



UvA-DARE (Digital Academic Repository)

Politics in Space

Methodological Considerations for Taking Space Seriously in Subnational Research

Harbers, I.; Ingram, M.C.

DOI

[10.1017/9781108678384.002](https://doi.org/10.1017/9781108678384.002)

Publication date

2019

Document Version

Final published version

Published in

Inside Countries

[Link to publication](#)

Citation for published version (APA):

Harbers, I., & Ingram, M. C. (2019). Politics in Space: Methodological Considerations for Taking Space Seriously in Subnational Research. In A. Giraudy, E. Moncada, & R. Snyder (Eds.), *Inside Countries: Subnational Research in Comparative Politics* (pp. 57-91). Cambridge University Press. <https://doi.org/10.1017/9781108678384.002>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Politics in Space

Methodological Considerations for Taking Space Seriously in Subnational Research

Imke Harbers
Matthew C. Ingram

Throughout the twentieth century, methodological nationalism has been the predominant form of thinking about political phenomena. In recent years, there has been a critical reevaluation of how readily social scientists, and especially scholars of comparative politics and international relations, accepted the nation-state as the most important level of analysis. Letting go of the simplifying assumption that the primary causes and consequences of political phenomena are located in the national arena enables scholars to more adequately map and explain the spatially uneven nature of contemporary political and economic transformations (Snyder, 2001). Indeed, over the past two decades a rich research program has emerged in which scholars draw on the subnational approach to better understand phenomena such as state formation, democratization, and development. Despite its undisputed potential, however, the subnational approach also creates specific challenges for researchers throughout the research cycle that have yet to be resolved.

This chapter explores how insights from Geographic Information Systems (GIS) and spatial analysis can help us work through some of these challenges. Furthermore, we highlight how a spatial perspective can strengthen the subnational approach by opening up new opportunities for theory development and analysis. While some of the techniques discussed below have been available since the 1980s (Doreian, 1980, 1982; Cliff & Ord, 1981; Anselin, 1988), political science has been slower than other social science disciplines to adopt a spatial perspective. Moreover, current work in comparative politics has tended to use GIS primarily for visualizing data at the level of subnational jurisdictions, without recognizing its potential for

We thank Richard Snyder, Agustina Giraudy, and Eduardo Moncada; workshop participants at Brown University in 2013 and Harvard University in 2014, especially Lily Tsai and Daniel Ziblatt; and two anonymous readers for their comments on an earlier draft. We are grateful to Alejandro Trelles and Miguel Carreras for sharing their municipal-level data on homicide rates from Mexico. All remaining errors are our own.

theory development or analysis.¹ Throughout this chapter, we include citations to key texts in spatial analysis to enable further reading and a more in-depth look at the techniques discussed.

Our central argument is that there is much to gain by thinking more explicitly about how the phenomena we study are situated in space and how space may, in turn, structure or condition outcomes and causal relationships of interest. By “space” we mean the geographic connectivity among units of observation. This connectivity can be conceptualized in multiple ways (e.g., contiguity, distance), but we emphasize its geographic or territorial nature (see Appendix on spatial weights). The spatial nature of connectivity comes into clearer focus if contrasted with the relational nature of connectivity in the field of network analysis. For instance, in a study of voting behavior, network analysts would be more interested in the associational ties and social closeness among individuals (e.g., Huckfeldt & Sprague, 1995), whereas spatial analysts would be more interested in their geographic proximity (e.g., Darmofal, 2006). Closeness, distance, and proximity can have overlapping connotations in both spatial and network research, but the key difference is that network connectivity is relational or affective whereas spatial connectivity is geographic.

“Taking space seriously” has key conceptual as well as theoretical implications. A major implication is the need for a more thorough recognition of the structural dependence that exists among units of observation. To be sure, this recognition is also important in international relations and cross-national comparative studies,² but the analytic shift is perhaps most important at the subnational level, where units of analysis have boundaries that are more porous than international borders. In quantitative research or mixed-methods designs with a quantitative component – which have recently become the norm in subnational comparative analyses (Moncada & Snyder, 2012) – treating these units as independently distributed is often untenable. Analyses drawing on estimation techniques that do not account for this spatial dependence then run the risk of obtaining incorrect, biased estimates and missing key factors influencing phenomena of interest. Yet most studies in comparative politics have so far left the issue of spatial dependence unaddressed.

More importantly, however, the value added of spatial analysis for subnational research lies not only in getting “right” answers to existing questions but also in bringing to the table exciting methods for (a) seeing existing questions in a new light and (b) identifying new and interesting questions that might otherwise go unnoticed. Taking space seriously thus also implies considering how spatial

¹ Among political science subfields, the international relations literature has been most proactive about embracing the notion of “politics in space” (e.g., Cederman & Gleditsch, 2009), and there has been a push to make GIS and related tools accessible to a larger audience (e.g., Gleditsch & Ward, 2005; Gleditsch & Weidman, 2012; Ward & Gleditsch, 2018). In comparative politics, see Franzese and Hays (2008).

² See, e.g., Ward and Gleditsch (2018) for spatial dependence in international relations or Hafner-Burton et al. (2009) for network dependence.

dependence structures outcomes and relationships of interest, and it invites researchers to question the assumption that units are self-contained. Moreover, in conventional, large-N analyses, the relationship between an explanatory variable and the dependent variable is generally assumed to be the same among all units, and the relationship in one unit is assumed to be unaffected by the outcome or explanatory variables in nearby units. As discussed in the introduction to this volume, the research program on subnational democracy, for instance, has looked for the causes of subnational undemocratic regimes mostly within the units themselves or in the vertical interactions between subnational units and higher levels of government. In light of the permeability of subnational borders, however, it may also be valuable to explore more systematically horizontal interactions among units and issues of spillover, diffusion, contagion, and similar phenomena among observations. Schedler (2014), for instance, raises the question of how the spread of violence in Mexico subverts democracy. Applying a spatial lens to violence, as we show in this chapter, allows analysts to more fully theorize and assess how such phenomena are also spatial processes – i.e., causal processes structured by space.

In making our plea for a “spatial turn” in subnational research, we recognize that we are guilty ourselves of the sins we are exposing, namely, of practicing the “dark art” of treating subnational units as independently distributed observations and of not considering the effect of spatial structures on outcomes and relationships of interest (e.g., Ingram, 2013, 2016; Harbers, 2014). Further, we acknowledge that a spatial perspective is not a simple, cool trick, nor does it reduce to a quick methodological fix. Instead, this chapter is intended as a contribution to a conversation about how to study subnational politics in a more disciplined and self-conscious way by examining the implications of spatial thinking across three stages of research design: (1) conceptualization; (2) theorizing; and (3) analysis. Within each of these phases there are lessons to be learned from taking space more seriously, and there are important analytic costs of not doing so.³

The chapter follows the structure outlined in Table 2.1. Looking ahead, a major concern in the subnational literature is the marked variation within countries in outcomes of interest, especially democracy and security. Throughout the chapter, we draw on the substantive examples of subnational democracy and the territorial dimension of violence. As the editors highlight in their introduction to the volume, these are areas of research where insights from the subnational approach have been particularly valuable. In our discussion, we highlight current practices and explain how taking space seriously can open new directions for research. The first section of the paper focuses on conceptualization where closer attention to how concepts are related to spatial and institutional

³ In our discussion, we assume some basic familiarity with the vocabulary of spatial analysis. For readers unfamiliar with the logic of spatial analysis, we have included a brief appendix on spatial weights.

TABLE 2.1 *Strengthening Subnational Research: Implications of a Spatial Perspective for Conceptualization, Theory, and Analysis*

	Current Practices and Associated Challenges for Improving Subnational Research	Contributions of a Spatial Perspective to Improving Subnational Research
Conceptualization	<p><u>Practice:</u> Emphasis on adapting national-level concepts for subnational units</p> <p><u>Challenge:</u> Make explicit how space and/or institutional categories are related to the phenomenon of interest</p>	<ul style="list-style-type: none"> • Distinguish unbound from institutional phenomena • Identify appropriate level of analysis
Theory	<p><u>Practice:</u> Explaining causes and consequences of territorial variation (i.e., unevenness) but treat subnational units as independent of one another</p> <p><u>Challenge:</u> Conceptualize nature of spatial dependence among units</p>	<ul style="list-style-type: none"> • Recognize relationship between structures of spatial dependence and outcomes of interest • Elucidate whether sources of territorial variation are place-based or propagation-based, and what the underlying causal process entails
Analysis	<p><u>Practice:</u> Mixed-methods designs, where quantitative analyses draw on estimation techniques for time-series cross-sectional data</p> <p><u>Challenge:</u> Incorporate spatial dependence and interactions into analyses</p>	<ul style="list-style-type: none"> • Identify and specify connectivity among subnational units • Detect nature of spatial dependence in data: spatial error, spatial lag of DV, or mixed process, including spatial lag of IVs (i.e., correlated relationship, endogenous interaction, or exogenous interaction; Manski, 1993)

categories can help clarify the causes and consequences of territorial unevenness. Core research questions include whether this variation has local, contextual sources – what we call *place-based* processes – or whether the variation is due to factors that help or hinder the diffusion, spread, transfer, or spillover of the outcome of interest – what we call *propagation-based* processes. The section on theory argues that more deliberate attention to the role of space can generate shifts in our framework of analysis and yield valuable insights about causation in both place- and propagation-based processes. Lastly, we examine tools for exploratory and confirmatory spatial analysis, focusing on how a variety of techniques can advance quantitative analyses of spatial patterns in the data.

Across all three stages of research design, we draw on concrete examples to illustrate our points, including an extended analysis of homicide rates across Mexico's municipalities.

2.1 CONCEPTUALIZATION: INCLUDING SPACE IN CONCEPT FORMATION

The subnational turn has sparked a lively debate about concept formation. One important debate in subnational comparative analyses centers on whether and how concepts initially formulated at the national level can “travel” to subnational units. The classic issue in comparative politics concerns whether concepts can usefully be applied to different historical and cultural contexts without the risk of “stretching” (Sartori, 1970). This debate has recently been broadened to the question of whether concepts can travel across levels of analysis (e.g., Hilgers, 2011; Gibson & Suárez-Cao, 2010). Sartori (2005 [1976]) was highly skeptical about applying concepts developed for the national arena – like democracy – to subnational units. He specifically cautions against “jump unit fallacies,” where “a sub-state, i.e. a member of a federal state, is made equal to a sovereign state.” Discussing politics in the US South, Sartori stated that “with respect to ‘democracy’ . . . the single states are granted only a subordinate and limited autonomy. Hence Florida or Louisiana or Mississippi . . . are not states in the sense in which Mexico and Tanzania are such” (Sartori, 2005 [1976], p. 73). In light of a wealth of empirical evidence demonstrating territorial unevenness in democratization (e.g., Lankina & Getachew, 2006; Gervasoni, 2010; Giraudy, 2010; Schedler, 2014), however, the idea that we cannot meaningfully study intra-country variation in democracy is clearly unsatisfactory. Scholarship has therefore consciously discussed conditions under which concepts originally developed for the national arena can be applied to subnational units and whether acknowledging the presence of multiple levels of analysis (and power) creates the need to refine concepts, also at the national level (e.g., Harbers & Ingram, 2014).⁴

Beyond this debate, a more fundamental challenge arising from taking space seriously is choosing the appropriate subnational level of analysis. As Soifer's Chapter 3 in this volume points out, the issue of *unit selection* is particularly pressing in subnational research. Because many subnational analyses move into uncharted methodological territory, existing conceptualizations and theories may offer little guidance on whether the theory about a causal relationship of interest is *unit-independent*, *unit-specific*, or *unit-limiting*. In light of this, we

⁴ Refining concepts appears to be the course of action recommended by Sartori (2005 [1976], p. 74), because in federal systems “each level is of itself incomplete and/or reflective of the other level. With respect to ‘democracy’, for instance, the state level has a wholly subordinate jurisdiction (a clear case of incompleteness).” Sartori's own commitment to multi-level concepts remains haphazard, however, as Gibson and Suárez-Cao (2010) point out. Even though the national level is supposedly incomplete without the subnational level, Sartori makes no attempt to incorporate this in his typology of party systems.

echo Soifer's advice that scholars make explicit which theoretical and practical considerations have entered into choosing the units for analyses. In the following paragraphs we also outline why distinguishing between *institutional* and *unbound* phenomena may be helpful at the stage of concept formation. Here a spatial perspective allows us to recognize more explicitly how a phenomenon of interest is anchored in space.

Even though GIS software appeared in the social sciences only around the 2000s, the idea of analyzing how political and social phenomena relate to space is by no means novel. A classic example familiar to most political scientists is John Snow's investigation of the 1854 cholera epidemic in London. By visualizing where in the city cholera victims lived, Snow was able to identify the Broad Street water pump as one of the culprits in the outbreak. GIS facilitates such analyses of spatial patterns by providing software that can store and process large quantities of information and connect them to space in meaningful ways. The potential contribution of GIS for the social sciences arises from the ability to connect non-spatial observations and their properties to a specific location. GIS is therefore "a methodological and conceptual approach that allows for the linking together of spatial data, or data that is based on a physical space, with non-spatial data, which can be thought of as any data that contains no direct reference to physical location" (Parker & Asencio, 2009, p. 1). The process by which non-spatial data is linked to spatial locations is called "geocoding" or "georeferencing."

The issue of how the phenomena we study are related to institutional or spatial categories has received relatively scant attention in comparative politics. Most comparativists intuitively choose to study subnational politics within the boundaries of territorially delimited jurisdictions, and the decision to focus on these units as the relevant objects of inquiry has often seemed so natural that it is almost nonconscious. Spatially uneven processes such as democratization have therefore generally been studied by focusing on provinces, or states – i.e., "the territorially-defined subunits of a political system" (Snyder, 2001, p. 94). Yet formal jurisdictions or administrative units are by no means the only lens through which we can study spatially uneven processes, and whether they are always the most appropriate lens deserves careful consideration.⁵

⁵ We are, of course, by no means the first to point out that social, economic, and political phenomena do not necessarily align with administrative divisions, either at the national or the subnational level. Moncada and Snyder (2012) highlight that "coping with spatially complex, uneven, and unbound processes and flows" presents important challenges for comparative research. Rodrigues-Silveira (2013, p. 4) recently introduced the term "institutional unboundedness" to denote a "territorial mismatch between state administrative boundaries and social, political and economic processes." This is slightly different from distinguishing types of phenomena, as we propose, because the idea of a "mismatch" suggests that the relevant comparison is between institutional boundaries and broader processes, not whether the occurrence of the phenomenon is tied to the institution itself. For examples in criminology literature, see Mears and Bhati (2006).

To appreciate why this approach has been so prevalent, it is useful to consider the origins of many concepts in comparative politics. Comparativists – like Sartori – have generally studied concepts such as regimes, party systems, and parliaments in which the state and its jurisdictional boundaries play a key role. Methodological nationalism tended to assume that national borders circumscribed the most relevant social and political phenomena. This “whole nation bias” (Rokkan, 1970; also Snyder, 2001) took for granted an alignment of institutional and spatial categories. Theories of the state as well as of democracy thus tended to assume “a high degree of homogeneity in the scope, both territorial and functional, of the state and of the social order it supports” (O’Donnell, 1999, pp. 137–138). This assumption of homogeneity was always an analytic shortcut, even for advanced industrial countries. What is striking, though, is that not just phenomena clearly associated with the jurisdiction of the state were conceptualized and measured at the national level. In addition to regimes, party systems, and parliaments, phenomena such as crime rates, child mortality, and ethno-linguistic fractionalization were also conceived of as properties of countries. As Soifer highlights in his chapter, those in the latter category share the feature of being aggregates and thus raise concerns about the modifiable areal unit problem (MAUP).

Subnational research provides an opportunity to unpack more systematically how the phenomena we study are related to spatial categories. Rather than simply adapting categories to a lower level of aggregation, we suggest instead that it may be more useful to consider whether the phenomenon of interest is indeed related to institutional categories and circumscribed by jurisdictional boundaries. In geocoding non-spatial data it is important to make explicit to which spatial feature the phenomenon or attribute in question belongs. Comparativists have often intuitively linked non-spatial data to polygons representing formal, subnational administrative jurisdictions – sometimes without realizing that this move constitutes a conceptual choice, and that other options are available.

In line with the discussion of varieties of territorial units in the introduction to this volume, Map 2.1 offers an illustration of three different ways to leverage subnational research designs and increase the number of observations – as Snyder (2001) suggests – by focusing on subnational units. The first two maps in Map 2.1 reflect instances of subnational jurisdictional units – i.e., states and municipalities – and will be familiar to many comparativists. These units “have clearly demarcated, legally constituted boundaries” (see Introduction). Moreover, the boundaries are endogenous to particular institutional arenas and political processes.⁶ The third map divides Mexico into equal squares according to the PRIO-GRID, a unified spatial data structure for conflict research, which

⁶ Even though all subnational boundaries are endogenous to political processes, the extent to which subnational or national factors, as well as considerations of regional community or scale efficiency, determine jurisdictional design varies considerably across countries and over time. For an insightful discussion of this issue, see Hooghe and Marks (2016).



MAP 2.1 Different Approaches to Mapping Subnational Units in Mexico

has a resolution of 0.5 x 0.5 decimal degrees latitude/longitude or about 50 x 50 km at the equator. In contrast to the first two maps, these gridcells are “insensitive to political boundaries and developments, and they are completely exogenous to likely features of interest” (Tollefsen, Strand, & Buhaug, 2012, p. 363). Thus, whereas the units in the first two maps are politically meaningful, gridcells are intended to be arbitrary divisions of the territory.

The key issue to consider at the stage of concept formation is whether the phenomenon of interest is necessarily tied to an institutional arena. Even though political institutions are part and parcel of political science thinking, not all concepts are equally attached to the domain of institutions. Whereas some phenomena, such as cabinets or party systems, cannot be conceived apart from a political institutional setting, others, like criminal violence or disease, may not be circumscribed by institutional jurisdictions. In many other instances, the difference between institutional and unbound phenomena may not be so clear-cut, making it especially important to think about alternative levels of analysis.

Let us illustrate what this might look like with an example. One of the key insights provided by subnational research is that democratization within countries is a spatially uneven process. There is no consensus, however, about how democracy varies within countries. At least two ways are possible. In the first, variation in democracy is captured at the level of subnational jurisdictions. In the second, we might observe considerable variation even within subnational jurisdictions. Each of these two interpretations of unevenness implies distinct choices at the stage of concept formation.

According to the first logic, intra-country variation in democracy occurs because subnational jurisdictional units have democratic characteristics to varying degrees. Giraudy (2013) – following Goertz (2006) – conceptualizes subnational democracy in terms of four secondary dimensions: turnover; contestation for the executive; contestation for the legislature; and clean elections. These dimensions are all necessary and jointly sufficient to classify a regime as democratic. All dimensions contain an explicit reference to institutional categories, thus implying that democracy varies at the level of that particular institutional framework. An important implication of this conceptual choice is that only federal or politically decentralized countries display subnational variation in democracy (see also Lankina & Getachew, 2006; Gervasoni, 2010; Behrend, 2011).

An alternative view classifies democratic unevenness on the basis of variation in secondary dimensions. Goertz (2006, p. 107), for instance, identifies four secondary-level dimensions of national democracy: (1) competitiveness of participation; (2) executive recruitment; (3) constraints on the executive; and (4) political liberties. Whereas the first three dimensions are associated with institutions, the fourth dimension, political liberties, is not necessarily linked to an institutional arena. Liberties are under threat where the rule of law is weak and citizens fear for their safety and bodily integrity. Large-scale criminal, interpersonal, or state-sanctioned violence, such as exists in contemporary

Mexico, thus subverts democracy (Schedler, 2014). Yet, while levels of violence can be influenced by jurisdictional boundaries (Snyder and Durán-Martínez, 2009), violence itself is not tied to specific institutional settings (Messner et al., 1999; Baller et al., 2001; Mears & Bhati, 2006; Deane et al., 2008). Both violence and liberties can therefore vary within jurisdictional units, and, following this logic, even unitary countries and politically centralized countries can display subnational variation in democracy. This approach opens up the possibility of studying unevenness in democracy at the level of non-jurisdictional units, such as squatter settlements, shanty towns, or areas controlled or governed by non-state actors – e.g., gangs or criminal organizations.

An illustration of this approach is O'Donnell's (1999) conceptual map of the state according to what he calls blue, green, and brown areas, where each color denotes progressively greater deficits in the rule of law. State capacity and the quality of institutions often vary significantly within countries (e.g., Harbers, 2015). While residents of shanty towns may formally be entitled to the same rights and protections as residents of upper-middle-class neighborhoods, they generally cannot expect proper treatment from the justice system. This, O'Donnell argues, results in "low intensity citizenship" – often structured along territorial lines – which undermines democracy.⁷ Even within jurisdictions, we are therefore likely to encounter blue, green, and brown areas – and hence, variation in one of the constituent dimensions of democracy.

The answer to the question about how democracy varies within countries thus depends, in part, on the conceptualization of democracy. Our purpose is not to take sides in the debate about which concept of democracy is preferable but to point out that making explicit the ways in which concepts relate to spatial features is not a trivial matter. How, why, and at which level variation arises is important if we are to develop a better understanding of the causes and consequences of spatially uneven processes. Moreover, how the concept is defined can be helpful in identifying the appropriate subnational unit for comparative analyses. If we are interested in examining how variation in civil liberties shapes democracy, it is probably appropriate to collect data at smaller levels of aggregation than provinces or states and perhaps at an even lower level of analysis than municipalities (e.g., localities, neighborhoods).

The distinction between institutional and unbound phenomena arises at a very early stage in the research cycle, and it is different from concerns about the modifiable areal unit problem. Soifer's Chapter 3 focuses primarily on relationships between variables and on spelling out at which level the mechanism proposed by a given theory operates. Our focus until now, by contrast, has been univariate and conceptual. Questions about how

⁷ O'Donnell (1999) identifies Bolivia, Colombia, and Peru as countries characterized by extreme territorial heterogeneity. Yet, when the paper was first published in 1993, political decentralization in these countries was in its infancy and limited primarily to the municipal level. "Brown areas" therefore do not correspond to subnational jurisdictions.

phenomena are anchored in space thus arise even *prior* to the formulation of causal arguments.

What level of analysis is appropriate depends fundamentally on the research problem at hand. Ideally, theory relevant to answering the research question would guide selection. Yet, practical considerations like data availability may restrict choices (Baller et al., 2001, p. 569; see also Soifer, Chapter 3 in this volume). Still, from a spatial perspective, researchers should be wary of two common pitfalls when selecting levels of analysis: (1) selecting areas that are *too large*; and (2) selecting areas that are *too small*. If the research examines an individual-level phenomenon, even small areal units may overlook individual-level variation, and any attempt at causal inferences may be vulnerable to an ecological fallacy (King, 1997). Similarly, if the chosen areal unit is too large, meaningful variation in an outcome of interest at a lower level of aggregation remains unseen. Alternately, if units are too small and a researcher examines a phenomenon covering geographic areas larger than the chosen level of analysis, then splitting this area into smaller pieces will artificially produce spatial autocorrelation.

While remaining cognizant of these pitfalls, scholars might also consider the following practical criteria when choosing levels of analysis: (1) what level of analysis maximizes the number of observations; (2) what is the lowest level of analysis that still offers contiguous areas across the full national territory – i.e., complete contiguity; (3) what level of analysis maximizes comparability with existing studies; (4) what level of analysis maximizes boundary stability over time, thereby facilitating longitudinal studies⁸; and (5) what is the lowest level of analysis that still offers data availability, maximizing opportunities to “scale up” in the future. For most research questions, there will be trade-offs among these criteria, and scholars may seek to balance them in different ways. For many studies, multiple levels of analysis are plausible choices. Explicit attention to these issues, however, and a discussion of how tension between criteria for selecting levels of analysis is resolved will contribute not only to transparency but also to the accumulation of knowledge about core concepts and theories.

2.2 THEORY: RECOGNIZING SPATIAL DEPENDENCE IN CAUSAL ARGUMENTS

The key analytic insight of spatial analysis is the spatially dependent structure of data. Contrary to conventional regression analysis where individual observations are regarded as distributed independently, spatial analysis explicitly acknowledges that observations are connected in space. In this regard, spatial analysis shares important conceptual and analytic features with network, temporal, and multilevel analysis. Whereas network analysis

⁸ See Weidmann, Kuse, and Gleditsch (2010) for an approach that accounts for changing boundaries over time.

emphasizes social ties among observations, temporal analysis examines the influence of past aspects of a unit on present aspects of the unit, and multilevel analysis examines the embeddedness of observations in vertical structures, spatial analysis allows scholars to examine horizontal cross-unit interactions between observations located in space. Spatial dependence, similar to other types of dependent structures, is seen as both substantively meaningful and a methodological challenge that requires the use of diagnostic tools to determine appropriate modeling techniques.

To be sure, in his seminal piece on subnational research, Snyder (2001) highlighted the interdependence of subnational units. Yet his article was geared more toward small-N, controlled comparisons, and the implications of dependent or independent data structures may not resonate as strongly with scholars pursuing small-N work as they might for scholars pursuing large-N statistical analyses.⁹ More recently, Moncada and Snyder (2012) note that much subnational work has progressed to mixed-methods designs, integrating small-N, qualitative techniques with large-N, quantitative ones (see also the Introduction to this volume). Still, the issue of the dependent structure of subnational data and the nature of spatial dynamics has received limited attention in subnational research. Thus, despite drawing scholarly attention to the analytic leverage gained from “scaling down” and to the added leverage of multi-method research designs, subnational research in comparative politics – especially quantitative work – has largely ignored the structural dependence among observations. As we noted in the introduction to this chapter, it is precisely in the subnational context where we might expect territorial boundaries to be permeable and geographic units to be dependent – and therefore spatial analysis to be especially relevant.

Acknowledging that subnational units may be dependent raises questions about how outcomes of interest are distributed across space and why different types of spatial patterns emerge. Outcomes of interest may cluster in space in three principal ways. Figure 2.1 shows three stylized graphs (a–c, from left to right); each square within each of the three graphs represents a territorial unit. If there is no clustering, then we observe spatial randomness (a). That is, the outcome of interest exhibits no dependence on the underlying spatial structure. However, if high values of the outcome of interest tend to appear close to other high values, and low values near other low values, then the data exhibit clustering of similar values (b). By contrast, if high and low values appear near each other, then there is clustering of dissimilar values, as in c (see, e.g., Griffith, 1987, p. 37; Darmofal, 2015). These stylized patterns offer simplified versions of the different types of spatial patterns in cluster maps generated by

⁹ Though Franzese and Hays (2008, pp. 756, 760 n. 33) warn that even qualitative studies that neglect interdependence are vulnerable to biased estimates in the form of an inflated impression of the weight of nonspatial factors.

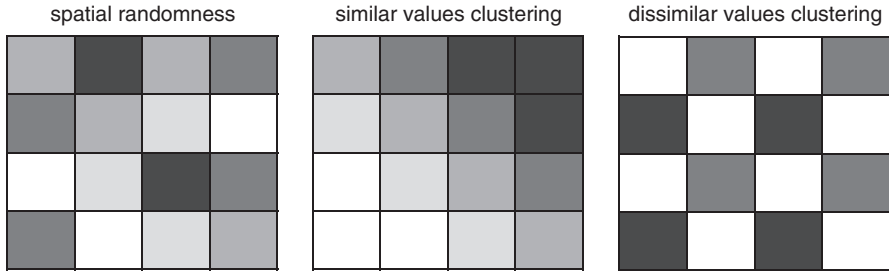


FIGURE 2.1 Varieties of Clustering Patterns

using the local indicators of spatial autocorrelation (LISA values, which we discuss in the analysis of Exploratory Spatial Analysis in Section 2.3).

Turning to more theory-oriented concerns, “[t]o interpret spatial patterns, we need spatial theories” (Logan et al., 2010, p. 15). From a spatial perspective, the causal process producing a spatial pattern of interest may come in two forms: (1) place-based, and (2) propagation-based. While theoretically distinct, empirically these two types of processes are not mutually exclusive, and may be present at the same time. Questions about place-based processes ask whether there is something about a particular area or region that creates a similar data-generating process for a set of units within that area, thereby producing a similar pattern in an outcome of interest across neighboring territorial units. In this respect, place-based relationships are instances of what Manski calls “correlated relationships” or what Franzese and Hays call “common exposure.” In practice, questions about place-based processes essentially try to identify a regional omitted variable. For instance, fertile soil is conducive to agriculture, which in turn contributes to the emergence of certain social structures (e.g., reliance on non-free labor and inequality) and thus encourages particular types of political order (e.g., strong local elites). Soil characteristics may vary across regions but not as rapidly as variation in administrative boundaries, so a study of a large set of adjacent units that overlooks their soil characteristics might miss an important determinant of political patterns within the units of observation. The notion of “neighborhoods” implied by this example is common in studies of a wide range of phenomena. Conversely, different geographic or place-specific conditions might help explain divergent patterns of electoral behavior in neighboring units if, for instance, there is a structural or geologic feature of the terrain that produces a different pattern of interactions within that unit.¹⁰ In this manner, research on place-based processes resembles work in the field of

¹⁰ For an example of how soil characteristics have shaped voting patterns in the United States, see www.npr.org/blogs/kurlewich/2012/10/02/162163801/obama-s-secret-weapon-in-the-south-small-dead-but-still-kickin.

international relations that explores whether there is a particular “stock” or characteristic of an area covering multiple countries and whether, in turn, this regional characteristic has far greater explanatory power than unit-specific properties or attributes (e.g., Kopstein & Reilly, 2000). A further example of a place-based process draws from Schedler’s (2014) study of patterns of violence in Mexico. In his review of existing research, Schedler notes the “labor supply” of young men as an untested correlate of violence. If this supply is spread over, or proximate to, many neighboring units of observation, then a key place-based source of violence may be obscured by studies that ignore this demographic feature.

The second kind of causal process that produces spatial patterns, propagation-based, involves the spread or diffusion of a phenomenon of interest between or among territorial units. Unpacking this idea further, diffusion can occur in both (a) the outcome of interest (Manski’s “endogenous interaction,” i.e., endogenous spread or diffusion) and (b) a predictor of the outcome of interest (Manski’s “exogenous interaction,” i.e., exogenous spread or diffusion). That is, the outcome of interest propagates itself (endogenous) or, alternatively, a change in a causal factor in nearby units produces a change in the outcome of interest in the focal unit (exogenous). When analyzing a propagation-based process, emphasis is given to cross-space data-generation, e.g., spatial contours that affect the connectedness or dependence among units and therefore help or hinder the spread of either the outcome of interest itself or of a predictor of this outcome located in nearby units. Propagation-based explanations resemble arguments in the international relations field that focus on “flows” as opposed to “stocks” (e.g., Kopstein & Reilly, 2000). Returning to Schedler’s review of research on violence in Mexico, he notes two major unanswered questions: (1) we do not know the boundaries or contours of the problem of violence in Mexico; and (2) while violence has generally been concentrated geographically, it has begun to spill over or diffuse to a larger number of units, yet we do not adequately understand how this happens. While spatial approaches can help answer both of these questions, they lend themselves especially well to addressing the second one, which concerns diffusion processes. Overall, a spatial perspective allows us both to conceptualize and theorize more effectively how geography affects the causal processes we are interested in.

These examples of place-based and propagation-based processes align with the way spatial dependence can be modeled mathematically. For instance, “unmeasured causes of crime [might] cluster in geographic space,” resulting in a place-based process. Conversely, a “diffusion process that causes crime to spill over from one district to a neighboring district” yields an endogenous propagation-based process. “The former process is modeled by a ‘spatial error’ term, while the latter process is more closely approximated by a ‘spatial lag’ term” (Messner et al., 2011, p. 9, citing Anselin & Bera, 1998; Baller et al., 2001). More specifically, endogenous

propagation is modeled by a spatial lag of the dependent variable; whereas exogenous propagation would be modeled by the spatial lag of an independent variable (LeSage & Pace, 2010).

In sum, spatially confined phenomena are associated with a place-based causal process whereas spatially interconnected phenomena are associated with a propagation-based causal process. The next section considers how the spatial error model captures the effect of spatially confined, yet unmeasured, variables and is thus especially useful for identifying relevant omitted variables and gaining leverage to generate new hypotheses and develop theory. The spatial lag model, on the other hand, captures propagation and diffusion effects. For both types of phenomena, theoretical arguments should explicate the causal process. Spatial analysis, which provides tools for assessing causal processes of diffusion and reciprocal influence across subnational units, is therefore especially useful for what the editors of the volume call a *horizontal strategy* (Table 2.2, Quadrant III) and a *reciprocal horizontal strategy* (Table 2.2, Quadrant VI) for subnational research.

2.3 ANALYSIS: IDENTIFYING SPATIAL DEPENDENCE

Because research in comparative politics has increasingly deployed mixed-methods designs, and time-series cross-sectional data have become a standard data structure, it is worth revisiting what a spatial perspective means for analytic techniques in the study of subnational politics. Specifically, how might scholars go about operationalizing the spatial structure of data and assessing the consequence of this structure, just as earlier methodological research placed a premium on operationalizing the temporal dynamics and serial autocorrelation present in such data structures (e.g., Beck & Katz, 1996; also Beck, Gleditsch, & Beardsley, 2006).

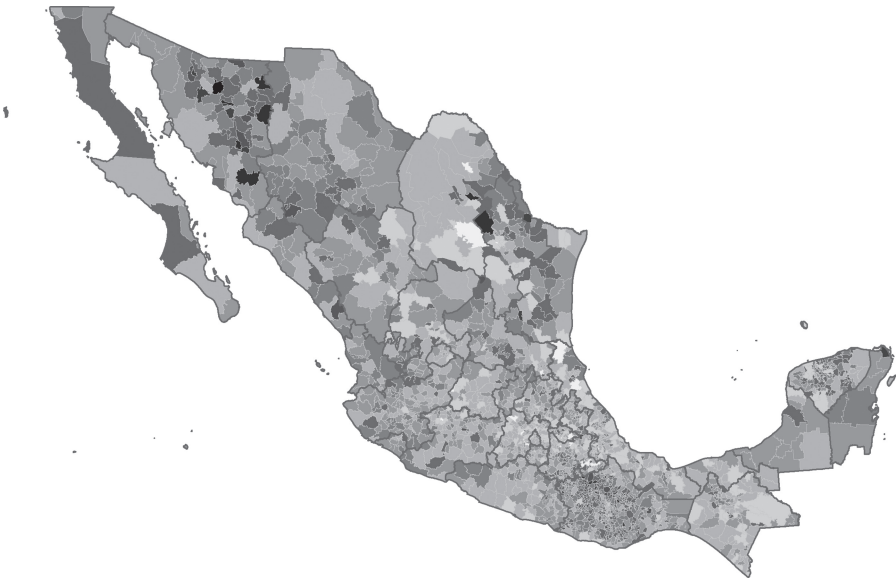
As with conceptualization, there are myriad ways to analyze “politics in space.” We start with some exploratory techniques but then focus on deductive, confirmatory approaches, acknowledging that these and other tools can also be employed in a more inductive fashion. A running example explores homicide rates across Mexico’s municipalities (homicides per 1,000 people), because violence has important implications for the territorial dimension of democracy (Schedler, 2014). From a place-based perspective, attributes or characteristics of a particular location may be predictors or determinants of violence in that space. From a propagation-based perspective, on the other hand, the attributes of a particular area encompassing several units may not matter, but the *connectedness* among communities may lead to high levels of violence in one community to increase violence in nearby communities (endogenous relationship), or, alternatively, the predictors of violence may exert an effect across territorial units, thereby influencing violence in neighboring units (exogenous relationship). Notably, existing research suggests that spatial patterns of violence were diminishing over time within Mexico since the 1980s

and were largely absent by 2003 (Snyder & Durán-Martínez, 2009, pp. 266–267). Thus, 2010 presents a “least likely” scenario for finding spatial patterns of violence. If any spatial patterns are found, then, they are that much more remarkable and deserving of attention.

Exploratory Spatial Analysis

Exploratory techniques or exploratory spatial data analysis (ESDA) is “a critical first step for visualizing patterns in the data, identifying spatial clusters and spatial outliers, and diagnosing possible misspecification in analytic models” (Baller et al., 2001, p. 563). Although maps are not a necessary step, “[g]raphical displays provide an auxiliary method [to data tables] that may allow patterns to be discovered visually, quickly” (Ward & Gleditsch, 2018, p. 23). For instance, the decile map in Map 2.2 visualizes municipal-level data on homicide rates in Mexico for the year 2010.¹¹

In the decile map, light shading identifies municipalities with low homicide rates, and the color darkens as the homicide rate increases. The darkest shades identify the municipalities with the highest homicide rates. Even a cursory



MAP 2.2 Decile Map of Homicide Rates in Mexican Municipalities (2010)

¹¹ Homicide data is from Trelles and Carreras (2012), and the municipal shapefile and georeferenced data are from INEGI (www.inegi.org.mx/geo/contenidos/geoestadistica/catalogoclaves.aspx; last accessed April 5, 2013).

glance at this map reveals concentrations of darker, violent areas in (1) the upper west coast of Mexico (across the states of Nayarit, Sinaloa, and Sonora), (2) the northeast (covering parts of three states: Coahuila, Nuevo León, Tamaulipas), (3) southern Mexico, and (4) portions of the Yucatán peninsula. Moreover, there are a few areas in northern, central, and southern Mexico that are almost clear of any color, i.e., have low homicide rates.

Helpful variants of this kind of visualization include standard deviation maps – maps that identify units that are one or more standard deviations above or below the mean. Even a quick glance at this kind of map would help identify spatial units that represent outliers or extreme values.

Two additional techniques include global and local tests of spatial autocorrelation, which can be used to assess the degree of structural dependence among units. Specifically, global and local tests of spatial autocorrelation posit a null hypothesis of no spatial dependence among observations, i.e., spatial randomness, and then test whether this null hypothesis is supported. One global test is the global Moran's I , which examines whether there are any regular patterns among geographically connected units (Moran, 1948; Cliff & Ord, 1981). If there are no regular patterns of spatial association, then the statistic is not significant. On the other hand, if there are significant spatial associations, the statistic can be positive or negative. A positive global Moran's I indicates that territorial units that are connected exhibit similar values on the outcome of interest; a negative result indicates territorial units that are connected have divergent or dissimilar values.

Table 2.2 lists global Moran's I values (and corresponding z-value) for homicide rates across Mexico's municipalities in 2010, as well as the average values for three time periods (2007–2009, 2001–2006, and 1995–2000), and the year with the highest value in the available data, 1996. All values are statistically significant at the .01 level.¹²

Looking only at 2010, the high z-value allows us to confidently reject the null hypothesis of spatial randomness in the data. This suggests that standard regression techniques would not only be inappropriate but would also overlook a key characteristic of the phenomenon of violence. Further, the highest Moran's I values appear prior to 2000, i.e., prior to the end of the PRI's 71-year rule that marked the national transition to democracy. Complementing Snyder and Durán-Martínez's (2009) suggestion that state-sponsored protection rackets that may have existed prior to 2000 were dissolved by the weakening of the PRI in the 1990s, these municipal-level data

¹² The population at risk in each Mexican municipality can vary considerably, so it is important to account for the variance instability of rates. Following Baller et al. (2001, p. 589) and Anselin (2005, p. 148), we do this by implementing an empirical Bayes (EB) standardization as suggested by Assunção and Reis (1999). Also, longitudinal comparisons are inappropriate if the underlying structure of geographic units changes considerably over time (see, e.g., Darmofal, 2006, p. 131 n.6). This is not the case with Mexico's municipalities during this time frame. Values generated in GeoDa v1.4.0.

TABLE 2.2 *Global Moran's I Values for Homicide Rates in Mexico's Municipalities, 1995–2010*

year	Moran's <i>I</i>	z-value
2010	0.0940	7.63
2007–2009avg	0.1003	
2001–2006avg	0.1002	
1995–2000avg	0.1306	
1996	0.1720	13.97

support their findings based on state-level data that spatial clustering of violence appears to be decreasing over time in Mexico. While these substantive findings are compelling and merit further exploration, we focus instead on the methodological lessons that: (1) longitudinal comparisons of spatial clustering is appropriate if the underlying spatial/geographic structure among units is stable (see note 12); and (2) the decision to focus on municipalities or states, or any other level of analysis, is a critical consideration, and ultimately depends on one's research questions and existing theory.

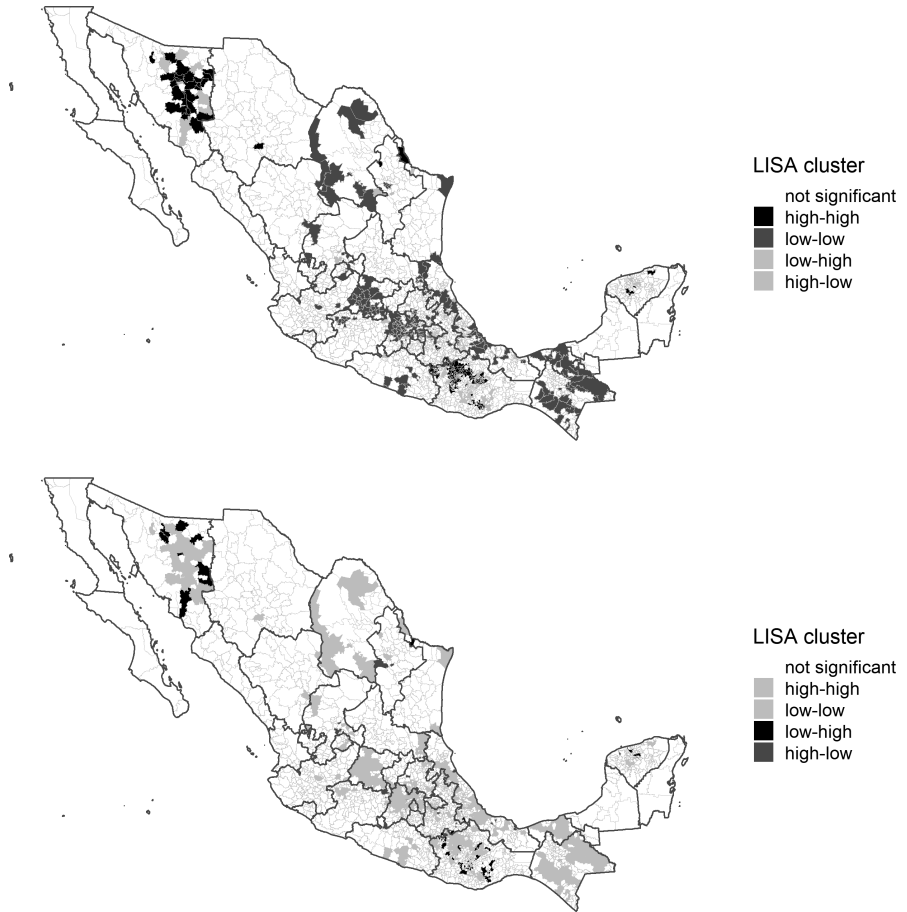
In addition to global tests of spatial autocorrelation, another useful technique for exploratory analysis of spatial dependence, is the *local* Moran's *I*, or LISA (Anselin, 1995). A LISA statistic provides information about the correlation of an outcome of interest among a focal unit *i* and the units to which *i* is connected, *j* (e.g., *i*'s neighbors), whether the association is positive (i.e., similar values) or negative (i.e., dissimilar values), and whether the association is statistically significant (see Appendix on spatial weights). Thus, LISA statistics help identify local clusters or spatial patterns of an outcome of interest. To be clear, while the global Moran's *I* may suggest little overall spatial autocorrelation in the data, LISA values can help identify smaller geographic areas where positive or negative clustering occurs.¹³

LISA statistics can be analyzed on their own to detect extreme values, but visualizing these statistics – for example, with LISA cluster maps – can offer a quick and instructive way to proceed. Depending on the depth of one's knowledge of the subject at hand, LISA cluster maps and other visualizations can be very revealing and may even serve to test hypotheses. If deeper or broader contextual knowledge is absent, however, then any visualization exercise is purely exploratory.

Map 2.3 reports LISA cluster maps in two panels.¹⁴ In both panels, blank areas indicate regions of spatial randomness in the distribution of violence,

¹³ The global Moran's *I* is the mean of all LISA values (Anselin, 2005, p. 141).

¹⁴ Color version is available from authors.



MAP 2.3 LISA Cluster Map of Homicide Rates

whereas shaded areas indicate nonrandom, statistically significant spatial clusters. All cluster associations are significant at least at the .05 level.¹⁵ Note also that the shaded municipalities constitute the *core* of spatial clusters. That is, the shaded municipalities have a statistically significant relationship with the municipalities that border them, including those that

¹⁵ All LISA statistics were generated using a conditional permutation approach (Anselin, 1995) with 999 permutations. Estimation was implemented in Python (version 3.5.2; Python Software Foundation, 2016) using the PySAL package (PySAL Developers, 2017; Rey & Anselin, 2007). All figures were generated in R (R Core Team, 2017), using ggplot package (Wickham, 2009).

are clear (i.e., have no shading). Thus, the outer boundary of the cluster extends into clear municipalities bordering shaded ones, and the true size of the spatial cluster is, in fact, larger than the shaded cores (see, e.g., Anselin, 2005, p. 146).

A LISA cluster map also contains information about the substantive content of spatial clusters. According to Anselin (2005, p. 140), this kind of map is “[a]rguably the most useful graph” in spatial analysis. In the first panel (top), for example, black identifies municipalities with higher-than-average homicide rates that are surrounded by municipalities with similarly high homicide rates (high-high). Medium shading, on the other hand, identifies municipalities with lower-than-average homicide rates surrounded by municipalities with similarly low rates (low-low). Other types of statistically significant clusters (discussed next) appear in light shading to distinguish them from the nonsignificant areas.

In addition to identifying statistically significant neighborhoods of high violence and low violence, LISA statistics also allow us to identify spatial outliers. The second panel (bottom) shows these outliers. Municipalities with low homicide rates, surrounded by ones with high rates (low-high) now appear with the darkest shading, and, conversely, municipalities with high homicide rates surrounded by ones with low rates (high-low) have medium shading. Again, the other types of statistically significant clusters (high-high and low-low, discussed earlier) appear in the lightest shading to distinguish them from non-significant areas.

Returning to the substantive issue of homicide rates in Mexico, the LISA cluster map presented in Map 2.3 provides strong evidence that complements our earlier, cursory evaluation of a map of the raw data (Map 2.2). Whereas we earlier identified the upper west coast of Mexico (from Nayarit to Sonora) as the clearest “hot spot” of violence, we now see that the northern portion of this geographic area constitutes the largest and clearest high-high spatial cluster. Adjacent to it, however, are less violent municipalities that do not fit the regional pattern and are therefore identified in the second panel as outliers, specifically as low-high clusters.

Additional insights can be gleaned by examining extreme values of LISA statistics – that is, by looking for the strongest, statistically significant positive and negative associations among focal and surrounding units. For instance, the five highest LISA values are all statistically significant: Four are from Oaxaca and one from Sonora, and all identify cores of high-high spatial clusters. Thus, from both the LISA cluster map and the examination of extreme LISA values, Sonora and Oaxaca would seem to be promising cases both for in-depth qualitative analysis and for more focused quantitative analysis. Moreover, visible clusters can also be seen that extend beyond state boundaries, including high-violence clusters in Coahuila and Nuevo León, and low-violence clusters across the country. These cases provide opportunities to explore whether formal state boundaries succeed (or fail) in containing violence. Notably, unlike the United States, where studies of homicide rates at

the county level show that the south is a high-violence region and the northeast is a low-violence region (Land et al., 1990; Baller et al., 2001), no single region in Mexico can be similarly singled out.

Overall, mapping an outcome of interest in the ways illustrated here can generate valuable insights about its geographic distribution. This, in turn, provides a useful starting point for further qualitative and quantitative analysis.¹⁶

Spatial Regressions and Diagnostics

The techniques of exploratory analysis outlined in the previous paragraphs can generate a variety of insights about both the core research question and about case selection for further research. Depending on the question, these exploratory tools may even serve to test hypotheses about spatial patterns in the outcome of interest. Econometric techniques take the analysis several steps further, allowing us to examine key questions about the spatial nature of subnational politics.

Continuing with the example of homicide rates in Mexico, we offer a basic OLS regression analysis, diagnostics based on this regression, and then two core versions of spatial regressions that can be used to examine different underlying spatial dynamics: (1) a spatial error model; and (2) a spatial lag model. Only one of these – the spatial lag model – is related to diffusion, that is, the propagation-based spread of an outcome of interest from one place to another, and it is thus important to (a) distinguish between these two models and, (b) based on diagnostics of the basic OLS model, determine which model is most appropriate. Beyond these two core spatial models, we also identify several extensions of basic spatial regressions, including (a) a spatial Durbin model with a lagged dependent variable and lagged independent variables, (b) a geographically weighted regression that allows predictors of interest to vary in their effects across spatial units, (c) a spatio-temporal regression that includes temporal as well as spatial lags in order to analyze spatial processes longitudinally, and (d) spatial regression-discontinuity designs, in which a geographic boundary is treated as randomly separating a control area from a treatment area. There are other models we do not address, including multilevel models and models with complex forms of both spatial and network dependence. Disentangling the effects of different forms of dependence – separating the spatial component of an effect from a temporal, relational, or vertical (multilevel) component – is not a simple task, but there are feasible approaches. Multiple statistical models aim to do this, along with multiple estimation strategies for each model and diagnostics to facilitate interpretation (Franzese & Hays, 2008; Bivand & Piras, 2015; Darmofal,

¹⁶ See Harbers and Ingram (2017a, 2017b) for specific strategies, including case selection strategies, using diagnostics of spatial regressions.

2015). Our goal here is not to provide an exhaustive assessment of all models and estimation strategies but to offer an introductory orientation to promising tools for exploring the spatial dimension of subnational politics.

Table 2.3 reports five models.¹⁷ Drawing on existing research on the structural covariates of homicide rates in the United States (Land et al., 1990; Baller et al., 2001), the key independent variables in the models capture population pressures (total municipal population and proportion of the population that is male), socioeconomic pressures (average years of education, income per capita), and unemployment pressures (percent of the population that is *not* economically active). Following Graif and Sampson (2009), the models also capture migration pressures (proportion of the population that was born in another state), and building on conflict studies that find mountainous and other rough terrain is conducive to higher levels of violence (e.g., Fearon & Laitin, 2003), two variables capture elevation and the unevenness of the terrain (altitude and the standard deviation of altitude). The goal is not to provide the best specification of a model of homicide rates but instead to offer a reasonable approximation of such a model with the purpose of assessing the role of space in explaining patterns in this violence.

Across the models, the predictors of homicide rates are of less interest for our current purposes than determining the nature of spatial autocorrelation. The first model reports a basic OLS estimation (OLS1). Moran's *I* of the residuals measures remaining spatial autocorrelation unaccounted for by the variables in the model; the value of 0.0731 is highly statistically significant ($p < 0.001$), strongly suggesting a positive spatial association not captured by the model. OLS2 reports an OLS model with a region dummy (0,1) where 1 captures states in Map 2.2 with high levels of violence (Chihuahua, Durango, Guerrero, Jalisco, Michoacán, Nuevo León, Sinaloa, Sonora, and Tamaulipas). That is, OLS2 reflects a common practice among researchers who attempt to account for regional effects by adding a simple dummy variable for a particular region or set of geographic units.¹⁸

However, diagnostics of the basic OLS model (OLS1) based on a classical Lagrange Multiplier (LM) test identify what kind of model may best capture the spatial dependence in the data. In other words, examination of the residuals helps test for spatial autocorrelation and determine which form of spatial dependence is present in the data. The LM test accomplishes this task, and two forms of the test distinguish between “spatial lag” and “spatial error” types of spatial dependence (Anselin, 1988; Baller et al., 2001, p. 590).¹⁹

¹⁷ All models run in R (R Core Team, 2017); spatial statistics and models generated using package *spdep* (Bivand & Piras, 2015).

¹⁸ A similar model with a dummy for all northern Mexican states bordering the United States, not reported here, showed no meaningful differences.

¹⁹ Prior studies refer to the spatial error model also as a “spatial disturbance” model (see, e.g., Baller et al., 2001). Network analysts also refer to these models as “network effects” and “network disturbance” models (e.g., Dow, 2007).

TABLE 2.3 *Models of Municipal Homicide Rates in Mexico*²⁰

y = homicide rate	OLS1	OLS2	SEM	SLM1	SLM2
(Intercept)	1.95** (0.66)	1.91** (0.66)	1.96** (0.69)	1.83*** (0.49)	1.83*** (0.65)
Population	-0.61*** (0.01)	-0.61*** (0.01)	-0.61*** (0.02)	-0.56*** (0.02)	-0.56*** (0.02)
% male	-1.30 (1.15)	-1.25 (1.16)	-1.59 (1.22)	-0.48 (0.87)	-0.47 (1.15)
% out of state	0.00 (0.02)	0.00 (0.02)	0.00 (0.02)	0.01 (0.01)	0.01 (0.02)
Education	-0.07*** (0.02)	-0.07*** (0.02)	-0.07*** (0.02)	-0.07*** (0.02)	-0.07*** (0.02)
Economic Inactivity	-0.52 (0.38)	-0.53 (0.38)	-0.54 (0.39)	-0.57 (0.36)	-0.57 (0.37)
Income	0.45*** (0.04)	0.45*** (0.04)	0.45*** (0.04)	0.39*** (0.04)	0.39*** (0.04)
Inequality (2005)	-1.25*** (0.34)	-1.25*** (0.34)	-1.18** (0.37)	-1.07** (0.33)	-1.07** (0.33)
Altitude	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Uneven terrain	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.02 (0.01)	-0.02 (0.01)
Region		-0.02 (0.04)			
Lambda			0.17*** (0.03)		
Rho				0.16*** (0.02)	0.16*** (0.03)
N	2455	2455	2455	2455	2455
RMSE	0.8163	0.8164	0.8076	0.8039	.8057
AIC	5982.26	5984.02	5954.19	5930.83	n/a
Moran's I (residuals)	0.0731	0.0733	-0.0057	-0.0068	-0.0076
Moran's I (p-value)	0.000	0.000	0.671	0.611	0.564

Note: *p<.05 **p<.01 ***p<.001

For all practical purposes, researchers need only consult the LM tests. Here, LM tests using a spatial weights matrix (W) based on simple rook-1 contiguity

²⁰ All measures are for 2010 unless otherwise noted. Population = total population, logged; % males = percentage of total population that is male; % out of state = percentage of total population that reported being born out of state; Education = average years of education; Economic Inactivity = % of total population classified as not economically active; Income = per capita income in 2005 in US dollars (logged); Inequality = Gini coefficient for 2005; Altitude = average altitude above sea level among localities in municipality; Uneven terrain = standard deviation of altitude among localities in a municipality.

(see Appendix) produced the following results: LMerror = 34.19, $p < .001$; robust LMerror = 0.34, $p = 0.559$; LMlag = 57.82, $p < .001$; robust LMlag = 23.97; $p < .001$. The simple versions of the LMerror and LMlag tests provide an initial indication about whether spatial dependence exists in the data. Both are significant, suggesting both forms of spatial dependence are present. In these situations, we consult the robust forms of both (Anselin et al., 1995). Doing so reveals that only the robust LMlag is significant. This is good evidence that the spatial dependence follows a lag structure (i.e., Manski's "endogenous interaction") rather than an error or disturbance structure. This is also an indication that a diffusion or spillover process may be at work. Conversely, if the LMlag test were not significant, this would be a strong indication that there is no diffusion at work and that looking for diffusion may not be a fruitful avenue of further research. At the same time, the evidence we found of a spatial lag structure does not allow us to specify the exact nature of the diffusion process, because the LM test does not help identify the mechanisms of diffusion.

The next three models illustrate improvement in goodness of fit as the models capture the appropriate spatial dynamics. Although the spatial error model (SEM) was not the preferred model indicated by the LM tests, even this model improves the fit, reducing RMSE (or standard error of the regression), reducing AIC, and reducing Moran's I to the point that it is no longer statistically significant. Finally, the two spatial lag models (SLM1 employing maximum likelihood estimation, and SLM2 employing two-stage least squares) are the most indicated specifications and show further improvement in fit. Thus, even though LM tests indicate a spatial lag model (SLM) would be a better option, the SEM model turns out to be a better fit than either of the nonspatial models (i.e., OLS1 and OLS2), illustrating a point made by Franzese and Hays (2008a, p. 760) that, even if modest interdependence is present, any spatial model is better than a nonspatial one because, as they put it, "ignoring interdependence when appreciably present is usually far worse than imperfectly including it."²¹

Various extensions of the basic spatial models above (SEM and SLM) are possible, including a spatial Durbin model, geographically weighted regression (GWR; Fotheringham, Brunson, & Charlton, 2002), longitudinal models with a temporally lagged spatial lag (e.g., Franzese & Hays, 2008; Ward & Gleditsch, 2018), and geographic regression-discontinuity designs (Keele & Titunik, 2015). Spatial Durbin models (SDM) incorporate a spatial lag of the dependent variable and also spatial lags of the independent variables (LeSage & Pace, 2010; Ellhorst, 2010; Yang et al., 2015). In this way, these models afford rich opportunities for exploring the spillover effect of the outcome of interest and of the dynamic local (direct) and neighbor (indirect) effect of explanatory variables. SDMs are thus considered the "state of the art" of spatial analysis

²¹ Elsewhere, Franzese and Hays are more forceful: "Given any noticeable interdependence, then, nonspatial [least-squares] is an unmitigated disaster" (2008b, p. 6).

(Ellhorst, 2010; Yang et al., 2015). For an analysis of homicide in Mexico employing this technique, see Ingram (2014).

GWR offers a qualitatively different kind of analysis. Even when spatial lag or spatial error components of a model are significant, these models assume that the other covariates have a uniform effect across all units studied. That is, independent variables are assumed to have globally invariant or stationary effects. For instance, in the preceding analysis, education is assumed to have the same effect on violence across all Mexican municipalities. This assumption may be patently untenable in some situations, and it may be especially “inappropriate for modeling political behavior in a geopolitically diverse polity” (Darmofal, 2008, p. 957). GWR offers an alternative, allowing for “spatial variability of regression results across a region” so that “rather than [having to] accept one set of ‘global’ regression results, [researchers can produce] ‘local’ regression results from any point within the region so that the output from the analysis is a set of mappable statistics which denote local relationships” (Fotheringham, Charlton, & Brunson, 1998). Thus, where theory leads analysts to anticipate that the effect of a key explanatory variable may vary in significance or magnitude across spatial units or, alternatively, that interaction among key variables may produce different effects across space (e.g., Darmofal, 2008), a geographically weighted regression is appropriate. Ingram and Marchesini da Costa (2017) offer an analysis of homicide across Brazil’s municipalities using this approach, and Harbers and Ingram (2017) provide comprehensive guidance for using GWR to inform case selection in mixed-methods research designs.²²

Spatiotemporal models examine whether a spatial lag has a meaningful effect over time, offering spatial variants of the increasingly popular time-series cross-sectional analyses in research on subnational politics (Harbers, 2014; Giraudy, 2010; Ingram, 2013, 2016). A wide range of spatial panel or “space-time” models are available (e.g., Darmofal, 2015, chap. 8).

Finally, spatial or geographic regression-discontinuity designs (GRDs) rely on a geographic boundary that “splits units into treated and control areas and analysts make the case the division into treated and control areas occurs in an as-if random fashion” (Keele & Titiunik, 2015, p. 128), approximating randomized control trials, widely viewed as the “gold-standard” for causal inference. GRDs are thus a type of natural experiment that leverages the increasingly popular regression-discontinuity (RD) approach for causal identification, and the more convincing the case that the boundary occurred in an as-if random manner the more compelling the results. GRDs also resemble GWR in that GRDs estimate local effects on either side of a boundary whereas GWR estimate local effects for each unit (Keele & Titiunik, 2015, p. 152). To be

²² GWR can be implemented in R with packages *spgwr* (Bivand & Yu, 2017), *gwrr* (Wheeler, 2013), or *GWmodel* (Gollini et al., 2013), or with a variety of standalone software packages (e.g., GWR4).

sure, GRDs are difficult to execute well, even more so than non-geographic RDs. Some of the assumptions of RDs are less likely to hold in a geographic context, and geographic boundaries often overlap with other bound and unbound phenomena, generating “compound treatments” which can make causal identification tricky (Keele & Titiunik, 2015, p. 133).²³

Moreover, a tension exists between (a) GRD’s emphasis on causal identification and (b) the goal of taking spatial interdependence seriously, as advocated throughout this chapter. Specifically, the experimental logic of GRDs rests on the assumption of the independence of units and also the stable unit treatment value assumption (SUTVA), which in part holds that there is no interference among units or, more precisely, that treatment in one unit does not affect outcomes in another unit. This assumption is violated if there is any feedback, spillover, transfer, or diffusion between units. Thus, the identification or anticipation of any spatial interdependence undermines the logic of GRDs. As Franzese and Hays observe succinctly regarding propensity-score matching techniques, “if interdependence, then not SUTVA” (2008a, p. 760). This tension between the goal of improving causal identification by using a quasi-experimental method like GRD, which assumes independence of units, on the one hand, and the goal of strengthening our understanding of how spatial dependence influences key phenomena of interest to social scientists, especially at subnational levels, on the other, suggests the fruitfulness of future research that focuses on how to combine GRD with MAUP (see Soifer’s Chapter 3 in this volume), changing levels and scales of analysis, bound and unbound units, the nature of boundaries, the nature of spatial interaction, and GRD. Lastly, the effective use of GRDs in subnational research will also likely require strong local knowledge and substantive expertise (Keele & Titiunik, 2015, p. 128). This makes GRDs an especially promising tool for mixed-methods approaches that combine statistical analysis with in-depth fieldwork and also for collaborative research between GRD and substantive experts. As with many of the spatial econometric techniques presented here, their value and richness hinges on substantive knowledge of the subject matter and its spatial features.

2.4 CONCLUSION

This chapter develops a more self-conscious consideration of the role of space in the study of subnational politics, organizing the discussion across three core

²³ Keele and Titiunik use the example of media markets in the United States (citing Huber and Arceneaux [2007] and Krasno and Green [2008]), noting that the boundaries of these markets tend to overlap with county and, at times, state boundaries (Keele & Titiunik, 2015, p. 134). Researchers who want to examine the as-if random exposure to political ads from different media markets in the United States would thus need to consider the compound treatment of media market, county, and potentially state.

stages of research design. In the conceptualization stage, attention to space helps both to distinguish unbound from institutional phenomena and to select an appropriate subnational level of analysis. In the theorizing stage, attention to space helps clarify how structural dependence shapes outcomes and causal relationships of interest. Specifically, we distinguish between place-based and propagation-based processes to help clarify causal propositions about phenomena of interest. Lastly, in the analysis phase, a spatial perspective helps to identify spatial interdependence in the outcome of interest and also to differentiate among types of spatial dependence (e.g., spatial lag vs. spatial error). Summing up, taking space more seriously promises valuable insights for conceptualizing, theorizing, and analyzing subnational politics.

Although the spatial perspective arrived later in political science than in other social sciences, it is an exciting and welcome development because it explicitly addresses the dependent structure of the data which scholars in this field necessarily encounter. All data are embedded in some larger structure, which is why international relations scholars also find spatial analysis so useful. The spatial perspective is especially important in subnational research because the strength and density of the spatial dependence of observations is likely to be far stronger within a single country than across countries. Spatial tools and analysis thus hold great conceptual, theoretical, and empirical promise for subnational research in comparative politics.

APPENDIX

Constructing Spatial Weights Matrices and Spatial Lags

The operationalization of spatial weights is a key step in spatial analysis. Put simply, spatial weights capture the nature of connections among units. From a purely geographic standpoint, these weights capture the distance among units. So, for instance, a neighbor that is directly adjacent to unit A is deemed to be “contiguous.”

Software packages for spatial analysis include various options for assessing contiguity or, more broadly, connectedness. Classic options include *rook contiguity* and *queen contiguity*. As the names imply, “rook” and “queen” are references to how these pieces can move in the game of chess. From the perspective of rook contiguity (Panel a in Figure 2.A1), unit 6 has four neighbors: 2, 5, 7, and 10. The weights matrix (Figure 2.A2) reflects these neighbors as 1s when reading across (or down) from 6. Note that units along boundaries have only three neighbors, and units in the corners have only two neighbors.

Queen contiguity (Panel b in Figure 2.A1) implies a more expansive definition of neighbors, including those units with which 6 shares only a vertex. Thus,

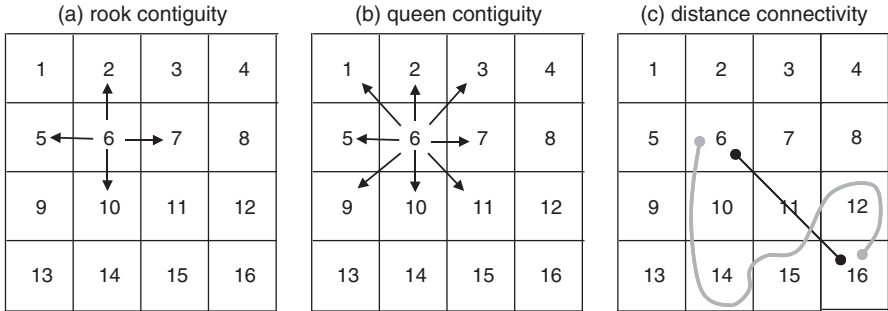


FIGURE 2.A1 Types of Contiguity in Spatial Analysis

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0
2	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0
3	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0
4	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
5	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0
6	0	1	0	0	1	0	1	0	0	1	0	0	0	0	0	0
7	0	0	1	0	0	1	0	1	0	0	1	0	0	0	0	0
8	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0
9	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0
10	0	0	0	0	0	1	0	0	1	0	1	0	0	1	0	0
11	0	0	0	0	0	0	1	0	0	1	0	1	0	0	1	0
12	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	1
13	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0
14	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1	0
15	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1
16	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0

FIGURE 2.A2 Corresponding Spatial Weights Matrix (W) for Rook Contiguity

from the perspective of queen contiguity, unit 6 has eight neighbors: 1, 2, 3, 5, 7, 9, 10, and 11. Again, the corresponding weight matrix for queen continuity (not depicted here) reflects these units as 1s and all other units as 0s. Variations of both rook and queen contiguity can expand to include adjacent territorial units that may be two or more “neighbors” away. For instance, in Panel b of Figure 2.A1, rook contiguity-2 would add units 8 and 14 to the list of unit 6’s neighbors.

Further, spatial weights matrices can also be specified using distance metrics. In these cases, the weights matrix is no longer binary but can include ordinal or continuous measures. One of the commonly used options is a spatial weights matrix based on Euclidean distance – the straight-line distance between units. Panel c in Figure 2.A1 illustrates this option as the straight line between units 6 and 16. Based on geographic coordinates (latitude and longitude), this distance is relatively straightforward to calculate. For very large spaces (e.g., regional or global analyses), calculations adjust for the curved surface of the earth to generate a more accurate measure of distance, the geodesic distance. Thus, both Euclidean and geodesic distances can be used to generate weight matrices different than those based on rook or queen contiguity. These distances can also be truncated at certain thresholds so that no effect is felt beyond a certain distance.

Yet another option raised in the measurement section is the possibility of generating additional spatial weights based on geographic features, georeferenced infrastructure, or other theoretically relevant aspects of the terrain. For instance, road infrastructure may be a more theoretically relevant way of measuring connectedness than either rook and queen contiguity or Euclidean (or geodesic) distance. This point is illustrated by the curving grey line in Panel c in Figure 2.A1, which represents a hypothetical road. Note that to get from unit 6 to 16 via the road, one must first pass through units 10, 14, 15, and 12. Several relevant implications flow from this fact. First, while queen contiguity and Euclidean distance would register unit 11 as being adjacent or close to unit 6, the road distance registers units 14 and 15 as closer than 11. Moreover, while queen contiguity and Euclidean distance would show unit 16 as the equivalent of two units away from unit 6, road distance shows these two units to be the equivalent of five units apart! Road networks are readily available in GIS format in many countries and can also be used to calculate travel time between observations. Further, other geographic features that are also readily available in GIS may be theoretically relevant depending on the research question (e.g., rivers and waterways, telecommunications infrastructure, sewage and other public utility infrastructure). For recent work employing a road network to construct spatial weights, see Zhukov (2012).

Regardless of how the spatial weights matrix is constructed, it is generally standardized by summing across each row and then dividing each element in the row by this row-sum. For instance, in the above rook contiguity matrix W , the sum of row 1 is 2. The row-standardized spatial weights matrix is then represented as follows in Figure 2.A3.

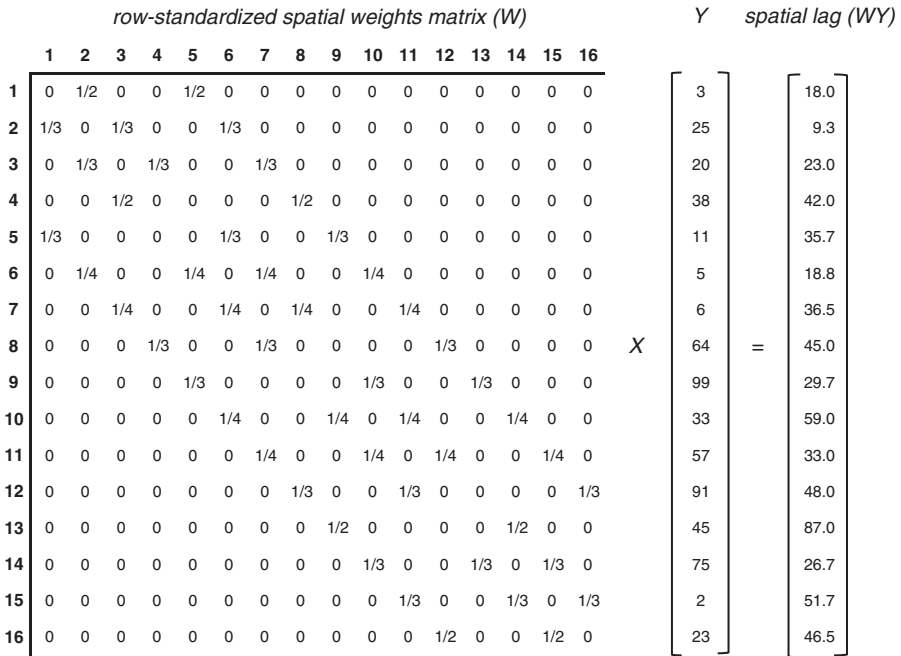


FIGURE 2.A3 Row-Standardized Spatial Weights Matrix (W) and Spatial Lags for Rook Contiguity

To generate the spatial lag for each observation, we then multiply the row-standardized matrix (16x16) by a vector Y (16x1) of a hypothetical outcome of interest in each relevant unit. Doing so for each observation (1–16) generates a vector of spatially weighted lags, or spatial lags (16x1). Note that the weight matrix and resulting spatial lags would look very different if the matrix were constructed based on queen contiguity, Euclidean distance, or road distance. How spatial closeness or proximity is conceptualized, and therefore how the spatial weights matrix is constructed, hinges ultimately on the research question and relevant theory. For an in-depth discussion of conceptualizing and constructing spatial weights, see Beck, Gleditsch, and Beardsley (2006), Darmofal (2015), and Neumayer and Plümper (2016)

BIBLIOGRAPHY

Anselin, Luc. (1988). *Spatial Econometrics: Methods and Models*. Dordrecht, Netherlands: Springer Science-Business Media.
 Anselin, Luc. (1995). Local indicators of spatial association – LISA. *Geographical Analysis* 27(2), 93–115.

- Anselin, Luc, & Bera, A. K. (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics. In A. Ullah & D. E. A. Giles (Eds.), *Handbook of applied economic statistics*, pp. 237–289. New York, NY: Marcel Dekker.
- Anselin, Luc. (2005). *Exploring spatial data with GeoDa: A workbook*. Available at: https://geodacenter.asu.edu/og_tutorials (last accessed April 2, 2013).
- Assunção, R., & Reis, E. A. (1999). A new proposal to adjust Moran's I for population density. *Statistics in Medicine* 18(16), 2147–2161.
- Baller, Robert D., Anselin, Luc, Messner, Steven F., Deane, Glenn, & Hawkins, Darnell F. (2001). Structural covariates of U.S. County homicide rates: Incorporating spatial effects. *Criminology* 39(3), 561–590.
- Beck, Nathaniel, Gleditsch, Kristian Skrede, & Beardsley, Kyle. (2006). Space is more than geography: Using spatial econometrics in the study of political economy. *International Studies Quarterly* 50(1), 27–44.
- Beck, Nathaniel L., & Katz, Jonathan N. (1996). Nuisance vs. substance: Specifying and estimating time-series-cross-section models. *Political Analysis* 6, 1–36.
- Behrend, Jacqueline. (2011). The unevenness of democracy at the subnational level: Provincial closed games in Argentina. *Latin American Research Review* 46(1), 150–176.
- Bivand, Roger, & Piras, Gianfranco. (2015). Comparing implementations of estimation methods for spatial econometrics. *Journal of Statistical Software* 63(18), 1–36.
- Bivand, Roger, & Yu, Danlin. (2017). spgwr: Geographically Weighted Regression. R package version 0.6–31. Available at: <https://CRAN.R-project.org/package=spgwr> (last accessed December 7, 2018).
- Cederman, Lars-Erik, & Gleditsch, Kristian Skrede. (2009). Introduction to special issue on “disaggregating civil war.” *Journal of Conflict Resolution* 53(4), 487–495.
- Cliff, A. D., & Ord, J. K. (1981). *Spatial Processes: Models and applications*. London, England: Pion Ltd.
- Darmofal, David. (2006). The political geography of macro-level turnout in American political development. *Political Geography* 25(2), 123–150.
- Darmofal, David. (2008). The political geography of the New Deal realignment. *American Politics Research* 36(6), 934–961.
- Darmofal, David. (2015). *Spatial analysis for the social sciences*. New York, NY: Cambridge University Press.
- Deane, Glenn, Messner, Steven F., Stucky, Thomas D., McGeever, Kelly, & Kubrin, Charis. (2008). Not “islands, entire of themselves”: Exploring the spatial context of city-level robbery rates. *Journal of Quantitative Criminology* 24(4), 363–380.
- Doreian, Patrick. (1980). Linear models with spatially distributed data: Spatial disturbances or spatial effects? *Sociological Methods & Research* 9(1), 29–60.
- Doreian, Patrick. (1982). Maximum likelihood methods for linear models: Spatial effect and spatial disturbance terms. *Sociological Methods & Research* 10(3), 243–269.
- Dow, Malcolm M. (2007). Galton's problem as multiple network autocorrelation effects. *Cross-Cultural Research* 41(4), 336–363.
- Elhorst, J. Paul. (2010). Applied spatial econometrics: Raising the bar. *Spatial Economic Analysis* 5(1), 9–28.
- Fearon, James D., & Laitin, David D. (2003). Ethnicity, insurgency, and civil war. *American Political Science Review* 97(1), 75–90.

- Fotheringham, A. Stewart, Brunsdon, Chris, & Charlton, Martin. (2002). *Geographically weighted regression: The analysis of spatially varying relationships*. Chichester, England: Wiley and Sons.
- Fotheringham, A. Stewart, Charlton, Martin, & Brunsdon, Chris. (1998). Geographically weighted regression: A natural evolution of the expansion method for spatial data analysis. *Environment and Planning A* 30(11), 1905–1927.
- Franzese, Robert Jr., & Hays, Jude C. (2008a). Spatial interdependence in comparative politics: Substance, theory, empirics, substance. *Comparative Political Studies* 41(4/5), 742–780.
- Franzese, Robert Jr., & Hays, Jude C. (2008b). Contagion, common exposure, and selection: Empirical modeling of theories and substance of interdependence in political science. *Concepts & Methods* 4(2), 2–8.
- Gervasoni, Carlos. (2010). Measuring variance in subnational regimes: Results from an expert-based operationalization of democracy in the Argentine provinces. *Journal of Politics in Latin America* 2(2), 13–52.
- Gibson, Edward L., & Suárez-Cao, Julieta. (2010). Federalized party systems and subnational party competition: Theory and an empirical application to Argentina. *Comparative Politics* 43(1), 21–39.
- Giraudy, Agustina. (2010). The politics of subnational undemocratic regime reproduction in Argentina and Mexico. *Journal of Politics in Latin America* 2(2), 53–84.
- Giraudy, Agustina. (2013). Varieties of subnational undemocratic regimes: Evidence from Argentina and Mexico. *Studies in Comparative International Development* 48(1), 1–30.
- Gleditsch, Kristian Skrede, & Ward, Michael D. (2005). Visualization in international relations. In Alex Mintz & Bruce Russett (Eds.), *New directions for international relations: Confronting the method-of-analysis problem*, pp. 65–91. Lanham, MD: Lexington Books.
- Gleditsch, Kristian Skrede, & Weidmann, Nils B. (2012). Richardson in the information age: Geographic information systems and spatial data in international studies. *Annual Review of Political Science* 15(1), 461–481.
- Gleditsch, Kristian, & Ward, Michael. (2000). War and peace in time and space: The role of democratization. *International Studies Quarterly* 44(1), 1–29.
- Goertz, Gary. (2006). *Social science concepts: A user's guide*. Princeton, NJ: Princeton University Press.
- Gollini, Isabella, Lu, Binbin, Charlton, Martin, Brunsdon, Christopher, & Harris, Paul. (2013). GWmodel: An R package for exploring spatial heterogeneity using geographically weighted models. *arXiv:1306.0413*.
- Griffith, Daniel A. (1987). *Spatial autocorrelation: A primer*. Washington DC: Association of American Geographers.
- Hafner-Burton, Emilie M., Kahler, Miles, & Montgomery, Alexander H. (2009). Network analysis for international relations. *International Organization* 63(3), 559–592.
- Harbers, Imke, & Ingram, Matthew C. (2014). Democratic institutions beyond the nation state: Measuring institutional dissimilarity in federal countries. *Government and Opposition* 49(1), 24–46.
- Harbers, Imke, & Ingram, Matthew C. (2017a). Geo-nested analysis: Mixed-methods research with spatially dependent data. *Political Analysis* 75(3), 289–307.

- Harbers, Imke, & Ingram, Matthew C. (2017b). Incorporating space in multimethod research: Combining spatial analysis with case-study research. *PS: Political Science & Politics* 50(4), 1032–1037.
- Harbers, Imke. (2014). States and strategy in new federal democracies: Competitiveness and intra-party resource allocation in Mexico. *Party Politics* 20(6), 823–835.
- Harbers, Imke. (2015). Taxation and the unequal reach of the state: Mapping state capacity in Ecuador. *Governance* 28(3), 373–391.
- Hilgers, Tina. (2011). Clientelism and conceptual stretching: Differentiating among concepts and among analytical levels. *Theory and Society* 40(5), 567–588.
- Hooghe, Liesbet, & Marks, Gary. (2016). *Community, scale and regional governance*. Oxford, England: Oxford University Press.
- Huber, Gregory A., & Arceneaux, Kevin. (2007). Identifying the persuasive effects of presidential advertising. *American Journal of Political Science* 51(4), 957–977.
- Huckfeldt, R. Robert, & Sprague, John. (1995). *Citizens, politics and social communication: Information and influence in an election campaign*. New York, NY: Cambridge University Press.
- Ingram, Matthew C. (2013). Elections, ideology, or opposition? Assessing competing explanations of judicial spending in the Mexican states. *Journal of Law, Economics, and Organization* 29(1), 178–209.
- Ingram, Matthew C. (2014). *The local educational and regional economic foundations of violence: A spatial analysis of homicide in Mexico's municipalities*. Mexico Institute Working Paper Series, Woodrow Wilson International Center for Scholars, Washington, DC.
- Ingram, Matthew C. (2016). *Crafting courts in new democracies: The politics of subnational judicial reform in Brazil and Mexico*. Cambridge, England: Cambridge University Press.
- Ingram, Matthew C., & da Costa, Marcelo Marchesini. (2017). A spatial analysis of homicide across Brazil's municipalities. *Homicide Studies* 21(2), 87–110.
- Keele, Luke, & Titiunik, Rocio. (2015). Geographic boundaries as regression discontinuities. *Political Analysis* 23(1), 127–155.
- King, Gary. (1997). *A solution to the ecological inference problem*. Princeton, NJ: Princeton University Press.
- Kopstein, Jeffrey, & Reilly, David A. (2000). Geographic diffusion and the transformation of the postcommunist world. *World Politics* 53(1), 1–37.
- Krasno, Jonathan S., & Green, Donald P. (2008). Do televised presidential ads increase voter turnout? Evidence from a natural experiment. *The Journal of Politics* 70(1), 245–261.
- Land, Kenneth C., McCall, Patricia L., & Cohen, Lawrence E. (1990). Structural covariates of homicide rates: Are there any invariances across time and social space? *American Journal of Sociology* 95(4), 922–963.
- Lankina, Tomila, & Getachew, Lullit. (2006). A geographic incremental theory of democratization: Territory, aid, and democracy in postcommunist regions. *World Politics* 58, 536–582.
- LeSage, James P., & Pace, R. Kelley. (2010). Spatial econometric models. In Manfred M. Fischer & Arthur Getis (Eds.), *Handbook of applied spatial analysis: Software tools, methods, and applications*, pp. 355–376. Berlin, Germany: Springer.
- Logan, John R., Zhang, Weiwei, & Xu, Hongwei. (2010). Applying spatial thinking in social science research. *GeoJournal* 75(1), 15–27.

- Manski, Charles F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies* 60(3), 531–542.
- Mears, Daniel P., & Bhati, Avitash S. (2006). No community is an island: The effects of resource deprivation on urban violence in spatially and socially proximate communities. *Criminology* 44(3), 509–548.
- Messner, Steven F., Anselin, Luc, Baller, Robert D., Hawkins, Darnell F., Deane, Glenn, & Tolnay, S. E. (1999). The spatial patterning of county homicide rates: An application of exploratory spatial data analysis. *Journal of Quantitative Criminology* 15(4), 423–450.
- Messner, Steven F., Teske, Jr., Raymond H.C., Baller, Robert D., & Thome, Helmut. (2011). Structural covariates of violent crime rates in Germany: Exploratory spatial analyses of Kreise. *Justice Quarterly*, 30(6), 1015–1041. doi:10.1080/07418825.2011.645862
- Moncada, Eduardo, & Snyder, Richard. (2012). Subnational comparative research on democracy: Taking stock and looking forward. *Newsletter of the Comparative Democratization Section of the American Political Science Association* 10(1), 1–9.
- Moran, Patrick Allen Pierce. (1948). The interpretation of statistical maps. *Journal of the Royal Statistical Society B*, 10, 243–251.
- Neumayer, E., & Plümpner, T. (2016). W. *Political Science Research and Methods* 4(1), 175–193.
- O'Donnell, Guillermo. (1999). *Counterpoints: Selected essays on authoritarianism and democratization*. Notre Dame, CA: University of Notre Dame Press.
- Parker, Robert Nash, & Asencio, Emily K. (2009). *GIS and spatial analysis for the social sciences: Coding, mapping and modeling*. New York, NY: Routledge.
- PySAL Developers. (2017). PySAL—Python Spatial Analysis Library. Release 1.13.0 (May 2017). Available at: <http://pysal.github.io/index.html> (last accessed June 14, 2017).
- Python Software Foundation. (2016). Python. Version 3.5.2. Available at: www.python.org/ (last accessed June 14, 2017).
- R Core Team. (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: www.R-project.org/ (last accessed December 10, 2018).
- Rey, S. J., & Anselin, L. (2007). PySAL: a python library of spatial analytical methods. *Review of Regional Studies* 37(1), 5–27.
- Rodrigues-Silveira, Rodrigo. (2013). *The subnational method and social policy provision: Socioeconomic context, political institutions and spatial inequality*. Working Paper No. 36. desiguALdades.net, Berlin, Germany.
- Rokkan, Stein. (1970). *Citizens, elections, parties*. Oslo, Norway: Universitetsforlaget.
- Sampson, Robert J., & Graif, Corina. (2009). Neighborhood social capital as differential social organization: Resident and leadership dimensions. *American Behavioral Scientist* 52(11), 1579–1605.
- Sartori, Giovanni. (1970). Concept misformation in comparative politics. *American Political Science Review* 64(4), 1033–1053.
- Sartori, Giovanni. (2005 [1976]). *Parties and party systems: A framework for analysis*. Colchester, England: ECPR Press.
- Schedler, Andreas. (2014). The criminal subversion of Mexican democracy. *Journal of Democracy* 25(1), 5–18.

- Snyder, Richard. (2001). Scaling down: The subnational comparative method. *Studies in Comparative International Development* 36(1), 93–110.
- Snyder, Richard, & Durán-Martínez, Angélica. (2009). Does illegality breed violence? Drug trafficking and state-sponsored protection rackets. *Crime, Law and Social Change* 52(3), 253–273.
- Tollefsen, Andreas Forø, Strand, Håvard, & Buhaug, Halvard. (2012). PRIO-GRID: A unified spatial data structure. *Journal of Peace Research* 49(2), 363–374.
- Trelles, Alejandro, & Carreras, Miguel. (2012). Bullets and votes: Violence and electoral participation in Mexico. *Journal of Politics in Latin America* 4(2), 89–123.
- Ward, Michael, & Gleditsch, Kristian Skrede. (2018). *Spatial Regression Models*. 2nd ed. Thousand Oaks, CA: SAGE Publications.
- Weidmann, Nils B., Kuse, Doreen, & Gleditsch, Kristian Skrede. (2010). The geography of the international system: The cshapes dataset. *International Interactions* 36(1), 86–106.
- Wheeler, David. (2013). gwrr: Fits geographically weighted regression models with diagnostic tools. R package version 0.2–1. Available at: <https://CRAN.R-project.org/package=gwrr> (last accessed December 10, 2018).
- Wickham, Hadley. (2009). *ggplot2: Elegant graphics for data analysis*. New York, NY: Springer-Verlag.
- Yang, Tse-Chuan, Noah, Aggie J., & Shoff, Carla. (2015). Exploring geographic variation in S.S. mortality rates using a spatial durbin approach. *Population, Space, and Place* 21(1), 18–37.
- Zhukov, Yuri M. (2012). Roads and the diffusion of insurgent violence: The logistics of conflict in Russia's North Caucasus. *Political Geography* 31(3), 144–156.