estimates as an inverse function of their plausibility. In the context of a survey on the water quality within and across different lakes, Stow, Lamon, Qian, Soranno & Reckhow (2008) pooled estimates of the mean chlorophyll concentrations in specific lakes to generate the overall mean. In lakes yielding few observations or a large variance: "The multilevel estimate provides a sensible way of discounting the information that is less trustworthy. This approach is consistent with the way that Bayes theorem pools information from different sources. The information represented in the overall mean can be seen as the prior for an individual lake, while data from the lakes are treated as observations." (p. 117)

As we will explain, there are important similarities between the (Bayesian) shrinkage and spatial weighting functions. Both approaches rely on weighted means to make other information available in the data system useful for constructing particular contextual values. However, as we will also describe, the precise reasons for which the weights are used and, more importantly, the method by which they are computed differ substantially between the two approaches. Unlike multilevel analyses using spatially weighted context data, multilevel analyses using Bayesian shrinkage appear to remain faithful to the classic multilevel framework, where contextual units are permutable. Weighting for the overall mean implies being blind to the fact that depending on the relationships between three specific contexts, A, B and C, the data collected in context A might be more relevant than those collected in context B to adjust the estimates in context C.

This does not mean that all contexts are treated equally. In particular, contexts involving larger sample sizes are treated as 'more relevant' than those with small sample sizes for adjusting the estimates for all other contexts. However, the within-context sample sizes derive from the sampling design and can be completely independent of the population characteristics. If we want to push the spatial analogy further, the weighting scheme underlying data shrinkage in current multilevel modelling can be thought of as a researcher-driven space, rather than one driven by reality. The 'centrality' in the system is defined by attributes of the sampling design rather than by the structure of the system in the population. Therefore, while (Bayesian) shrinkage does effectively avoid estimates that are highly dependent on random variations, this method does not easily lend itself to clear substantial interpretation when applied to observations that are nested within contextual units of variable levels of binomial interdependence[i].

A third quality of the existing multilevel techniques should be mentioned because it is directly relevant to the analysis using spatially dependent context data. While previous generations of multilevel regression have assumed statistical independence between contextual-level measures, Banerjee, Carlin and Gelfand (2004) described how to estimate multilevel models with spatially dependent random effects, and Savitz and Raudenbush (2009) added a spatially dependent error term to multilevel regression analysis to generate so-called spatially autoregressive models.

While these relatively recent developments have successfully *neutralised* spatial dependency as an impediment to multilevel analysis, the next step is to *exploit* spatial dependency as an interesting phenomena in itself. In fact, one avenue toward 'spatialising' social analyses that has been relatively underestimated thus far (with the notable exception of the work of Chaix, Merlo & Chauvin, 2005) consists of incorporating spatial functions into the definition of contextual predictor variables. The way in which spatial models are presented in authoritative handbooks reflects this current situation well. While LeSage and Pace (2009) mention the possibility of so-called 'spatial-lag-of-X-models', i.e., models that compute "a spatial average of neighbouring home characteristics" (p. 30), they do not discuss this possibility (unlike models implying a spatial-lag-of-Y) in further detail. When summarising the existing motivations for using spatial analysis, they actually equate spatial models *per se* with "models containing spatially lagged dependent variables" (p. 42).

2.1 What multilevel analysis presupposes

The previous description of multilevel analyses has already alluded to the possible tensions between the goals inherent in the study of collective (war) experiences and the statistical assumptions underlying multilevel modelling. However, there are still more precise reasons why, despite the described qualities, classic multilevel regression appeared to be partially at odds with the nature of both the survey data in our hands and the social dynamics that we aimed to elucidate. The core of the problem lies in the implicit assumption inherent to multilevel analysis that the contextual influences stop at the boundaries of the contextual units (and are homogenous within the boundaries of these units). For example, using the rate of war-related events within a given region in a multilevel model as a contextual variable to explain collective worldviews *within* that region would imply that we assume that everyone's formative context within the region is equally affected by *all* of the events within the region and *not at all* by any events outside of the region.

If we follow that logic and, for example, define 'Croatia' as a contextual unit, then we assume that the dramatic war events in 1991 during the siege and subsequent devastation of the city of Vukovar, on the Eastern boundary with Serbia, affected the social context of the inhabitants of Zagreb, Istria, or Southern Dalmatia as much as the context of people living in Vukovar itself, while the impact of the same events becomes completely irrelevant only a few kilometres East from Vukovar, once one crosses the Croatian-Serbian boundary. Alternatively, if we chose a more local scale and define, for example, the Central Bosnian Lašva valley as a context in itself, then the massacres that occurred here in the spring of 1993 would affect the inhabitants of nearby Sarajevo as little as it affected the residents of Zagreb, Belgrade, or Ljubljana.

Working with such black-or-white contextual effects requires a strong rationale as to why the studied effect would stop at the contextual boundaries and precise knowledge regarding the drawing of these boundaries. The ideal case is when the aim is to determine the effect of a legal framework on social behaviour. In this case, the territory within which the law applies is easily identifiable by boundaries of states or federal entities, and the argument that the law is not expected to regulate behaviour in territories where it does not apply is straightforward. However, when it comes to war experiences (as well as, most plausibly, many other social phenomena), we look in vain for similar clear-cut boundaries at which the impact would stop abruptly. In practice, current multi-level analyses often involve choosing (convenient) definitions of contextual units and hoping that the shape of the actual contextual influences does not differ significantly from the shape of arbitrarily defined contextual units. In particular, national boundaries tend to be treated as a 'natural' (or unreflected) context for observing manifold collective processes. This practice creates two types of methodological problems. First, there is often a mismatch between the scale at which the contextual processes are conceptually explained and the scale at which they are empirically observed. Second, even when using an appropriate scale, the black-or-white logic of national (or any other) discrete contextual units artificially removes any contextual influences of events that happen to fall 'just on the other side' of the more-or-less arbitrarily defined boundaries from the analysis.

Among geographers, both types of pitfalls have been well known since Openshaws' (1984) seminal description of the *modifiable areal unit problem*, which encompasses scale effects (i.e., the strength and direction of ecological correlations depend on the level of aggregation) and zoning effects (i.e., even when the average size of contextual units is held constant, the precise drawing of the contextual boundaries can substantially affect relationships between aggregate indicators) (see also Fotheringham & Wong, 1991). In a recent study, Flowerdrew, Manley and Sabel (2008) generated 50 alternative ways of redrawing the boundaries of 21 wards within the English district of Swindon based on various geographic or social criteria. These authors then recalculated the ecological correlation between the empirical rate of long-term limiting illness and various sociodemographic predictors for each of the resulting 50 'pseudo-ward' systems. The findings impressively support the authors' conclusion that "it does matter where you draw the boundaries" (p. 1252). For example, the correlation coefficients between the rates of owner-occupied housing and long-term illness ranged from 0.22 to 0.86 across the 50 generated pseudo-ward systems, whereas the corresponding value for the official wards was only 0.19. The correlation between the rate of children under five and long-term illness ranged from -0.19 to -0.56 across the pseudo-wards, whereas the correlation with the official wards was null (0.02). These examples clearly demonstrate how the reliance on discrete contextual units with non-theoretically defined boundaries provides only partial, selective, and typically conservative insights into contextual effects.

In addition to this core incompatibility, two more technical inconsistencies between the nature of the data available to us to describe collective war experiences and the statistical assumptions underlying classic multilevel regression analysis must be mentioned. First, multilevel regression analysis assumes that higher-level units of analysis are randomly drawn from a specified population (see, e.g., Snjijders & Bosker, 1999). However, when geographically defined survey strata cover a complete territory, they do not represent a random sample. In the application that we will present in sections 4 and 5, the defined survey areas covered the entire territory of the former Yugoslavia and nothing beyond. Thus, they are better thought of as the full 'population' of contexts or, using the terminology that we will introduce in the next section, a consistent system of contextual units.

Second, the fact that in classic multilevel analysis, contextual values must be computed before the actual modelling begins implies that the model is blind to previous sources of error. When contextual values represent aggregations of micro-level measurements, one would clearly be better off trusting a model estimated using contextual values aggregated from many observations than a model aggregated from a few observations. However, if the point-value of the aggregate contextual value is the same, the final outcomes will be completely identical for context data based on a survey carried out among 10 or 10'000 individuals. Strictly speaking, the aggregate contextual values will be treated as if they were obtained from a full census in all cases, leaving conscientious data analysts with no alternative but to find data sources that comply closely enough with this ideal of precise contextual values or to refrain from using context data. Having an average of 50 observations available within each contextual unit to compute the estimates of collective war exposure, we did not have the first option at hand, but we did not want to resign ourselves to the second either.

3. Formalising spatially weighted context data (for multilevel analysis)

Rather than a fixed set of techniques, analyses using spatially weighted context data are better understood as a new approach to contextualising data analyses in the social sciences. Spatially weighted context data presuppose that the relevant social context is defined as a *consistent system* of contextual units (rather than as a random sample of permutable units), implying that the *relationships between the contextual units* are part of the model as much as the attributes of single contexts. This conceptualisation of the social context as an interdependent system is operationalised by calculating a *spatially weighted mean,* v_j^w, for each contextual unit *j* based on the corresponding observed values, v_k, for all C contextual units *k* in the system and a matrix of proximity weights w_{jk} between each possible binomial of contextual units *j* and *k* (Equation 1).

$$v_j^w = \frac{\sum_{k=1}^{C} v_k \times w_{jk}}{\sum_{k=1}^{C} w_{jk}} \qquad \text{(Equation 1)}$$

The relationship between the proximity weights and measured distances d between the contextual units is specified by a kernel function. In this way, a qualitative threshold *h* is introduced into a continuous distance function, which allows for the parameterisation of the scale of the relevant context. This threshold corresponds to the *bandwidth value* used in geographically weighted regression analysis, which similarly parameterises the scale (Fotheringham, Brunsdon & Charlton, 2002).

Equation 2 provides an example of a kernel function that has a set of desirable properties. As shown in Figure 1, this function allows for the generation of proximity weights w_{jk} between two contextual units j and k, which are typically two geographic regions that are part of the territory formed by all C regions. A value of 1 indicates that the distance d_{jk} between the two units is zero (the distance of a unit with itself), while the value tends towards zero when the distance tends to be infinite and tends towards 1/2 when d_{jk} tends towards h, which simultaneously constitutes the point of inflection and the steepest decline of the proximity function.

$$w_{jk} = \left(\frac{1}{2}\right)^{\frac{d_{jk}^{\,2}}{h^2}} \qquad \text{(Equation 2)}$$

It is important to note that spatially weighted context data defined in this way are applicable to both geographic and non-geographic conceptions of space. There is no need to confine the definition of 'distance' to physical distances expressed in kilometres. The approach can be flexibly adapted to a broad set of continuous distance functions to quantify how far, dissimilar, or disconnected the contextual units are from each other[ii].

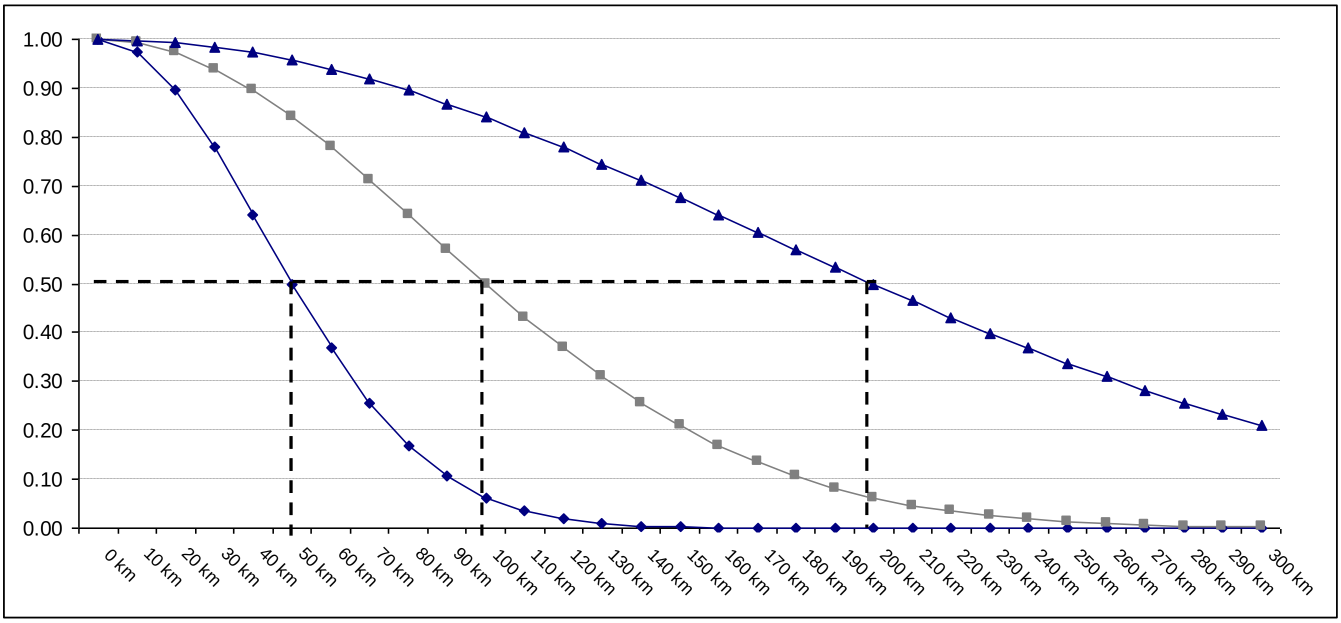


Figure 1: Kernel functions for spatial proximity weights based on geographic distance for three different scale values: h=200 km (top), h= 100 (middle), and h=50 km (bottom)

3.1 Multilevel modelling using spatially weighted context data

In the previous section, we argued that the primary advantage of using multilevel regression, compared to simple ecologic regression, as a tool to model context effects is that this method allows for the disentanglement of composition effects and 'true' (or net) contextual effects. Concretely, in multilevel analyses, fixed terms in an individual regression equation can be decomposed into fixed and random parts at the contextual level. Equation 3 presents the individual-level regression for an individual, *i*, embedded in a contextual unit, *j* (e.g., the region where i lives). The outcome variable Y_{ij} is expressed as a function of a set of n individual-level predictor variables, X_{ij}, context-specific estimates of the intercept and slope estimates, β_j, and an individual residual term, e_{ij}.

$$Y_{ij} = \beta_{0j} + \beta_{1j} X_{1ij} + \ldots + \beta_{nj} X_{nij} + e_{ij} \qquad \text{(Equation 3)}$$

In the present paper, we focus on the main effects of contextual variables using models in which only the intercept β_{0j} is decomposed. The vector of Equations 4.1.0 to 4.1.n shows the simplest case, in which the intercept is decomposed into one constant, γ_{00}, and one contextual-level residual term, u_{0j}, while slope parameters β_{1j} to β_{nj} are set equal each to a different constant (γ_{10} to γ_{n0}). Thus, unlike slope coefficients, the intercept value is allowed to vary across contexts, and the first interest generally lies in estimating the variance of u_{0j}.

$$\beta_{0j} = \gamma_{00} + u_{0j} \qquad \text{(Equation 4.1.0)}$$

$$\beta_{1j} = \gamma_{10} \qquad \text{(Equation 4.1.1)}$$

…

$$\beta_{nj} = \gamma_{n0} \qquad \text{(Equation 4.1.n)}$$

However, in general, the interest lies not only in observing whether there is some contextual-level variance but also in explaining this variance using specific contextual predictors. Accordingly, Equation 4.2 adapts the previous equation to the case in which one contextual-level predictor variable, Z_j, is introduced into the fixed part of the intercept decomposition.

$$\beta_{0j} = \gamma_{00} + \gamma_{01} Z_j + u_{0j} \qquad \text{(Equation 4.2)}$$

Equation 4.3 introduces the spatially weighted variant to the classic contextual-level equation. The contextual-level predictor term is now multiplied by the matrix W^g of weights w_{jk} divided by the sums of weights for all contextual units j ($\sum_{k=1}^{C} w_{jk}$), which specifies the relationship between j and each of the C contextual units *k* in the system according to Equation 2. The subscript *g* of W indicates that these weights are defined geographically, i.e., that d_{jk} corresponds to the geographic distance between *j* and *k* (e.g., to the distance in kilometres between the geographic centres of the two corresponding regions).

$$\beta_{0j} = \gamma_{00} + \gamma_{01} W^g Z_j + u_{0j} \qquad \text{(Equation 4.3)}$$

It is important to stress that even though the values of d_{jk} are assumed to be known and fixed, the values of W^g, and thus the estimation of γ_{01} (and γ_{00}), depend on the value of *h* in Equation 2. In this manner, scale is introduced into the multilevel regression model as an additional parameter, and the spatially weighted matrix allows for the estimation of the context effects as a function of scale.

However, we anticipate that once the scale effects have been potentially detected, geography alone will generally not be sufficient to explain what makes the events that occur within a certain radius different from those occurring further away. Thus, at some point, it will generally prove worthwhile to shift from geographic definitions of space to social definitions. Equation 4.4 presents the more general case, where social distances d_{jk} between contexts j and k are defined independently of the geographic distances. This small change is made visible by substituting the subscript *g* of *W* with *s* to indicate that socially defined distance matrices are used to weight the contextual values[iii].

$$\beta_{0j} = \gamma_{00} + \gamma_{01} W^{s} \ Z_{j} + u_{0j} \qquad \text{(Equation 4.4)}$$

Finally, to verify the extent to which the underlying social proximity between the contextual units actually mediates the geographic patterns, it can be interesting to estimate simultaneously the respective net contributions of geographically and non-geographically weighted indicators based on the same aggregate values, using Equation 4.5.

$$\beta_{0j} = \gamma_{00} + \gamma_{01} W_{g}^{g} Z_{j} + \gamma_{02} W_{s}^{s} Z_{j} + u_{0j} \qquad \text{(Equation 4.5)}$$

3.2 How much spatial dependency is explained?

A final issue to address is the extent to which different models using spatially weighted context data are actually able to explain the observed spatial dependency in the data. Geographers routinely quantify the level of spatial auto-correlation by computing a Moran's I coefficient. Equation 5 applies the Moran's I definition to spatial auto-correlation in the contextual-level residuals u_0. In this case, *j,k* represent all of the possible binomials of contextual units *j* and *k* in the system, and $\overline{u_0}$ is the average residual across all contextual units.

$$I_{u_0} = \frac{N}{\sum_j \sum_k w_{jk}} \frac{\sum_j \sum_k w_{jk} \left(u_{0j} - \overline{u_0}\right)\left(u_{0k} - \overline{u_0}\right)}{\sum_j \left(u_{0k} - \overline{u_0}\right)^2} \qquad \text{(Equation 5)}$$

As shown by this equation, the calculation of Moran's I requires the use of a full matrix of weights w_{jk} between contextual units *i* and *k*. Although defining such proximity weights dichotomously (e.g., distinguishing between 'neighbours' and 'non-neighbours') is the current practice, nothing precludes the use of continuously defined proximity matrices. Consequently, the integration of a spatial Kernel function (following Equation 2) into the estimate of spatial auto-correlation (following Equation 5) provides an interesting opportunity to parameterise the scale of spatial dependency, similar to the way in which the scale of contextual effects has been parameterised thus far. The resulting outcomes can be displayed in the form of a *spatial variogram*, in which the spatial dependency is represented as a decreasing function of the scale *h* at which two contextual units *j* and *k* are assumed to be close 'neighbours' (an example will be provided in section 5).

To facilitate the replication of analyses using spatially weighted context data and the application of these analyses to other fields, a specially designed *R-package* called *Spacom* (Junge, Penic & Elcheroth, 2012) integrates all of the computations presented in formulas 1 to 5 into a sequence of user-friendly commands and applies them to sample data (including all of the analyses presented in the two following sections). This package further integrates the ad hoc bootstrap resampling approach to the computation of confidence intervals, which will be presented below as an alternative to the standard errors available in current multilevel software. We can already anticipate that the bootstrap estimates will replace the described problematic statistical assumptions and are applicable to data that display non-independence between contextual measures, a full population of contexts, and non-precise aggregate contextual estimates.

4. Generation of spatially weighted context data in practice: describing collective war experiences across the former Yugoslavia

In this section, we will present the application of the spatial weighting technique introduced in section 3 to actual survey data on war experiences across the former Yugoslavia. Although this section will focus on a descriptive analysis using spatially weighted context data, it is important to keep in mind that an ultimate goal will be to separate the contextual and composition effects of war exposure. This objective requires the use of micro-level data about individual war experiences to estimate the contribution of composition effects to the aggregate outcomes. While the present section focuses on how micro-level data that are appropriately tied to relevant contexts can be collected and used to generate spatially weighted context data, in the next section, we will show how these context data can be combined with individual-level variables to separate composition and context effects created by war exposure.

4.1 Survey design

Between the spring and summer of 2006, more than six thousand adult respondents took part in the Transition to Adulthood and Collective Experiences Survey (TRACES), which was stratified into 80 continuous geographic areas. Two parallel surveys were conducted with fixed target sample sizes within each of these 80 areas. The first survey was conducted among a probability sample of the general adult population born in 1981 or earlier to collect micro-level data on life events, particularly war experiences ($\bar{n}$=50, N=3,975). The second survey was conducted among a probability sample of the 1968-74 birth cohort ($\bar{n}$ =28, N=2,254) to provide information about worldviews, and their individual-level correlates, among people whose formative experiences of early adult life (Schumann & Scott, 1989) occurred during the war years. While the respondents from the general adult sample were only administered a life events calendar, members of the cohort sample answered a longer attitudinal questionnaire in addition to the life calendar[iv].

To stratify the sample into 80 geographical areas, the following guidelines were followed:

- Areas are defined as geographically contiguous clusters of municipalities (i.e., which respect the boundaries of the municipalities in 2006).

- Areas are regional subdivisions within existing state boundaries and boundaries of different federal entities (e.g., Federation of Bosnia and Herzegovina vs. Republika Srpska).

- When intermediate levels of political subdivisions exist (e.g., counties in Croatia), areas correspond in principle to either one unit defined by this subdivision or a geographically contiguous cluster of these units.

- Regions that are highly heterogeneous with regard to geographical or historical factors, especially factors that affected the fate of inhabitants during the wars of the 1990s, were not grouped together into one area.

- Six urban areas are defined by the boundaries of the major cities: Belgrade, Ljubljana, Pristina, Sarajevo, Skopje, and Zagreb.

- Smaller political entities were over-sampled compared to larger political entities, i.e., the average numbers of inhabitants in areas within smaller countries are less than those within large countries.

- Regions populated mainly by major ethnic groups that are less numerous within the former Yugoslavia (Albanians, Bosniaks, Macedonians, and Slovenes) were over-sampled compared to regions populated mainly by the two most numerous ethnic groups (Croats and Serbs).

- Finally, to the extent that compliance with all of these guidelines still left room for different solutions, precise area boundaries were drawn such that the differences in the population sizes across areas that were equivalent with regard to the last two criteria were minimised.

Next, within each survey area, 15 sampling points were selected using a multi-stage cluster design. In most cases, a three-stage procedure for selecting the sampling points was used, involving the random selection of (1) municipalities within areas, (2) settlements within municipalities, and (3) sampling points within settlements. Finally, a random walk technique was applied to randomly select households, and a single eligible respondent was randomly chosen from within each selected household.

The described guidelines helped us to establish comparably defined areas across the entire territory and to approach an optimally stratified sampling design. As illustrated in Figure 2, areas with higher population density or higher expected variability in the demographic composition or war experiences (i.e., urban areas, border areas, ethnically ‘mixed’ areas, or areas where the overall minority population constituted a local majority) were over-sampled by defining smaller surfaces for these areas.

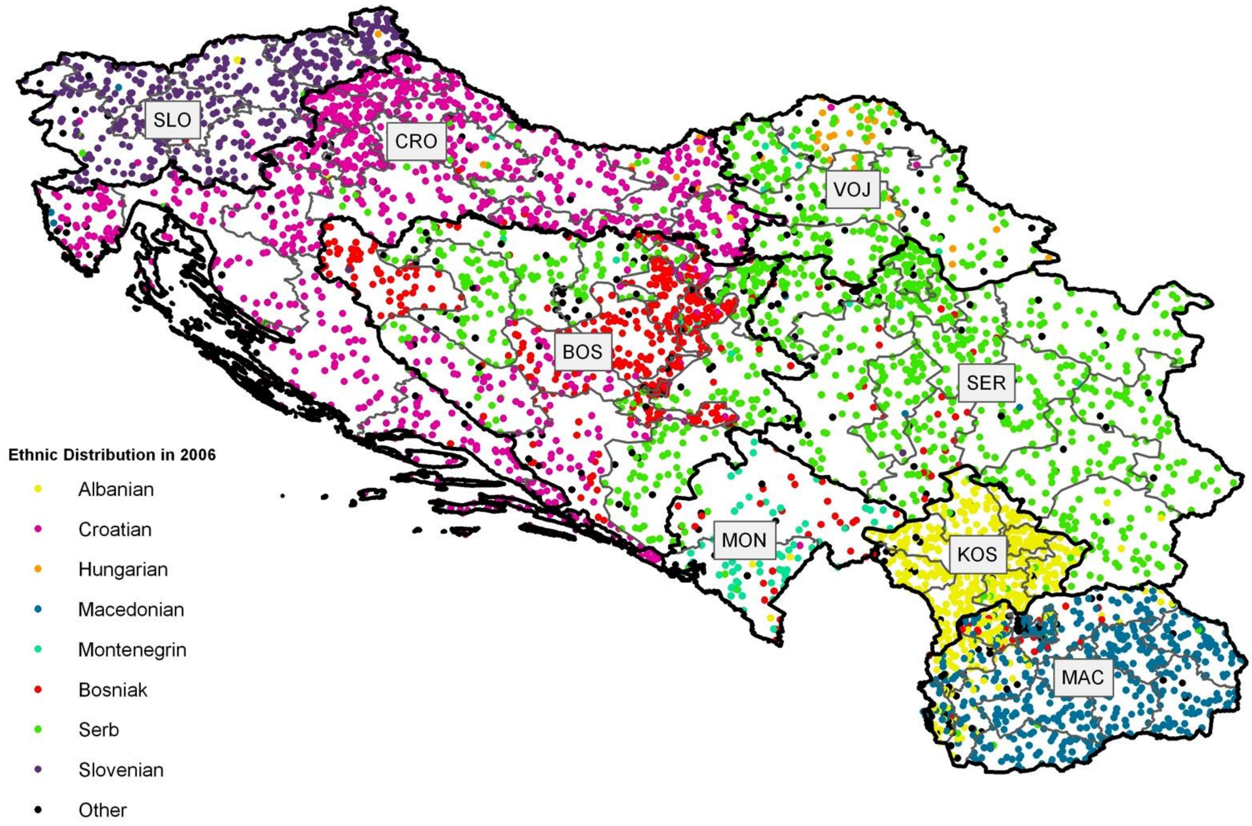
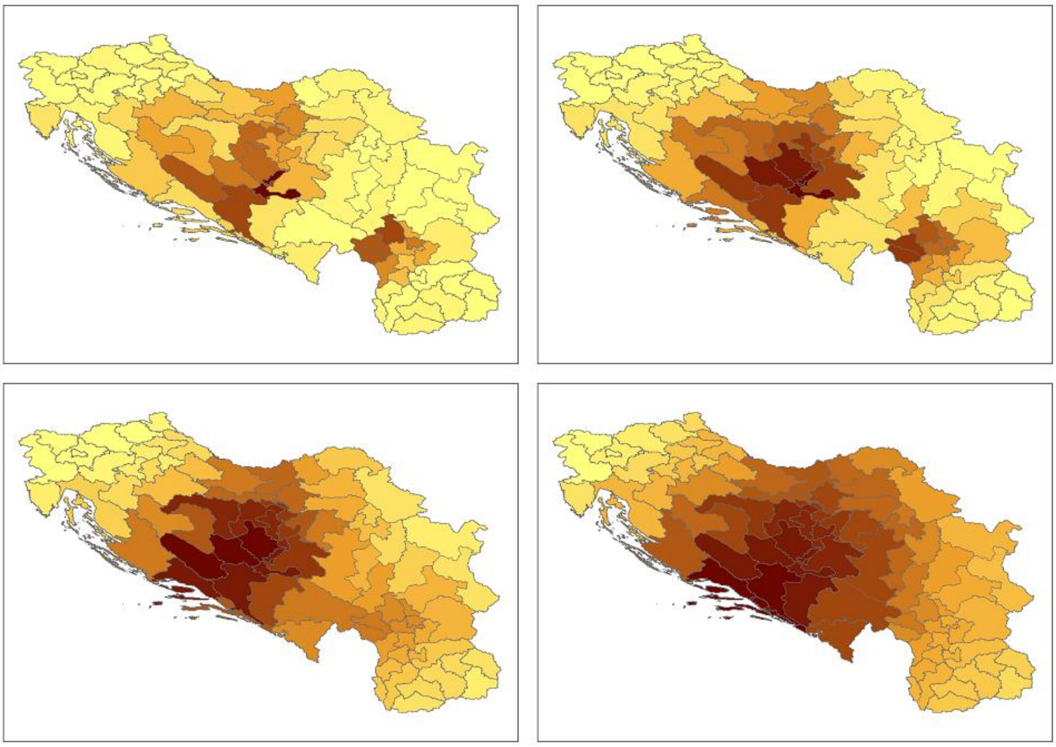


Figure 2: Sample composition of the 2006 TRACES survey as a function of geographic location and ethnic categories (dots represent individual respondents, lines represent boundaries of geographic survey strata)

Another important element of the survey design concerns the way in which past war events were recorded and associated with geographic locations. This task was complicated by the fact that most of the relevant events occurred more than a decade before the survey, and a substantial number of respondents had changed their geographic location, in many cases because of the war. The life events calendars allowed us to address both challenges[v]. The chronological years and seasons, birthdays, collective and personal marker events, and residential trajectories served as multiple temporal anchors. The recording of past locations allowed for the association of events with locations at the time of the events instead of the frequently inaccurate locations at the time of the survey (see Appendix 1). Using this method, an observation-based database was constructed, with individuals in three-month periods as cases and life events and geographic locations as variables[vi].

4.2 Spatialising descriptive analyses

All of the indicators of collective war experiences presented in this paper are based on aggregated responses to six survey questions on war events among the general adult sample. The respondents indicated whether they were forced to leave their home, had been captured, lost a close relative, had their property seriously damaged, were wounded, or had their house looted as a consequence of war (precise wordings are provided in Appendix 2). When at least one of these events happened to one respondent during a three-month period, the corresponding case in the observational database was coded as one. The case was coded as zero if none of the six events happened to the individual during the period. Based on this dichotomous event variable, an initial aggregate indicator was obtained by dividing the design-weighted[vii] estimate of the number of observations for a war event within each geographic area by the (design-weighted) total number of observations within the same area (i.e., the sum of the three-month periods that all of the respondents spent within this area). The relative magnitudes of the 80 resulting area-level values are projected onto a map of the former Yugoslavia in the upper left quadrant of Figure 3[viii]. This map allows for a visual interpretation of the main theatres of war (shaded dark brown) in Southern and Central Bosnia, the area of the Croatian boundary with Serbia in the North, and the Northwestern regions of Kosovo.



<u>Figure 3</u>: War experiences aggregated within the boundaries of survey areas (top left) or spatially weighted for geographic proximity across three different scales: 50 km (top right), 100 km (bottom left), and 200 km (bottom right)

In the next step, these primary contextual values, v_i, were spatially weighted according to Equations 1 and 2. The outcomes depicted in the three remaining maps of Figure 3 are all based on the geographic distance matrix, where d_{jk} is defined as the distance in kilometres between the geographic centres of areas *j* and *k*. The map in the upper right quadrant was obtained with a scale parameter of h=50 km, which roughly indicates that events occurring within a 50-km radius of the area's centre contribute substantially to the spatially weighted indicator, while at distances significantly beyond 50 km, the influence of events tends to become negligible (see Figure 1). This map displays a geographically smoothed version of the initial indicator, in which the overall patterns are emphasised and local idiosyncrasies, due to either true contextual peculiarities or random measurement errors, are reduced. Modelled in this manner, the collective impact of the war now appears to be structured concentrically around two separate and clearly identifiable main theatres in Central Bosnia and Northwestern Kosovo. The lower part of Figure 3 shows how this structure evolves if the value of the scale parameter is increased to h=100 km or h=200 km The dual-concentric structure then appears to evolve towards a mono-concentric structure, which is very strong in the latter case.

How precise are these descriptive estimates? To compute univariate confidence intervals that accurately reflect the impact of the relevant sample sizes on the precision of the estimates, we simulated the corresponding sampling distributions using a stratified bootstrap resampling procedure. Within each of the 80 areas, a number of (individual-by-time) observations that were equivalent to the initial number of observations within the corresponding area were randomly drawn with replacement from these observations. The initial aggregate indicators within each area were then recalculated, and spatially weighted indicators were computed as previously described. This procedure was repeated 1'000 times to generate a bootstrap distribution of the estimate. Figure 4 displays the 2.5th, 50th and 97.5th percentiles of the resulting distributions in each area, which serve as robust (median) point estimates and boundaries of the corresponding 95% confidence intervals. The comparison between the initial and spatially weighted estimates for three different values of h reveals how the estimates vary between two extremes as a function of scale. The unweighted estimates appear very imprecise (the confidence intervals are very large) and are partially erratic. This finding is not surprising given that these estimates are based on events reported by only 50 respondents. In contrast, when the estimates are weighted on a very large scale (200 km), they become very precise and rather uniform. Again, this finding is not surprising. When all of the spatial weights tend towards 1, i.e., when all of the events across the entire system tend to influence all of the estimates in an equivalent manner, all of the estimates simply converge towards the estimate of one overall mean based on the events reported by almost 4'000 respondents.

Good contextual indicators should avoid both maintaining high random error and discarding genuine variability between contexts. The two intermediate cases represented in Figure 4 appear to provide attractive compromise solutions with regards to both requirements. The estimates resulting from spatial weights at h=50 km are particularly interesting. They show that even if the indicators focus on events located within a relatively small radius, the precision of the resulting estimates can already be dramatically enhanced while preserving the variability between the contexts. Overall, the spatial weights appear to 'normalise' the values of a few outliers more than they make all values uniform. Consequently, these findings suggest that although a sample size of approximately 4'000 is insufficient for describing localised patterns across 80 individual regions with precision (as revealed by the large confidence intervals in the upper left quadrant), in combination with appropriate spatial weights, the same sample size can be sufficient for reliably interpreting 'smoothed' geographic patterns, such as the one depicted in the upper right quadrant of Figure 3.

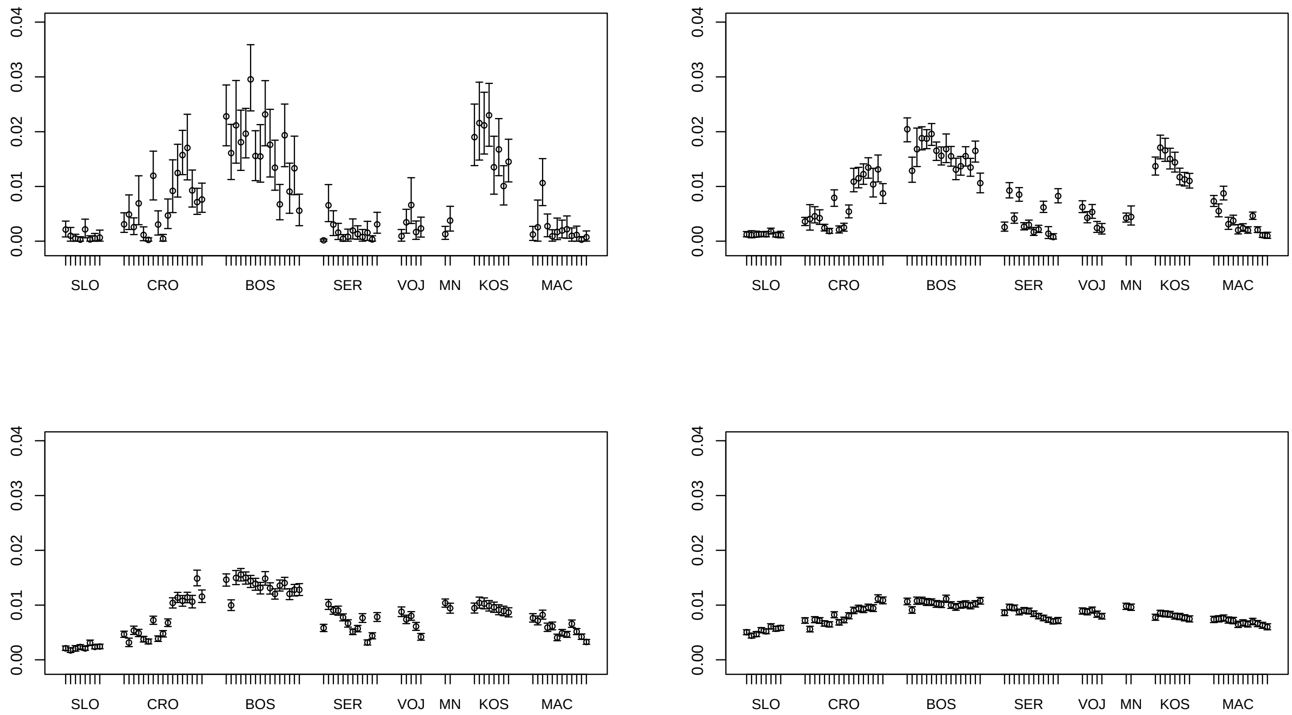


Figure 4: Boundaries of descriptive 95% confidence intervals for contextual indicators of collective war experiences, estimated by bootstrap resampling: unweighted (top left) and spatially weighted for geographic proximity at 50 km (top right), 100 km (bottom left), and 200 km (bottom right)

In this regard, the spatial weighting of context data can have consequences similar to those achieved using (Bayesian) shrinkage corrections in multilevel analyses, discussed in section 2, which weight the specific contextual information by the shared contextual information to a degree that is inversely related to the precision of the specific information. The difference is that with the existing shrinkage procedures, the precision of the estimates is the priority, and the resulting changes in the (implicit) specification of the relationships between the contextual units only constitute a possible and uncontrolled side effect. However, with spatially weighted context data, the priority is the substantially valid specification of relationships between contextual units, and repercussions on the precision of the estimates constitute a probable secondary advantage.

4.3 Socialising spatial descriptions

The next step consists of generalising the logic of spatially weighted context data to non-geographic definitions of space. Figure 5 illustrates the sensitivity of the identification of spatial patterns to the definition of 'distance'. Each of the three new maps that are compared to the geographically weighted estimate of the impact of collective war experiences (at h=50 km) are based on the *same* initial aggregate values v_j and the *same* formulas for weighting these values (Equations 1 and 2) but employ three *different* ways of defining the distance d_{jk} between contextual units *j* and *k*, resulting in three different distance matrixes. Each of these definitions exemplifies one particular approach to equating 'closeness' with a particular facet of social interdependence: similarity in the composition of the population in the first example, a sense of common identification in the second example, and actual contact opportunities provided by migratory flow in the third example.

The first approach defines *distance* as *dissimilarity* in the (ethnic) composition of a population. Following Equation 6, the distance d_{jk} between two areas *j* and *k* is defined as the sum of the absolute difference between the estimated population rates of self-declared members of ethnic group *g* within area *j* ($\hat{r}_{jg}$) before the war (in 1991) and of the same group in area *k* ($\hat{r}_{kg}$).

$$d_{ij} = \sum_{g=1}^{6} \left| \hat{r}_{jg} - \hat{r}_{kg} \right| \quad \text{(Equation 6)}$$

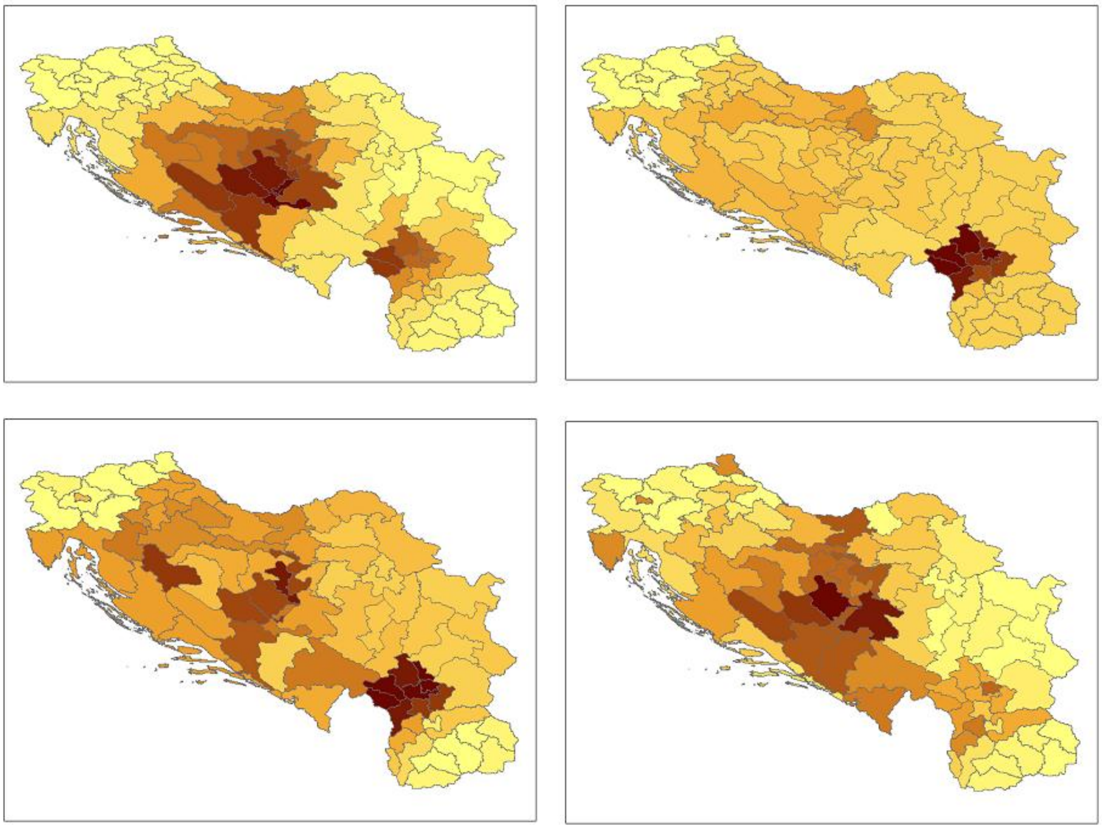


Figure 5: War experiences spatially weighted for geographic proximity (h=50 km, top left) and three types of non-geographic social proximity: territorial identification (top right), ethnic composition (bottom left) and migration flows (bottom right)

In this example, *h* was defined as the maximal distance possible between two areas with the same majority group (i.e., *h*=1). As shown in the lower left quadrant of Figure 5, such a definition of spatial proximity results in the representation of war experiences as bounded within territories that were inhabited mainly by members of those ethnic communities that were the most victimised by war: Albanians in Kosovo, Bosniaks in Central Bosnia, and Croats across various areas of Croatia and Herzegovina.

The second approach defines *distance* as *a lack of common identification*. To avoid anachronistic models, we need to assess the identification among the population before, rather than after, the relevant war period (which was also a period of rapid transformation of collective identities; see Elcheroth & Spini, 2010). Because subjective identification cannot be measured reliably in a retrospective study, we relied on historical survey data to construct this indicator. For the Yugoslav Public Opinion Studies of 1990, a representative sample of Yugoslav residents were asked to what extent they identified with their republic or province and with Yugoslavia as a whole. Both types of identification were recorded independently on a five-point Likert scale, where 1 meant 'very important' and 5 meant 'unimportant'.

In Equation 7.1, y_j and y_k represent the estimated weaknesses in Yugoslav identification in areas *j* and *k*, respectively, whereas r_j and r_k represent estimated weaknesses in republican (or provincial) identification. Thus, the distance between *j* and *k* is defined as the smallest value between the square roots of the between-area products of both types of territorial identifications.

$$d_{ij} = MIN\left(\sqrt{y_j \times y_k}, \sqrt{r_j \times r_k}\right) \text{ (Equation 7.1)}$$

For two areas that were *not* part of the same republic or province, only identification with Yugoslavia as a whole could be the basis of a common territorial identification, and the equation is simplified as follows:

$$d_{jk} = \sqrt{y_j \times y_k} \qquad \text{(Equation 7.2)}$$

In this case, the threshold value was defined as the distance between two areas sharing a common identity considered, on average, to be 'rather important' by both populations (i.e., *h*=2). This spatial weighting scheme results in a representation of the collective war experiences in which the differences between territories where the populations strongly identified with Yugoslavia as a whole on the eve of war are strongly smoothed. At the same time, it emphasises the unique positions of Slovenia and Kosovo, where identification with the republic or province was predominant in 1990. In the former case, the area appears to have been specifically spared from war experiences, and in the latter case, the area appears to have been specifically affected (see upper left quadrant of Figure 5).

The third approach defines distance as *the lack of contact* between two populations. According to this perspective, distance is reduced by migratory flows from one area to another. The recording of residential trajectories for the TRACES provides the opportunity to operationalise this definition of distance d_{jk} between areas *j* and *k* based on the number of reported moves from *j* to *k* ($n_{j \to k}$) and from *k* to *j* ($n_{k \to j}$). In Equation 8, $\hat{n}_{j \to k}$ represents the (design-weighted) estimated number of moves from area *j* to area *k* in the population between 1990 and 2006, and $\hat{n}_{k \to j}$ represents the estimated number of moves from *k* to *j*. N_j and N_k represent the population sizes[ix] of areas *j* and *k*, respectively. The second term of the product introduces a correction intended to enhance the robustness of the estimates by giving less weight to estimates based on a small number of observations. Given that the responses from 3'795 individuals served as the basis for estimating the moves between 3'160 possible area binomials, the small number of observations by cell were the rule rather than the exception. By correcting for small sample sizes, we intended to emphasise the proximities between those regions where a critical mass of moves was actually observed, i.e., to construct an indicator that is sensitive to systematic migratory flux and disentangles it from random patterns of individual dislocations.

$$d_{jk} = \frac{(N_j + N_k)/2}{\hat{n}_{j \to k} + \hat{n}_{k \to j}} \times \frac{4}{\ln(n_{j \to k} + n_{k \to j}) + 1} \quad \text{(Equation 8)}$$

When transforming these distances into spatial proximity coefficients, the scale parameter was defined as the distance between two areas related to the total number of estimated moves corresponding to 4% of the mean population across the two areas (i.e., *h*=25). The resulting representation partially reproduces the spatial pattern created by geographic distances, which can easily be interpreted: people are more likely to move to geographically close areas. However, the same representation also introduces a series of local deviations from the simple geographic pattern. For example, this representation highlights how the Kosovo war 'spilled over' into neighbouring areas in Macedonia and how the war was brought to Slovenian areas that accommodated many refugees during the war, especially in the urban centre of Ljubljana. Thus, this indication can be an interesting way to model how war can shape collective experiences, even among populations that did not experience (major) combat in their own territory.

5. Multilevel modelling using spatially weighted context data: the impact of collective war experiences on collective worldviews

In this section, we aim to show how spatially weighted context data can be incorporated into multilevel regression analyses to model the impact of collective experiences on collective worldviews, how they allow for the parameterisation of scale effects in these models, and how hypotheses about social mechanisms that create spatial structures can be empirically tested. Multilevel analysis allows for the combination of context data and individual data if the affiliation of individuals with contextual units can be established. In the example presented in this section, the (spatially weighted) context data about the collective war experiences described in the previous section are combined with individual-level data from the survey conducted among one cohort. Each individual respondent from the cohort study was attached to the area he or she lived in at the time of survey as his formative postwar context[x]. Because the sample in this second survey was geographically stratified by the same contextual units used for the general adult sample, which provided life events data to construct the contextual indicators, the combination of both data sources at the contextual level was straightforward[xi].

The present analyses focus on two related outcome variables: collective guilt assignment, i.e., the attribution of blame to other (ethnic) groups, and collective guilt acceptance, which, inversely, represents a critical view of the role of one's own group in past actions (Branscombe & Doosje, 2004). The strength of the individual endorsement of both types of worldviews was operationalised as the average of the individual responses on a seven-point Likert scale to a set of five survey items each (see Appendix 2 for the precise wordings).

The first two regression models shown in Table 1 represent an application of Equations 4.1 and 4.2 (in combination with Equation 3) to the explanation of collective guilt assignment ($GUILT\_ASS_{ij}$) as an outcome variable: the 'composition effects only' model (Equation 9.1) and the 'discrete contexts model' (Equation 9.2).

$$GUILT\_ASS_{ij} = \gamma_{00} + \lambda_{10} PERS\_WAR_{ij} + ... + \lambda_{n0} X_{nij} + u_{0j} + e_{ij} \quad \text{(Equation 9.1)}$$

In these models (as in all of the following models), the personal experience of war ($PERS\_WAR_{ij}$) is used as part of a set of n individual-level control variables, X_{ij}. As this dichotomous variable is based on the same six items as the micro-level measures that generated the aggregate contextual indicators, its inclusion in the model allows us to partial out the simple additive effects of personal experiences and to maintain only the additional impact of events experienced by others at the contextual level. Sex, age, level of education and combat experience are included as further individual-level controls to consider as much as possible the composition effects that do not stem from the aggregated variables themselves but from other confounding factors in the composition of the population that may vary across contextual units. While model 1 is limited to these composition effects, model 2 introduces, according to Equation 9.2, the initial (unweighted) aggregate indicator of collective experiences as a contextual predictor variable.

$$\begin{aligned} GUILT\_ASS_{ij} &= \gamma_{00} + \gamma_{01} COLL\_WAR_{ij} \\ &+ \lambda_{10} PERS\_WAR_{ij} + \ldots + \lambda_{n0} X_{nij} + u_{0j} + e_{ij} \end{aligned} \quad \text{(Equation 9.2)}$$

The standardised partial regression coefficient of the 'discrete context model' displays a small but significant positive effect, suggesting that members of the generation of young adults exposed to war are slightly more likely to collectively blame ethnic outgroups in areas where war events affected the population more often than in areas where fewer war events occurred.

The above application is only a crude initial test of the hypothesis that collective war experiences affect generational worldviews. It is based on a single indicator and, more problematically, on the collective experiences bound within very small, specific geographic areas. Figure 6 provides insight into the general picture. The partial regression coefficients in this figure are all generated by introducing a spatially weighted contextual-level predictor term, according to Equation 4.3, using the matrix of distances between the geographic centres of the contextual units, which leads to Equation 9.3, where $W^g\ COLL\_WAR_j$ represents the collective war exposure weighted by the geographic distances.

$$\begin{aligned} GUILT\_ASS_{ij} &= \gamma_{00} + \gamma_{01} W^g COLL\_WAR_j \\ &+ \lambda_{10} PERS\_WAR_{ij} + \ldots + \lambda_{n0} X_{nij} + u_{0j} + e_{ij} \end{aligned} \quad \text{(Equation 9.3)}$$

As mentioned before, spatially weighted context data can differ substantially depending on the definition of the scale parameter (*h*), even when the micro-level data, the aggregation process, and the distance matrices are held constant.

In the findings presented in Figure 6, this important property of spatially weighted context data is exploited to systematically explore the relationship between the scale and size of the contextual effects of interest, i.e., the net contextual effects of collective war experiences on both collective guilt assignment and collective guilt acceptance. In these analyses, the scale value was gradually increased by intervals of 25 km, while all of the other parameters were held constant.

Table 1: Multilevel predictors of collective guilt assignment: robust standardised regression coefficients and boundaries of 95% confidence intervals (median value, 2.5 and 97.5 percentiles of stratified resampling distribution)

(*): Deviance reduction estimated with Full Maximum Likelihood. All other values in the table estimated with Restricted Maximum Likelihood. Models 2 to 6: df=1; Model 7: df=2.

	Model 1 "Composition effects"	Model 2 "Discrete contexts"	Model 3 "Geographic space"	Model 4 "Territorial identification"	Model 5 "Ethnic composition"	Model 6 "Migratory flows"	Model7 "Combined spaces"
Individual-level predictors							
Personal experience of war trauma	0.09	0.08	0.08	0.07	0.08	0.09	0.07
	(0.05 - 0.14)	(0.03 - 0.13)	(0.03 - 0.13)	(0.02 - 0.11)	(0.03 - 0.13)	(0.04 - 0.14)	(0.02 - 0.12)
Combatant	0.04	0.04	0.04	0.04	0.04	0.04	0.04
	(-0.01 - 0.09)	(-0.01 - 0.08)	(-0.01 - 0.08)	(0.00 - 0.09)	(-0.01 - 0.09)	(-0.01 - 0.09)	(0.00 - 0.09)
Male	0.04	0.04	0.04	0.04	0.04	0.04	0.04
	(0.00 - 0.08)	(0.00 - 0.08)	(0.00 - 0.08)	(0.00 - 0.08)	(0.00 - 0.08)	(0.00 - 0.08)	(0.00 - 0.08)
Age in 1990	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(-0.03 - 0.05)	(-0.03 - 0.05)	(-0.03 - 0.05)	(-0.04 - 0.04)	(-0.03 - 0.05)	(-0.03 - 0.05)	(-0.04 – 0.04)
Level of education							
- Secondary	0.03	0.03	0.03	0.04	0.04	0.03	0.04
	(-0.02 - 0.09)	(-0.02 - 0.09)	(-0.02 - 0.08)	(-0.01 - 0.09)	(-0.02 - 0.09)	(-0.02 - 0.09)	(-0.01 – 0.09)
- Tertiary	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
	(-0.06 - 0.04)	(-0.06 - 0.04)	(-0.06 - 0.04)	(-0.06 - 0.04)	(-0.06 - 0.04)	(-0.06 - 0.04)	(-0.06 – 0.04)
Contextual-level predictors							
Collective experience of war trauma							
- Unweighted (bounded within areas)	-	0.07	-	-	-	-	-
		(0.02 - 0.13)					
- Weighted by geographic space (h=50 km)	-	-	0.09	-	-	-	-0.04
			(0.03 - 0.14)				(-0.10 – 0.03)
- Weighted by territorial identification	-	-	-	0.21	-	-	0.23
				(0.17 - 0.25)			(0.17 - 0.27)
- Weighted by ethnic composition	-	-	-	-	0.10	-	-
					(0.05 - 0.15)		
- Weighted by migratory flows	-	-	-	-	-	0.02 (-0.04 - 0.07)	-
Deviance reduction (Ref: Model 1)*	-	2.12	2.72	18.26	3.83	0.17	19.04
	-	(0.15-5.72)	(0.44-6.91)	(11.50-26.54)	(1.00-8.38)	(0.00-1.61)	(12.64-26.93)
% of explained contextual variance (Ref: Model 1)		1.27	2.22	23.09	3.48	-1.28	26.14

To generate the robust point estimates and 95% confidence intervals in this figure (as well as in Table 1), we extended the stratified resampling procedure described in the previous section for the computation of univariate confidence intervals. Multilevel regression estimates rely on two different micro-level data sources within each contextual unit. Therefore, we not only generated 1'000 bootstrap resamples of the observations that generated the context data but also 1'000 bootstrap resamples of the individuals in the cohort sample, who provided the individual-level predictor and outcome variables. Each resample of the resulting context data was then randomly associated with one unique resample of the individual-level data. In this manner, each regression model could be reiterated across 1'000 matched bootstrap resamples to generate a distribution of 1'000 reiterations of the same partial regression coefficients (for other applications of bootstrap resampling procedures in multilevel analyses, see Van der Leeden, Meijer, & Busing, 2008; and Moran, 2006). These distributions are summarised in Figure 1 using three values for each regression coefficient: the median of the distribution, which serves as a robust point estimate, and the 2.5th and 97.5th percentiles of the same distribution, which constitute the boundaries of the associated 95% confidence interval. Thus, we were able to estimate the precision of each estimate by directly simulating the relevant sampling distribution instead of using the distributional assumptions that generate estimates of standard errors in standard multilevel applications, which (as argued in section 2) are incompatible with the present data structure.

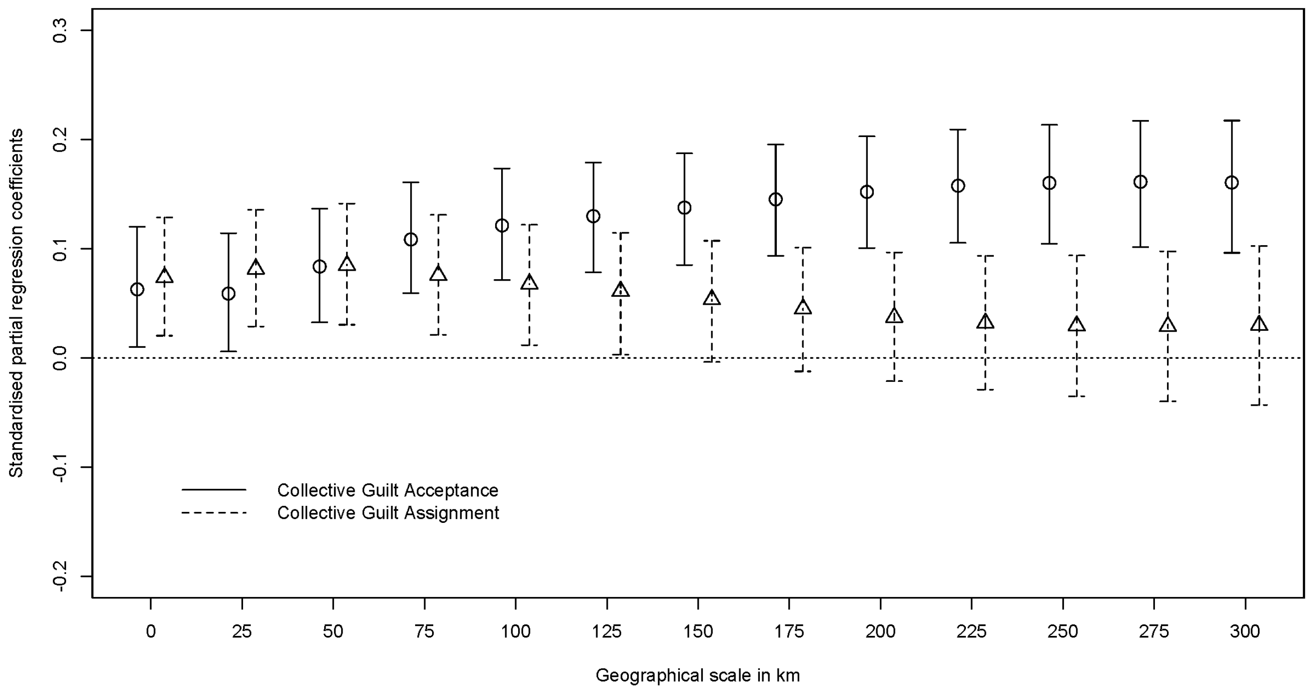


Figure 6: Partial regression coefficients for contextual effects of war experiences on collective guilt acceptance and assignment, spatially weighted for geographic proximity across scales ranging from 0 km (unweighted) to 300 km

The findings depicted in Figure 6 vividly illustrate how the size of the contextual effect of interest varies as a function of scale. They also demonstrate that the corresponding functions can have very different shapes and maximum values, even with two closely related outcome variables, such as collective guilt assignment and acceptance. The impact of collective war experiences on collective guilt assignment peaks at a radius of 50 km, diminishes beyond 50 km, and ceases to be significant beyond 100 km. However, the impact of the same collective experiences on collective guilt *acceptance* steadily increases and only reaches a plateau beyond 200 km. This finding suggests that although young adults' attributions of blame to outgroups are anchored in war events that were experienced by other people in a location that was relatively close, their critical assessments of their own group's role in wrongdoings are influenced by similar experiences that occurred in a space that was geographically much more extended. Any conventional contextual analysis approach would most likely have overlooked at least one of these two effects. A research design based on discrete contextual units at a relatively small regional level would have resulted in at least a significant underestimation of the effect of collective war experiences on collective guilt acceptance. In contrast, a design based on discrete contextual units at a larger regional level would have led to the conclusion that collective war experiences do not influence collective guilt assignment.

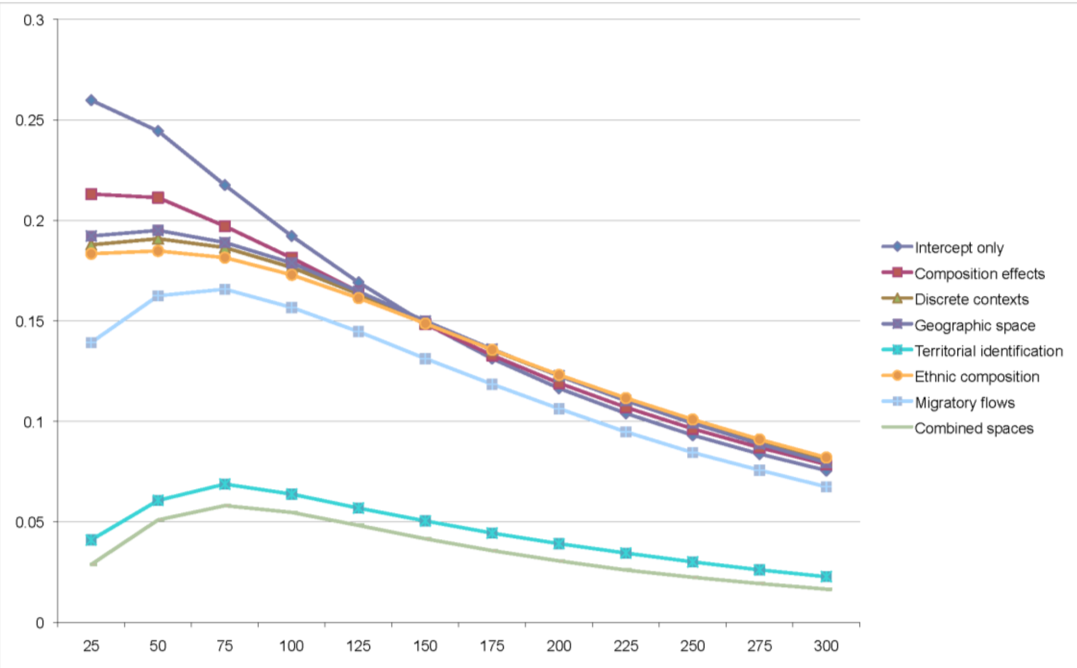
5.1 Spatial patterns and social interdependencies

What makes events that occur within a 50-km radius different from those that occur further away? To address this question, we need to shift from geographic definitions of space to more socially oriented definitions, according to Equation 4.4. Table 1 displays estimates of the contextual effects of the three types of non-geographically weighted indicators of collective war experiences introduced in the previous section (Models 4-6) in addition to the effect of the geographically weighted indicator at its maximal scale (h=50 km) on collective guilt assignment, as shown in Model 3. The strongest effect is observed for collective war experiences weighted by territorial identification. If both indicators (collective war experiences weighted by geographic space and by territorial identification) are entered simultaneously into the regression model, according to Equation 9.4, the effect of the socially weighted indicator appears to be equally strong, whereas the effect of the geographically weighted indicator becomes null.

$$GUILT\_ASS_{ij} = \gamma_{00} + \gamma_{01} W^{g} COLL\_WAR_{j} + \gamma_{02} W^{s(ident)} COLL\_WAR_{j} + \lambda_{10} PERS\_WAR_{ij} + \ldots + \lambda_{n0} X_{nij} + u_{0j} + e_{ij} \quad \text{(Equation 9.4)}$$

Thus, this evidence suggests that the apparent structuring effect of geographic proximity on a relatively small scale is mediated by an underlying social interdependency created by common identification. Thus, the most parsimonious way to describe the observed pattern is to state that a worldview in which other groups are to blame is anchored in traumatising war events that occurred within populations (once) tied together by a shared sense of affiliation. This conclusion is further corroborated by the amount by which the overall model deviance is reduced for each of the six contextual models (i.e., the difference between deviance associated with each of Models 2 to 7 and the deviance of the baseline Model 1): The model deviance is reduced much more for the two models using territorial identification as spatial weights (i.e., Models 4 & 7) than for all the remaining models. Collective war experiences weighted by territorial identification alone (Model 4) explain 23% of the variance across contextual units that remained unexplained by the composition-effects-only model (i.e., Model 1).

Finally, Figure 7 presents an example of a spatial variogram, which represents the residual spatial dependency (the level of spatial auto-correlation that is left unexplained by the model, according to Equation 5) for the various multilevel regression models reported in Table 1 as a function of geographic scale (from h=25 to h=300). The level of spatial dependency observed in the so-called 'intercept-only' or 'empty' model (i.e., which does not contain any explanatory variables) is added to provide the initial baseline value of the total spatial dependency in the data, or how much collective guilt assignment is spatially auto-correlated between the contextual units.



<u>Figure 7</u>: Spatial variogram, displaying residual spatial dependency as a function of decreasing geographic proximity, for different multi-level models explaining collective guilt assignment

The interpretation of the spatial variogram reveals three points. First, the observations are substantially spatially dependent across contextual units. The assumption of independence underlying statistical significance tests in standard (multilevel) regression analysis would have been violated here. This fact lends posterior validation to the assumed necessity of using alternative techniques for statistical inference, such as those proposed here or spatial autoregressive models. Second, these findings are in line with a basic law of geography (which also underlies the spatially weighted context data approach) that closer units are more likely to be correlated than those that are more distant. Third and most importantly, the findings highlight the vast differences across the tested models' capacities to appropriately account for the spatially dependent data structure. In particular, the model that only includes composition effects explains the spatial dependency to a very limited extent, suggesting that most spatial dependency must be accounted for by genuine context effects rather than by similarly composed populations in geographically close areas. The two models that include collective war experiences weighted by shared territorial identification stand out in their capacity to appropriately account for spatial dependency. Once this contextual predictor is included, the residual spatial dependency virtually disappears, particularly among geographically close areas. Thus, this observation converges with the interpretation of regression coefficients that common identification, rather than a similar ethnic population structure or concrete contacts created by migrations, appears to bind together singular war events into a collective war experience that affects collective worldviews (at least in terms of collective guilt assignment among a single generation).

6. Spatially weighted context data: potential, transfer, and boundary conditions

6.1 The rediscovery of social context

In the social sciences, increasing attention is being given to the fact that the individual as the sole unit of analysis for studying relevant patterns, dynamics and causal mechanisms is often insufficient. Ecological relationships are no longer conceived as simply 'fallacies'; they have benefited from increasing recognition and are now seen as interesting phenomena in themselves (see Schwartz, 1994; Chaix, Merlo & Chauvin, 2005). Conversely, awareness has been raised among social scientists of the equally misleading nature of *atomistic fallacies*, which means that conclusions about collective outcomes on the basis of individual-level relationships between observed variables are dubious (Diez-Roux, 1998; for an early detailed discussion of the issue, see Dogan & Rokkan, 1969).

Increasing interest in context effects has been paralleled, during the last two decades years, by increasing opportunities to study these effects empirically. Survey research appears to have gone through a 'comparative turn'. In Europe, new high-quality international survey programs, such as the *Comparative Studies of Electoral Systems* (since 1996), the *European Social Survey* (since 2002), and the surveys that are part of the *EU Statistics on Income and Living Conditions* program (since 2003) complement now the more classic surveys such as Eurobarometer or the European Values Survey. In Northern America, the federal structure of the United States or Canada implies that large federal surveys are often treated as comparative surveys in themselves.

Simultaneously, multilevel modelling techniques have been widely diffused in the social sciences. A new kind of routine analytical procedure appears to have emerged from these converging tendencies: cross-national survey data are commonly analysed by way of multilevel regression models, in which individual survey data provide micro-level predictors and outcome variables and national statistical indicators are used as macro-level predictors.

While this new routine is now conventionally treated as scientifically accepted practice, the fit between the data and assumptions underlying the method should always be cautiously examined. Although multilevel regression analysis assumes that higher-level units of analysis are randomly drawn from a specified population, it is generally unclear of which reference population countries in an international survey would constitute a random sample. Traditionally, multilevel regression models have further assumed statistical independence between contextual-level measures (but see Savitz & Raudenbush, 2009). This assumption may not be realistic for certain phenomena, like for example the diffusion of social behaviour facilitated by spatial mobility or social contacts across neighbouring countries or otherwise interdependent nation-states. Finally, when macro-level predictors used in cross-national analyses are survey-based estimates (for example, aggregations from national income or labour force surveys), a source of error is ignored in the model when they are statistically treated as if they were precise measures at the country level.

Beyond such 'technical' inconsistencies, one should particulary beware of relying to easily on nations or any other 'natural' units of analysis, even when explanatory mechanisms would suggest that infra- or trans-national dynamics cause the observed phenomena. For example, when cross-national correlations are used as indirect evidence for processes that are being theorised at a more proximal (or distal) level, such practice create the two types of methodological problems discussed in section 2. First, a mismatch between the scale at which contextual processes are conceptually explained and the scale at which they are empirically observed; second, even with an appropriate scale, the black-or-white logic of national (or any other) discrete contextual units artificially cuts off from the analysis any contextual influences of events than happen to fall 'just the other side' of more or less arbitrarily defined boundaries.

6.2 Spatial alternatives to multilevel analyses

A series of spatial solutions have been proposed to overcome the limitations related to discrete contextual units. The most prominent approaches propose one of the three following solutions: (1) let individual-level regression coefficients vary as a function of spatial coordinates, as is the case with *geographically weighted regression models* (see Brunsdon, Fotheringham & Charlton, 1998); (2) include values of the dependent variable for 'close' contexts as additional predictors, which is the logic of *diffusion or contagion models* (see, for example, Ward & Gleditsch, 2002); or (3) add a spatially dependant error term to classic or, more recently, multilevel regression analysis to generate so-called spatially autoregressive models (see, e.g., Savitz & Raudenbush, 2009).

These three approaches pursue variable goals and have allowed for genuine progress in the study of factors in such diverse outcomes as public health or housing prices (Brunsdon, Fotheringham & Charlton, 1998), violent conflict (Ward & Gleditsch, 2002), and crime rates (Savitz & Raudenbush, 2009). But none of them aimed to operationalise a substantive model based on collective experiences as explanatory mechanism for contextual influences. As a consequence, work presented in this paper started from the premise that the one avenue toward 'spatialising' social analyses that deserves more attention and effort than it has received so far, consists in incorporating spatial functions into the definition of contextual predictor variables.

6.3 Boundary conditions to spatially weighted context data

Different avenues can be imagined for establishing concrete analytic procedures that integrate spatially weighted context data, such as introduced here as an alternative to some of the shortcomings of classic multilevel analysis. However, the chosen procedures must comply with certain requirements. A first set of requirements concerns the definition of primary contextual units. To minimise zoning effects, these units must be defined significantly *below* the scale of the contextual effects of interest. Otherwise, distinct boundaries will still matter, and spatial weights will not redress the impact of arbitrarily defined boundaries. For example, if researchers are interested in the impact of collective experiences among people that share a territory that has more or less the size of a Swiss canton, the usage of cantons as primary contextual units would most likely create important zoning effects. These can be effectively avoided however by starting with observations at a municipality or district level, which are then spatially weighted at a scale that corresponds to the size of cantons. In practice, this requirement must be balanced against the constraint that sufficient information must be available to compute a complete matrix of distances between primary contextual units for one or several relevant dimensions. For many data sources, geographic coordinates will not be available on the level of precise measurement points or even neighbourhoods. However, in many cases, they may be available on the municipal or regional level. As far as non-geographic definitions of social distances are concerned, issues of applicability add to issues of availability. For example, if social distances are based on (dis-)similarities in certain policies, it would not be meaningful to define contextual units that cannot have policies on their own.

A second type of requirement arises when initial contextual measures are produced by the aggregation of micro-level survey data. In fact, while the use of aggregated survey data as contextual indicators represents already a frequent practice, it could even be further facilitated by the use of spatially weighted context data. As we have shown in sections 4 and 5, spatially weighted context data do not require that initial contextual values approach precise measures; consequently, even datasets with small to moderate sample sizes within contextual units can be appropriate for computing initial contextual values. However, it is clear that aggregate contextual indicators can only be unbiased to the extent that the micro-level measures that are aggregated are unbiased themselves. When context data are used as indicators of collective experience, they must therefore be based on accurate records of the micro-level events that produce the collective experience. The survey measures must have good construct validity and must accurately locate events within contextual units.

Furthermore, in the frequent case of past events that are assessed retrospectively, accurate recording of the timing of the events is generally presupposed, especially when there is significant spatial mobility of individuals across contextual units, such as circumstances in which the location of an individual response at the time of the survey cannot be used as a proxy for the individual's location at the time of the event. In section 4, we have therefore described a life calendars methods as a helpful device to increase the accuracy of recording past locations. Furthermore, surveys with sampling designs that are stratified by primary contextual units are more likely than simple random samples to provide sufficiently precise estimates of contextual effects when spatially weighted context data are used. Likewise, optimally stratified sampling designs (in which smaller spatial units are defined where there is more expected variability, as was the case with the sampling design described in section 4) are more likely to provide sufficiently precise estimates than are non-optimally stratified designs.

7. Conclusion: what do spatially weighted context data add?

Comparative analyses conducted using spatially weighted context data differ from classic multilevel regression analyses in three important ways. First, they rely on a set of contextual units defined as a consistent and interdependent social system instead of assuming the existence of a random sample of independent contexts. Second, they are conceived to study contextual influences that decrease with increasing distance rather than contextual influences that are bound within discrete and rather arbitrarily delimited contexts. Third, they parameterise the scale of contextual effects and thus enable the study of these effects as a function of scale, rather than constraining the scale of the studied effects in advance.

Analyses using spatially weighted context data also differ from the existing approaches to spatial analyses. In contrast to the prevailing spatial autoregressive paradigm, they do not attempt to 'partial out' spatial dependency from regression models to isolate the intrinsic effects of a set of (non-spatial) predictors. Instead, we propose a strategy that retains spatial dependency in the model, spatialises the contextual predictors to explain it substantially, and then determines the extent to which these spatialised predictors can account for the observed spatial dependency. Contrary to other current spatial applications, analyses with spatially weighted context data do not conceive of the identification of geographic spatial patterns as a goal in itself. Rather, it is a heuristic device to explain the structure of social interdependencies through which a set of single experiences organises into qualitatively distinct, consequential collective experiences. In section 5, we used concrete examples to show how geographically defined context data can help to detect new patterns, which can subsequently be explained from a sociological perspective. Therefore, spatially weighted context data should be viewed as a tool that can be flexibly adapted to various social definitions of ‘distance’, leaving room for theoretical imagination and facilitating the implementation of theory-driven definitions of context.

While there is no finite set of solutions for how to define the distance between two places or populations socially (what is a ‘good’ definition of social distance necessarily depends on the specific research agendas and studied social phenomena), three avenues appear promising to us. First, the findings reported in section 5 highlight that the spaces with which people identify are critical. This connects spatial analyses to theories of social identity and opens interesting insights into the social-psychological mechanisms through which concrete experiences are collectively remembered and interpreted. Second, common spaces are not only created by common imagination but also by common practices. The mobility and migration studies that generate data that are more fine-grained than those we utilised in analysing the former Yugoslavia might in other contexts allow for sufficiently precise representations of the extent to which spatial practices of people living in different places overlap. Third, people can also experience spaces by proxy and become close to places inhabited by relevant others. Network studies, which record geographical locations of socially related individuals, can therefore provide interesting opportunities to quantify the frequency of social ties between the inhabitants of two different places.

Most importantly, the findings from the empirical application of spatially weighted context data that we have presented in this paper have revealed that the examined collective war experiences have different types of impacts at different scales. Although a strong ‘local’ concentration of traumatising events appears to sustain outgroup blame, a high frequency of the same type of events within a more extended space appears to facilitate critical inward reflection. Thus, the interpretation of collective experiences of war varies substantially depending on the assumptions about the distance across which people are significantly affected by war events experienced by others. In cases like this, in which the empirical relationships vary across scale, traditional approaches to (multilevel) contextual analysis are extremely likely to neglect important parts of the pattern (and thus lead to the premature conclusion that there are no substantial context effects), whereas multilevel modelling with spatially weighted context data allows to view a more complete picture.

Furthermore, spatially weighted context data do not only help social scientists to determine where to ‘watch out’ for contextual effects; they can also substantially increase the precision of their sight. In more technical terms, the appropriate use of spatially weighted context data can considerably reduce the impact of measurement errors on contextual estimates or the number of observations required to reach a sufficiently low level of random error. In this sense, spatial weighting can also be seen as a possible complement or alternative to analytic procedures, such as Bayesian shrinkage, that aim to obtain the highest possible level of precision from a limited number of observations. Therefore, spatially weighted context data, such as implemented here with analytic tools made easily available for other researchers through the R-package *Spacom*, might also help overcoming situations where micro-level sources used to generate cross-nationally comparable context data are confined to a handful of resource-intensive international research programmes. Thus, the approach that we have outlined in this study is meant to run counter to the growing division of labour between survey data 'producers' and 'users' and to further encourage new small- to mid-scale survey projects in which researchers tailor theory-driven definitions of 'context' and measures of contextual 'indicators' to their own research questions in specific fields of substantive inquiry.

Notes

[i] Surveying the same population with two different sampling designs could in effect lead to important non-random differences in Bayesian-shrunk point estimates.

[ii] In section 4, we present a two-step approach. First, a geographic definition of space is used as a heuristic device to identify spatial patterns and generate a hypothesis about the underlying social mechanisms. Second, specific non-geographic social proximity functions are introduced to explain the observed spatial structures of collective experiences (see also Beck, Gleditsch & Beardsley, 2006).

[iii] Examples of quantifications of social distance between contexts are presented in section 4, but there is no a priori limited set of possible definitions of social distance. Whether a given definition of social distance is appropriate or not will always depend on its consistency with a study's theoretical framework, and whether it is practical will depend on the nature of the available data sources.

[iv] A detailed description of the survey methodology is provided in Spini, Elcheroth and Fasel (2011), and the full datasets can be accessed through the Data and Research Information Services from the Swiss Foundation for Research in the Social Sciences (www.unil.ch/daris).

[v] Conceived as a supportive tool for retrospective reports of personal events based on a model of autobiographical remembering as a structured cognitive activity, life events calendars provide multiple temporal anchors to support accurate remembering of events. These assets are best exploited through flexible, face-to-face interview techniques that adapt the sequence of event recording to the cognitive recollection process, whereby the interviewers reiterate previously unanswered items while the available episodic anchors become progressively richer and denser (see Axinn, Barber & Ghimire, 1997)

[vi] When respondents were unable to associate precise *dates* with specific events, these dates were imputed by a multiple regression model that took into account the reported dates for all other life events, the initial geographic location of the respondents in 1990, and their sex and age as multiple predictors. A set of diagnostic analyses provided by Honaker, King, and Blackwell (2007) showed that this imputational model had a particularly strong predictive value (surely due to the particularly strong dependency of war events on specific times and places) and that the additional measurement error introduced by the imputed dates was negligible. Therefore, to simplify subsequent calculations, we treated these imputed dates *as if* they were observed. We did not use imputed values for variables other than the dates of events.

[vii] Design weights correct for unequal selection probabilities due to geographical stratification, differential household sizes, and, in Kosovo (where, in the absence of appropriate frames for settlements, sampling points were generated by the random selection of geographical point coordinates), differential local population densities.

[viii] In the present application, Arc-GIS software was used to map the contextual values and to compare the geographic distance matrixes.

[ix] Census values or official estimates provided by the respective national statistical offices, for reference periods between 2001 and 2005.

[x] While the personal war experience variable thus indicates direct exposure of an individual to war-related events, independently of where they happened, contextual exposure reflects currently living in a social environment that has been shaped by collective war experiences in the past.

[xi] In theory, context data and individual-level data may stem from the same unique (micro-level) survey. However, in the present case, the second survey was necessary to meet the particular goals of the research project. As we aimed to describe the imprint of collective experiences on generational worldviews among the 'young adults of war' cohort, we had to substantially over-sample this cohort and, in addition to the life events calendar, administer an extensive attitudinal questionnaire to cohort members.

References

Axinn, W. G., Barber, J. S., & Ghimire, D. J. (1997). The neigborhood history calendar: A data collection method designed for dynamic multilevel modeling. *Sociological Methodology, 27*, 355-392.

Backe, K. M., O'Loughlin, J., & Ward, M. D. (2009). Social Distance in Bosnia and the North Caucasus Region of Russia: Inter- and intra-ethnic attitudes and identities. *Nations and Nationalism, 15* (2), 227-253.

Banerjee, S., Carlin, B., P., & Gelfand, A. E. (2004). *Hierarchical modelling and analysis for spatial data.* Boca Raton: Chapman & Hall.

Beck, N., Gleditsch, K.,& Beardsley, K. (2006). Space is more than geography: Using spatial econometrics in the study of political economy. *International Studies Quarterly, 50*, 27-44.

Branscombe, N. R.,& Doosje, B. (2004). *Collective guilt.* Cambridge: Cambridge University Press.

Brunsdon, C., Fotheringham, S., & Charlton, M. (1998). Geographically weighted regression : modelling spatial non-statinonarity. *The Statistician, 47*(3), 431-443.

Chaix B., Merlo J., & Chauvin, P. (2005). Comparison of a spatial approach with the multilevel approach for investigating place effects on health: the example of healthcare utilisation in France. *Journal of Epidemiololgy and Community Health, 59*, 517-26.

Courgeau, D., & Baccaini , B. (1998). Multilevel Analysis in the Social Sciences. *Population, 10*(1), 39-71.

Diez-Roux, A. V. (1998). Bringing context back into epidemiology: variables and fallacies in multilevel analysis.*Americain Journal of Public Health, 88*, 216-222.

Dogan, M., & Rokkan, S. (1969). Quantitative ecological analysis in the social sciences. Cambridge (Massachusetts) & London: The MIT Press.

Elcheroth, G. (2006). Individual-level and community-level effcts of war trauma on social representations related to humanitarian law. *European Journal of Social Psychology, 36*, 907-903.

Elcheroth, G., Doise, W., & Reicher, S. (2011). On the Knowledge of Politics and the Politics of Knowledge: How a Social Representations Approach Helps us Rethink the Subject of Political Psychology. *Political Psychology, 32(5),* 729-758*.*

Elcheroth, G.,& Spini, D. (2010). Political violence, intergroup conflict, and ethnic categories. In D. Bar-Tal (ed.), *Intergroup conflicts and their resolution: A social psychological perspective* (pp.175-194). New York: Psychology Press.

Flowerdew, R., Manley D., & Sabel, C. (2008). Neighbourhood effects on health: Does it matter where you draw the boundaries? *Social Science and Medicine, 66*(6), 1241-1255.

Fotheringham, A. S., & Wong, D. W. S. (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning, 23*(7) 1025 – 1044.

Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2002). *Geographically weighted regression.* Chichester: John Wiley & Sons.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models.*Cambridge: Cambridge University Press.

Honaker, J., King, G., & Blackwell, M. (2007). *Amelia II: A program for missing data*. Unpublished user's guide. Harvard University. Available athttp://gking.harvard.edu/amelia/

Hox, J. (2002). *Multilevel Analyses. Techniques and Applications.* Mahwah, New Jersey & London: Lawrence Erlbaum Associates.

Junge, T., Penic, S., &Elcheroth, G. (2012).*Spacom: Spatially weighted context data for multilevel modelling.*R package version 0.9-4. http://cran.r-project.org/web/packages/spacom/

LeSage, J. P. and Pace, R. K. (2009).*Introduction to Spatial Econometrics*.Boca Raton, FL.: CRC Press.

Lewis-Beck, M.S. (1990). *Economics and elections: The major Western Democracies.* Ann Arbor: The University of Michigan Press.

Moran, T. (2006). Statistical Inference for Measures of Inequality with a Cross-National Bootstrap Application, *Sociological Methods and Research, 34*(3), 296-333.

Openshaw, S. (1984).The modifiable areal unit problem. Norwich: Geo Books.

Savitz, N. V., & Raudenbush, S. W. (2009). Exploiting Spatial Dependence to Improve Measurement of Neighborhood Social Processes. *Sociological Methodology*, *39*(1), 151-183.

Schuman, H., & Scott, J. (1989). Generations and collective memories. *American Sociological Review, 54*, 359–381.

Schwartz, S. (1994). The fallacy of the ecological fallacy: the potential misuse of a concept and the consequences. *American Journal of Public Health, 84*(5), 819-24.

Snijders, T.A.B., & Bosker, R.J. (1999) *Multilevel Analysis. An introduction to basic and advanced multilevel modeling.* London: Sage

Spini, D., Elcheroth, G., & Fasel, R. (2008). The Impact of Group Norms and Generalization of Risks across Groups on Judgments of War Behavior. *Political Psychology, 29*(6), 919-941.

Spini, D., Elcheroth, G., & Fasel, R. (2011). TRACES: Methodological and technical report. *LIVES Working Papers, 4*.

Stow, C.A., Lamon, E.C., Qian, S.S., Soranno, P.A.,& Reckhow, K.H. (2008). Bayesian Hierarchical/Multilevel Models for Inference and Prediction Using Cross-Sectional Lake Data (pp. 111-136). In S. Miao, S. Carstenn, &M. Nungesser (eds.), *Real World Ecology: Large-scale and Long-term Case Studies and Methods.*Berlin: Springer.

Van der Leeden, R., Meijer, E., & Busing, F. M. T. A. (2008). Resampling multilevel models. In J. de Leeuw & E. Meijer (Eds.), *Handbook of multilevel analysis* (pp. 401–433). New York: Springer.

Ward, M. & Gleditsch, H. S. (2002). Location, Location, Location: An MCMC Approach to Modeling the Spatial Context of War and Peace, *Political Analysis, 10*(2), 244-260.

Appendix 1: Example of life events calendar

CALENDAR B
Bosnia-Herzegovina

Year	Quarter	Age	Marker event (optional)	Date of event	Location: Area (Code)	Location: Area (Code)	Why did you move?: Family	Why did you move?: Economic	Why did you move?: War	Why did you move?: Other	Lack of resources	Homeless	Unemployed	Separated from important people	Arbitrary treatment	Discrimination	Not allowed to express opinion	Exposed to threats	Forced to leave home	Imprisoned or kidnapped	Member of family killed	Damage to property	Wounded by the fighting	House looted	Carrying a weapon	Using a weapon	Quarter		Year
1990	1																										1	January, February, March	1990
1990	2																										2	April, May, June	1990
1990	3																										3	July, August, September	1990
1990	4																										4	October, November, December	1990
1991	1																										1	January, February, March	1991
1991	2																										2	April, May, June	1991
1991	3																										3	July, August, September	1991
1991	4																										4	October, November, December	1991
1992	1																										1	January, February, March	1992
1992	2		Beginning of war in BiH	X																							2	April, May, June	1992
1992	3																										3	July, August, September	1992
1992	4																										4	October, November, December	1992
1993	1																										1	January, February, March	1993
1993	2																										2	April, May, June	1993
1993	3																										3	July, August, September	1993
1993	4																										4	October, November, December	1993
1994	1																										1	January, February, March	1994
1994	2																										2	April, May, June	1994
1994	3																										3	July, August, September	1994
1994	4																										4	October, November, December	1994
1995	1																										1	January, February, March	1995
1995	2																										2	April, May, June	1995
1995	3		Dayton Agreement	X																							3	July, August, September	1995
1995	4																										4	October, November, December	1995
1996	1																										1	January, February, March	1996
1996	2																										2	April, May, June	1996
1996	3																										3	July, August, September	1996
1996	4																										4	October, November, December	1996
1997	1																										1	January, February, March	1997
1997	2																										2	April, May, June	1997
1997	3																										3	July, August, September	1997
1997	4																										4	October, November, December	1997
1998	1																										1	January, February, March	1998
1998	2																										2	April, May, June	1998
1998	3																										3	July, August, September	1998
1998	4																										4	October, November, December	1998
1999	1																										1	January, February, March	1999
1999	2																										2	April, May, June	1999
1999	3																										3	July, August, September	1999
1999	4																										4	October, November, December	1999
2000	1																										1	January, February, March	2000
2000	2																										2	April, May, June	2000
2000	3																										3	July, August, September	2000
2000	4																										4	October, November, December	2000
2001	1																										1	January, February, March	2001
2001	2																										2	April, May, June	2001
2001	3																										3	July, August, September	2001
2001	4																										4	October, November, December	2001
2002	1																										1	January, February, March	2002
2002	2																										2	April, May, June	2002
2002	3																										3	July, August, September	2002
2002	4																										4	October, November, December	2002
2003	1																										1	January, February, March	2003
2003	2																										2	April, May, June	2003
2003	3																										3	July, August, September	2003
2003	4																										4	October, November, December	2003
2004	1																										1	January, February, March	2004
2004	2																										2	April, May, June	2004
2004	3																										3	July, August, September	2004
2004	4																										4	October, November, December	2004
2005	1																										1	January, February, March	2005
2005	2																										2	April, May, June	2005
2005	3																										3	July, August, September	2005
2005	4																										4	October, November, December	2005
2006	1																										1	January, February, March	2006
2006	2																										2	April, May, June	2006
2006	3																										3	July, August, September	2006
2006	4																										4	October, November, December	2006

Appendix 2: Precise item wordings

War events

Have you been forced to leave your home and live elsewhere (as a consequence of war)?

Have you ever been imprisoned, kidnapped, or taken hostage (as a consequence of war)?

Has a member of your immediate family been killed during the armed conflict (son, daughter, father, mother, brother, sister, grandmother, grandfather, or grandchild)?

Has there been serious damage to your property (to your belongings) (as a consequence of war)?

Have you been wounded by the fighting?

Did you have your house looted (as a consequence of war)?

Collective guilt acceptance

I feel regret for my group's harmful past actions toward other groups.

I feel guilty about the negative things my ancestors did to other groups.

I feel regret for some of the things my group did to other groups in the past.

I believe that I should repair the damage caused to others by my group.

I can easily feel guilty for the bad outcomes brought about by members of my group.

Collective guilt assignment

Other groups have benefited at the expense of my group for generations.

It makes me upset that my group has been used to benefit other groups throughout history.

I feel entitled to concessions for past wrongs that other groups have done to my group.

Other groups that have benefited at the expense of my group owe us now.

It distresses me that my group suffers today because of the wrongs of former generations of another group.