Chapter 11
Geospatial Analysis in Higher Education Research

Nicholas W. Hillman

Like many working class communities across the country, Elkhart, Indiana, was hit especially hard during the Great Recession. Elkhart is a rust belt community located in the northernmost part of the state on the border of Michigan. Elkhart county is home to over 200,000 citizens, making it the sixth-largest county in the state (U.S. Census Bureau, 2016a). The labor force is comprised mainly of manufacturing and construction jobs, so when the economy slowed down during the Great Recession, plants began to close and mass layoffs ensued, resulting in the state's highest unemployment rate and one of the highest in the nation – peaking at 20 % (Bureau of Labor Statistics, 2016). These layoffs were not for short durations of time – people were often out of work for more than six months and many even longer (Katz, 2014).

With thousands of displaced workers living in Elkhart County, this is a prime time for people to go back to college. People who are out of work now have more time to invest in education and those who were previously on the margins of going to college would be likely to now enroll because their opportunity cost is lowered. However, this did not occur in Elkhart County: colleges located in the county increased their enrollments by only 140 total students during the Great Recession. How could this occur? Why would a county with such deep unemployment not experience a large increase in college enrollments? Researchers can answer these questions from a number of angles, but this chapter focuses on one in particular – *geography*.

Geography is an often overlooked factor shaping educational opportunity in the United States, yet it is one of the strongest forces affecting whether and where students attend college. Most students stay relatively close to home when they attend college; in fact, more than half of all college students enroll within just 20 miles of

---

N.W. Hillman (✉)
University of Wisconsin-Madison, 1000 Bascom Mall, Room 249, Madison, WI 53706, USA
e-mail: nwhillman@wisc.edu

M.B. Paulsen (ed.), *Higher Education: Handbook of Theory and Research*,
Higher Education: Handbook of Theory and Research 32,
DOI 10.1007/978-3-319-48983-4_11

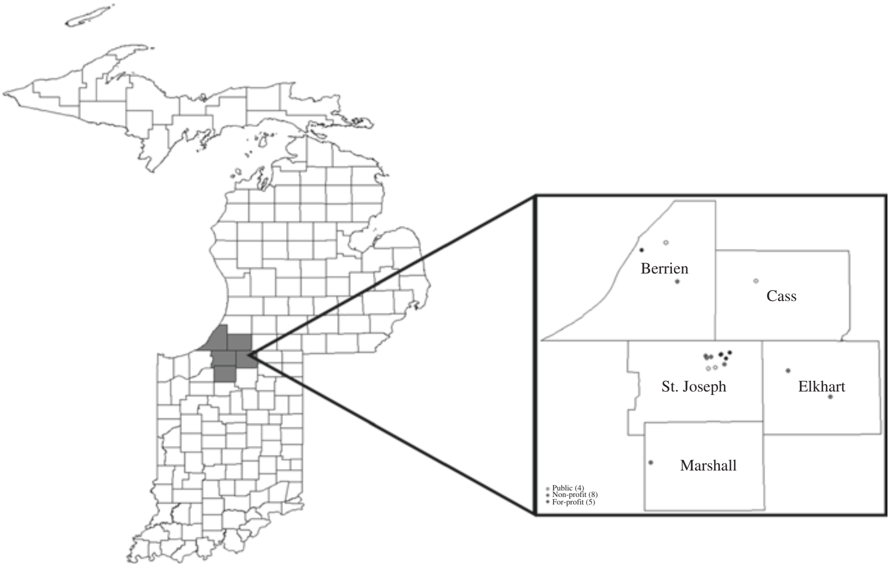

Fig. 11.1 Dot map of the South Bend-Mishawaka-Elkhart common statistical area

their permanent home address (Sponsler & Hillman, 2016). The image of a highly-mobile college student who conducts a national or even statewide search for college is the exception to the rule. It would be unlikely for a working class family living in Elkhart County to uproot and move long distances simply to attend college. This is the case nationwide, where today's college students are place-bound and for a number of reasons need to choose a college that is in close proximity to home and work.

Figure 11.1 displays where colleges are located in Elkhart and its neighboring counties. These five counties are not arbitrarily selected for this example; rather they have a "high degree of social and economic integration" and are part of a common statistical area where people cross over county and state lines on a daily basis (U.S. Office of Management and Budget, 2015). Clustering counties together in this way can help us observe a more complete picture of the geography of college opportunity in Elkhart. These five counties are commonly referred to as "Michiana" because they include areas both in Michigan and Indiana. Michiana contains a total of 17 colleges and universities reported in the U.S. Department of Education's Integrated Postsecondary Education Data System (IPEDS), including four public, eight private non-profits, and five private for-profit institutions.

Together, these institutions enroll approximately 53,000 total students with just over half (52 %) enrolled in the public sector. The largest institutions are the University of Notre Dame and Indiana University-South Bend (IUSB), which are both located in the area's largest county: St. Joseph County. Since most of the area's colleges are located in St. Joseph County, it is perhaps not surprising that the majority (63 %) of Michiana's college students are also enrolled in St. Joseph County. In both cases, St. Joseph County is the central place that attracts students from the

Michiana area. If a displaced worker living in Elkhart County wanted to enroll in a local college to retrain for the labor force, she would likely need to commute to St. Joseph County. But the distance between the population centers of these two counties is 21 miles, or a 30-min drive and there is no public transportation connecting the two communities. A displaced worker may find it difficult to afford the gasoline, car maintenance, and lost time by driving from home and school. Alternatively, Elkhart residents could choose to go across the state line to a nearby college in Michigan, but in addition to the same transportation barriers they would also face higher tuition charges because they would be charged out-of-state rates.

This example of Elkhart County and the Michiana geographic area illustrates a number of important features of local higher education marketplaces that can affect whether and where students attend college. It shows how the local supply of colleges (i.e., number, selectivity, sector, etc.) can affect educational opportunities. It demonstrates how distance from home to school can affect student enrollment demand, with further distances being associated with a lower likelihood of enrolling. It illustrates discontinuities that are created by state boundaries, where students on one side of a state line are charged different tuition rates than those living on the other side of the line. This chapter will explore such issues and examine how geography and place can be further integrated into higher education research. In so doing, higher education research can expand the way it engages with geography, both conceptually and empirically.

This chapter introduces readers to important research studies, data sources, and analytical techniques in order to integrate geography more in higher education scholarship. The example of Elkhart County and Michiana will be referenced occasionally, but new examples will also be introduced to illustrate key points. Also highlighted will be the ways similar phenomenon occur in communities across the country because the topics addressed in this chapter are not limited to this single rust-belt community. Similar geospatial patterns emerge across the country with each community having its own unique higher education marketplace that can be better understood employing a geospatial lens.

Goals of This Chapter

The primary goal of this chapter is to encourage researchers to find new and creative ways of applying geospatial analysis in higher education research. It provides a review of the existing research that has engaged with geospatial analysis and it outlines promising areas for further research. Even if readers do not plan to use geospatial analysis in their work, this chapter is written to introduce readers to key geographic concepts and strategies that could prove fruitful in advancing our understanding of college access and inequality. For researchers interested in incorporating geospatial analysis into their work, the chapter identifies existing data sources and software packages for conducting geospatial analysis. Also discussed are promising methods and empirical challenges often encountered when conducting geospatial analysis. The chapter is not intended to be an exhaustive or advanced treatment of geospatial analysis; rather, it offers an introduction to key concepts, data sources, and research strategies to expose and encourage readers to pursue geospatial analysis in higher education research.

In addition, the content herein will engage federal and state policies by discussing how geography interacts with various policy decisions. For example, the federal government has recently invested in a number of efforts designed to improve consumer information so students can make better decisions about where to enroll in postsecondary education. Embedded in the policy logic is the belief that students "shop around" for college and decide where to enroll based on their assessment of information such as tuition levels, graduation rates, financial aid availability, and job-placement rates. The College Scorecard, Financial Aid Shopping Sheet, Postsecondary Institution Rating System (PIRS) and College Navigator are examples of ongoing federal efforts to improve consumer information in the college marketplace. While students may use this information during their choice-making process, knowing that a college hundreds of miles away offers a better academic fit or more generous financial aid may be of little use to place-bound students. Even a perfectly informed student, living somewhere like Elkhart County, may have constrained choices if they are place-bound. Reflecting on these policy topics from a geospatial perspective may help advance public policy efforts to reverse educational inequality and to identify differences that exist in the postsecondary marketplace.

Geospatial analysis can be applied in many areas of higher education research, particularly around the topic of college choice. We still have much to learn about the extent to which proximity to college shapes students' decisions to prepare for, apply, enroll, and persist to degree completion. But geospatial analysis should extend far beyond this often studied area of inquiry. It can also help the field develop new questions, and answers, about local higher education markets and how colleges in the same market compete for students, set their prices, respond to local labor market demands, and create spillovers to their local communities. Geospatial analysis can also be useful for designing quasi-experimental studies to evaluate the impacts of policy interventions and even for merging higher education data with other georeferenced data sources that have yet to be fully explored. Applying a geospatial lens may help researchers find new ways of thinking about colleges as part of their local "built environment," where critical geography and theories of spatial inequality can offer new lines of inquiry for the field. Geospatial analysis can even deliver more user-friendly data visualizations to help researchers explain complex educational phenomenon, or be used to detect patterns in data that may have otherwise been overlooked. There are several reasons why integrating geospatial analysis more closely into the field of higher education is important, and this chapter's main goal is to provide examples and identify promising strategies to help advance this line of inquiry.

What Is Geospatial Analysis?

The term "geospatial analysis" is becoming increasingly popular in social science research, driven in large part by the growing number of software platforms and geocoded data available for research. But geospatial analysis has been around for a long time; some believe its origins began in 1855 when John Snow created the famous Broad Street Map that linked London's fatal cholera outbreak to contaminated water supplies (Snow, 1855). Prior to Snow's map, people did not know what caused cholera to spread and some hypothesized it was an airborne illness. To test his hypothesis that water and not air was spreading the disease, Snow first created a street map of the most afflicted neighborhood and then documented the number of cholera outbreaks for each home address. He then layered the location of water pumps to this map and was able to use statistical analysis to demonstrate that the outbreaks clustered non-randomly around certain water pumps that were subsequently found to be the source of the cholera problem.

The idea of layering social data (e.g., cholera cases) with geographic data (e.g., street maps) is now ubiquitous, with Google Maps and various social media platforms integrating maps and data on social activity. Internet access allows people to easily view the location of a particular place and its surrounding areas, traffic and weather patterns, time it will take to get to a location, and even recommendations for other activities or points of interest located nearby. Layering geographic data in this way has clear benefits for consumers, but in the case of Snow's maps it also served a public health purpose and provided insights into how people interact with their local environment.

For this chapter, geospatial analysis focuses on the various ways in which human behaviors shape, and are shaped by, their geographic environment. More specifically, geospatial analysis includes the statistical analysis of geographically-referenced data (Murayama, 2012). Common geographical references include street addresses, ZIP codes, Federal Information Processing Standards (FIPS) codes for counties or states, as well as latitude and longitude. When researchers include georeferenced data in a statistical analysis, then they are stepping into the realm of geospatial analysis. They can then estimate whether the distance between points is a significant predictor of a particular outcome, or they can examine whether patterns found in one geographic area differ significantly from the patterns observed in another area. They can also use georeferenced data to produce maps that are either static or interactive that can be used as a communication tool for sharing research findings.

Geospatial analysis is a subfield within the social sciences that, at its core, seeks to understand how human behavior interacts with its geographic environment. Geospatial analysis and human geography more broadly can be found across a wide range of academic disciplines and is not constrained to the field of geography. The "Chicago School" of sociology has a tradition tied directly to geography, where scholars study urban environments to understand how the social structures of a particular community shape individual and group behaviors (Fyfe & Kenny, 2005).

Sociologists and anthropologists may also engage with geography to understand why neighborhoods are segregated or how the location of industries shape people's life chances (D. Massey, 2005; D. S. Massey & Denton, 1988). Demographers examine population mobility and density to understand how people move across different regions of the country (Rosenbaum, Reynolds, & DeLuca, 2002; Molloy, Smith, & Wozniak, 2011). Epidemiologists use geography to document and anticipate how viruses or other public health problems spread throughout communities (Elliott, 2000; Meade, 2010). Economists have used geography to model natural experiments, where policy changes affecting one geographic region (but not another) can help explain the impacts of a particular policy decision (Card & Krueger, 2000; Dynarski, 2002).

Incorporating geospatial analysis into higher education can be helpful in conducting research and developing theories, but it can also aid in communicating research findings. By incorporating maps and geospatial analysis into higher education research, the field can become even more interdisciplinary in ways that push both theory and research design. There are several more ways to engage with geospatial analysis and this chapter offers an introduction to promising approaches.

How Geography Shapes Higher Education Markets

Geography and "place" have always been relevant to higher education in the United States. From the development of Land-Grant colleges and the creation of a dual system of segregated higher education of the nineteenth century to the expansion of community colleges in the twentieth century, the site selection and location of colleges has been a subject of both policy and scholarly attention. This section highlights some of the key moments in that history, emphasizing how the supply of colleges has changed over time and the way geography shapes students' demand for college. By discussing the location of colleges and how proximity shapes students' choices, we can gain greater insight into the marketplace of higher education. If every community had an equal mix of colleges in terms of quality, accessibility, and affordability, then geography would not bear on this marketplace. There would be little to no geographic friction because prospective students would have equal opportunity structures nearby.

However, states developed their public higher education systems very unevenly over time, with those that were newly admitted to the Union in the nineteenth Century being more likely to have larger public systems than the eastern states of the colonial era (Goldin & Katz, 1998). Similarly, the vestiges of segregation and Jim Crow policies of the South through the early twentieth century created dual and unequal public higher education systems that states are still remediating today (Gasman & Hilton, 2012). These historical differences matter because they helped shape today's higher education marketplace where each state has a differing mix of public two-year and four-year colleges operating alongside private non-profit and for-profit colleges. As a result, states will likely face their own unique circumstances and challenges related to geography depending on their own historical context.

The Supply of Colleges

Table 11.1 displays the number of public, non-profit, and for-profit colleges in the United States and the share of students enrolled in each sector. Whereas the majority of colleges are private, the majority of students enrolled in college attend public institutions. Community colleges enroll more students than public four-year colleges, suggesting that geography may play a more important role in public institutions given that the former were constructed to serve their local community's educational and workforce needs.

This is not to say geography is irrelevant to private colleges; it certainly matters, especially for minority serving institutions. But private institutions are more likely to offer exclusively online education and to recruit from broader geographic regions than are public institutions. For example, seven percent of public college students enroll exclusively online whereas 13 and 29 % of private non-profit and for-profit students (respectively) enroll exclusively online. Technology may be a useful supplement to in-person learning, but research to date shows enrolling exclusively online has more negative than positive outcomes (Alpert, Couch, & Harmon, 2016; Bettinger, Fox, Loeb, & Taylor, 2015; Bulman & Fairlie, 2016; Figlio, Rush, & Yin, 2013). Despite recent growth in distance education (particularly within for-profit colleges), the location of colleges relative to population centers is fundamental to understanding the higher education marketplace.

One of the key moments in expanding the supply of higher education came via the 1862 and 1890 Morrill Acts. The first Act granted states 30,000 acres of federal public land for each representative and senator, so long as the state created "at least one college where the leading object shall be…to promote the liberal and professional education of the industrial classes in the secular pursuits and professions of life"

Table 11.1 Number of U.S. colleges and 12-month and exclusively online enrollment (in millions), by sector

	Public		Private		Total
	Four-year	Two-year	Non-profit	For-profit	
Number of colleges	689 (15 %)	934 (20 %)	1652 (35 %)	1451 (31 %)	4726 (100 %)
12-month enrollment (mil.)	9.7 (35 %)	10.1 (37 %)	4.8 (17 %)	2.9 (10 %)	27.5 (100 %)
Exclusively online enrollment (mil.)	0.7 (24 %)	0.7 (24 %)	0.6 (21 %)	0.8 (30 %)	2.8 (100 %)

Source: U.S. Department of Education (2015). *Digest of Education Statistics*, Tables 308.10, 311.15, and 317.10

Note: Row percentages in parentheses

(Key, 1996). Prior to these Acts, the federal government would routinely and strategically sell land as a way to generate revenue, but the Morrill Land Grant Acts granted land with the expectation that the creation of a university would generate economic spillovers that ultimately would achieve the same end. Research would eventually document these spillovers, where county-level population density and labor productivity increased significantly after creating these universities (Liu, 2015).

One of the earliest reports on the supply of colleges and their ties to economic activity came in 1917 with the Commission on the Distribution of Colleges. Due in large part to the expansion of railroads that occurred during the latter part of the nineteenth century, the commission argued that colleges should be closely tied to economic activity that takes place along railways. As the railway expanded, so did the number of colleges and universities and the Commission ultimately recommended that "new institutions should always be located on the main lines of travel, near centers of population" (Thomas, Hughes, & McConaughy, 1921). This rationale and the idea of using colleges to stimulate local economic development was common in the late nineteenth and early twentieth centuries during periods of westward expansion, industrialization, and migration (Brown, 1995).

The idea that colleges can generate regional or even statewide or national economic spillovers is a longstanding argument for the public support of higher education. This economic rationale dates as far back as the European Middle Ages. Europe's first university, University of Bologna, was created in 1088 and another 50 European universities emerged over the next 400 years of the Middle Ages (Cantoni & Yuchtman, 2014). The establishment and growth of medieval universities was both a cause and consequence of commercial activity, where universities attracted people to move (or stay) in a particular city while they also produced new human capital for cities to grow (De Long & Shleifer, 1993; North & Thomas, 1973). Using Germany as a case study, Cantoni and Yuchtman (2014) find that specialized training in law that occurred in German universities helped establish new markets in cities located near the university. The economic spillover of having a highly trained workforce was that cities near universities were able to flourish economically because "legally-trained administrators and judges reduced the cost of establishing markets for cities and territories" (Cantoni & Yuchtman, 2014, p. 878).

More recently, a similar economic rationale has been applied to the expansion of American community colleges. In the late 1960s, there was just over 600 community colleges nationwide, but within a decade that number grew to over 900 (U.S. Department of Education, 2016). Doyle and Gorbunov (2011) examined why this growth occurred, hypothesizing it could be a function of economic demand, social stratification, political ideology, or existing structures of the state's higher education marketplace (i.e., organizational ecology). The authors conclude that community colleges expanded primarily because of economic demands, which is consistent with prior examples of economic spillovers. They also found the existing number and mix of colleges in a given state shapes whether to establish new community colleges – those with more four-year colleges experienced slower growth in their community college systems.

But the location of colleges is not merely an extension of economic reasoning. It also interacts with race and class in important ways. Southern states denied educational opportunities to Black people through Jim Crow policies that have legacies today where public Historically Black Colleges and Universities are still remediating the vestiges of segregation (Gasman & Hilton, 2012). And when states have created new public colleges and universities, they have done so in uneven ways that may actually reinforce inequality by locating well-resourced colleges near white neighborhoods and locating poorly-resourced ones near communities of color (Briscoe & De Oliver, 2006; Green, 2010). These are examples of what Massey (2005) calls the "spatial division of labor" and why Soja (2010) argues geography cannot be disentangled from politics of public policymaking. Geographical differences in the supply of colleges are outcomes of both political and economic processes.

Student Enrollment Demand

When students decide where to attend college, geography plays an important but under-examined role. College choice theories have long included geography as a factor shaping students' enrollment decisions, where Manski (1983), Hossler and Gallagher (1987), Cabrera and La Nasa (2000), and Perna (2006) all include some measure of distance or location in their theoretical models. These theories explain that students' choices are subject to what geographers call "friction of place" where the further one lives from a college and the mix of colleges located nearby will shape students' enrollment decisions. This friction can occur in any stage of the college choice process – it can affect students' predispositions, search process, and ultimately their choice if they are admitted to multiple institutions.

Students are more likely to cast a wide net when searching for colleges, but as the time to enroll approaches prospective students tend to narrow their lists down to colleges located closer to home (Astin, 1980). While this pattern generally holds today, some argue that geography now plays a weaker role on student choices because transportation costs and rapid growth in technological innovation make it easier and more affordable to be mobile (Hoxby, 1997; Long, 2004). From this perspective, the college-choice process may be more about helping students find a college that is well-matched to their academic performance, with some arguing that students should apply to and enroll in the most selective college possible. When students attend a less-selective institution than they could be admitted to, their "undermatch" can have material consequences by reducing their likelihood of graduating and by increasing the net price students pay (Howell & Pender, 2016; Hoxby & Avery, 2013).

Over 40 % of undergraduates enroll in a less-selective college than what they could have been admitted to, indicating a large degree of "undermatch" taking place in the higher education marketplace (Smith, Pender, & Howell, 2013). To have all students make optimal choices based on undermatching theories would require

Table 11.2 Undergraduate distance from home address to college

	Public		Private		Total
	Two-year	Four-year	Non-profit	For-profit	
Number of miles (mean)	32	84	252	283	106
Number of miles (median)	8	18	43	18	13
Percent enrolled < 20 miles away	78 %	52 %	38 %	51 %	61 %

Source: U.S. Department of Education (2016). National Postsecondary Student Aid Survey: 2012. National Center for Education Statistics: Washington, DC

Note: Data uses excludes students enrolled exclusively online (ALTONLN2)

nearly half of all undergraduates to sort into different colleges. There are promising strategies to help students find better college matches, including waiving college application fees, encouraging more students to take the SAT or ACT exam, and delivering personalized guidance in the college-choice process (Avery, Howell, & Page, 2014; Bettinger & Baker, 2014; Castleman, Page, & Schooley, 2014; Smith, 2014). However, many students need to stay close to home because of familial or work responsibilities that may make it difficult to "shop around" for a better college match. Today's college students are increasingly "non-traditional" and the friction of place may matter even more for students with family or work responsibilities. For these students, the proximity to college – and not information about how well matched they are – may be the most important factor in deciding whether or where to attend. Turley (2009) argues that existing college choice theories are not sufficiently nuanced for these students and that researchers should "stop treating the college choice process as though it were independent of location and start situating this process within the geographic context in which it occurs" (p. 126). Because of this, geography can help us build on theories of college choice that examine how distance and proximity to college shape educational destinations.

Table 11.2 uses 2012 National Postsecondary Student Aid Study data to show the mean and median distance between the student's permanent home address and their college location. The first row indicates the mean distance for each sector, where those attending public two-year and four-year colleges live 32 and 84 miles away, respectively. Students enrolled in the private sector enroll much further away, and the mean distance for all undergraduates is approximately 106 miles. Outliers and large values can skew these means upward, so the second row shows median distances and reveals even more interesting patterns regarding how far students travel to college. Using medians reveals that college enrollment decisions are even more localized than we might think: the median distance between home and college is only 13 miles. In fact, approximately 61 % of undergraduates enroll in a college less than 20 miles away (third row in this table). Theories of college student choice and models of student enrollment demand should continue to engage with geography because for many students, choices occur in local (not national) marketplaces.

Supply and Demand in Higher Education Markets

From medieval universities to the modern community college, geography has played a fundamental role in shaping educational opportunity and the life chances of people living near colleges. These outcomes are shaped both by the supply of colleges and students' demand for a college education. Traditional college choice theories focus on the process of selecting a college where students first develop predispositions, then search for a college, and finally make a choice after weighing the perceived and real costs and benefits (Cabrera & La Nasa, 2000). But college choice is also a function of geography because the location of colleges may carry more weight in the choice-making process, particularly for students who need to stay close to home for college. In this case, students may develop college-going predispositions and conduct a well-informed search, yet still have their choices constrained by what is available nearby.

On the supply side, simply having a college nearby can induce more students to enroll and can result in raising local educational attainment levels (Card, 1993; Kling, 2001; Rouse, 1995). On the demand side, the monetary and non-monetary costs associated with enrolling are reduced for students which makes it more likely for prospective students to invest in human capital (Griffith & Rothstein, 2009; Turley, 2009). When researchers incorporate geography into their studies, it is possible to discover new ways of understanding and explaining the evolving nature of the higher education marketplace. Students living in a particular geographic area will have different opportunity structures than those in other areas. Similarly, students will be affected differently by geography and place when some need to stay close to home, whereas others may travel far distances for college. These topics will be explored in the following section, but the main purpose is that by integrating geography into higher education research, we not only learn about how distance shapes behaviors but we also learn about variations in local higher education marketplaces.

Geospatial Analysis in Higher Education

There are four primary ways researchers integrate geography into the field of higher education: descriptive maps, quasi-experimental design, distance elasticity, and geostatistics. This section offers a brief introduction to these research strategies by providing examples of how researchers have employed the techniques in their work and by demonstrating the underlying intuition behind each technique. In so doing, this section illustrates innovative ways to engage with geography in higher education research while offering ideas for researchers to incorporate geospatial analysis into their work. This discussion is not comprehensive, so additional literature is suggested for those who want to explore the topic in more detail.

Descriptive Maps

Choropleth Maps

Among the more common types of maps applied in higher education research is the choropleth (or area) map. This thematic map uses colors to represent a particular value or feature of a given area and can be an effective visualization tool summarizing how geographic areas vary from each other. For example, states differ with respect to which political party is in control of the governor's office and a choropleth map can quickly display which states are led by Democrats or Republicans by simply coloring them blue or red (respectively). These maps can also be useful for describing how state policies spread across the country. Because states are "laboratories of democracy," one state might adopt a particular policy and eventually others will follow if the policy is perceived to be politically viable (Shipan & Volden, 2008). In higher education research, these maps are used to document how the finance policies such as college savings accounts, merit-based financial aid programs, and performance-based funding diffuse across state lines over time (Cohen-Vogel, Ingle, Levine, & Spence, 2007; Doyle, McLendon, & Hearn, 2010; Hillman, Tandberg, & Gross, 2014; Lacy & Tandberg, 2014).

Researchers can also use these maps to indicate how counties or other geographic regions vary on a particular measure. For example, to document the degree of racial segregation that occurs within the state of Texas, Cortes (2010) uses county-level choropleth maps to show how the state's Hispanic population is concentrated in southwestern counties, whereas the Black population is concentrated in the east. Choropleth maps are also used to document the number of colleges in a particular commuting zone (Hillman, 2016) or the number of colleges adopting the Common Application (Smith, 2014).

In each of these cases, readers are able to quickly see how geography can shape educational opportunity and where prospective students have more college choices nearby. Cortes' (2010) maps show that most of the state's universities are concentrated in the central and southern areas of the state – far away from where the state's minority populations live – which suggests banning affirmative action would only reinforce structural inequalities that already exist in the state. Hillman's (2016) maps show clusters of counties that have few or no colleges nearby, whereas Smith's (2014) maps show that colleges adopting the Common Application are clustered around the East Coast. In each example, these maps effectively display that the market structure for higher education differs across some geographic space (the country) and that not all communities have equal opportunities nearby.

Figure 11.2 provides an example of a choropleth map showing educational attainment levels of each Michiana county. Lighter shades represent higher educational attainment, darker shades represent lower educational attainment; readers can quickly see that St. Joseph and Berrien Counties have the highest attainment levels. Not only does Elkhart County have a small number of colleges (Fig. 11.1), but they also have a high share of residents with no formal education beyond high school.

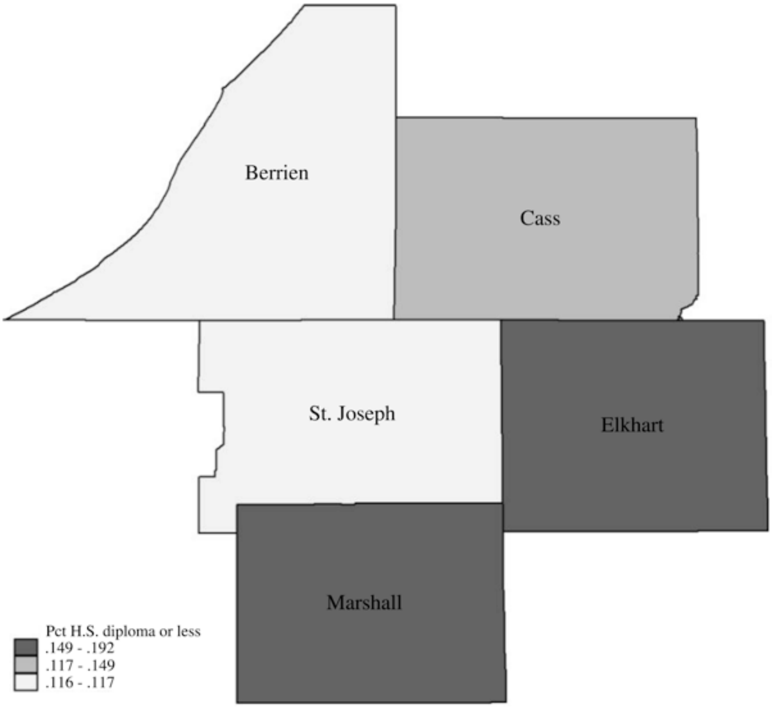

Fig. 11.2 Share of Michiana county-level adult population with a high school diploma or less

There are many plausible explanations about why these counties differ in terms of educational attainment, so choropleth maps can be useful tools for investigating these explanations and documenting the extent to which areas vary from one another.

Dot Maps

When researchers want to plot a specific place on a map, like the location of a college in Fig. 11.1, they may prefer to use dot maps. This visualization can help readers quickly identify how close or far a college is from another college, or how far it is from different communities within a particular geographic region. For example, De Oliver (1998) uses dot maps to illustrate how the location of the new University of Texas-San Antonio campus reinforces existing neighborhood segregation. When the state established this new university in 1969, it was unclear whether the campus would be built in the city center or in the suburban fringe. But when state officials decided to build the new campus in the northwest fringe of the city, rather than the urban core, working class and Latino residents of the city's south side contested the decision on grounds that it imposed additional costs on urban residents who would likely be unable to get back and forth to the new campus.

Jepsen and Montgomery (2009) use dot maps to show where community colleges are located around Baltimore, Maryland, finding that adult students are discouraged to enroll in college due to distance and transportation barriers. Their analysis found that just 2 miles of additional travel time between home and college can discourage students from enrolling. By showing the location of colleges relative to major highways and population centers, the study helps readers interpret the results and understand why working adults might be discouraged from enrolling even when there are several community colleges located in the greater Baltimore area. Similarly, Liu (2015) plots the location of the different types of Land Grant Colleges and Universities across the United States to estimate the spillover effects they have for their regional economy. Brewer et al. (2016) use dot maps to locate Tribal Colleges and Universities and to explain how federal land policies both compelled and restrained tribal governments to engage in agricultural development. Dot maps can be effective tools in displaying where colleges are located relative to other important sociodemographic indicators such as population centers, public transportation, and even other colleges and universities.

González Canché (2014) uses dot maps to show where colleges are located across the country, which in turn allowed him to examine how neighboring colleges might affect the tuition charged in a local area. The study found non-resident tuition prices are converging along geographic lines, where colleges set non-resident tuition prices according to the prices neighboring institutions charge. Using a geospatial lens, González Canché is able to reveal that colleges both compete against and are influenced by their neighbors. College pricing models are shaped by their local marketplace, where the amount charged to non-residents is conditioned by what nearby colleges charge. More specifically, being located near broad-access colleges tends to allow colleges to charge higher non-resident tuition whereas being neighbors with selective research universities may reduce the amount of non-resident tuition a college can charge. Thinking about pricing policy as a function of local market conditions, rather than by political or governance decisions, allows researchers to explore how spatial trends bear on a wide range of educational outcomes.

Heat Maps

Sometimes neither a choropleth nor dot map will visualize what the researcher needs to communicate, so heat maps can be useful alternatives. This type of map is similar to a dot map, but instead of simply displaying a point on a map it shows how densely a particular variable clusters around each point. This density is often color coded similar to choropleth maps, where the highest-density areas are shaded differently than lower-density areas. This strategy helps researchers identify geographically clustered areas and the extent to which these clusters are statistically different from others. If one geographic area has a greater density or frequency of "hot spots" than another, then these patterns may not be due to random noise in the data. These patterns can be either first-order (environmental) processes where one variable causes the clustering to occur, or they can be second-order (interaction) processes where the presence of a point increases the likelihood of observing another point (De Smith, Goodchild, & Longley, 2015).

Hites, et al. (2013) provides one of the only examples of heat maps in higher education literature. The analysis shows "hot spots" of crime on a college campus, coded by the street address where the crime occurred and the time of day and type of incident (e.g., violent crime, property crime, or other crime). Using geospatial analysis, the research team was able to discover patterns that not only helped document where events occurred, but information that was also useful for campus administrators interested in preventing, prioritizing, and responding to crime activities on campus. This study illustrates a first-order geographic process and demonstrates how geospatial analysis can be useful in both diagnosing and responding to problems related to campus safety. Maps can be powerful communication tools to help frame or contextualize a problem and, in turn, can be useful for mobilizing policy responses and interventions.

Quasi-Experimental Design

Many quantitative studies use survey or administrative data that are generated through observational (rather than experimental) processes. As a result of using observational data, researchers are unable to randomly assign individuals into "treatment" or "control" groups. This often results in studies that suffer from self-selection bias, where it is difficult to disentangle correlation from causation. When researchers want to draw causal inference but cannot randomly assign subjects into treatment and control groups, they must rely on alternative quasi-experimental strategies (Angrist & Pischke, 2015; DesJardins & Flaster, 2013).

Distance and geographic boundaries can be particularly useful variables in quasi-experimental designs, namely in the application of instrumental variables, difference-in-differences, and regression discontinuity methods. The distance between home and college is often used as an instrumental variable, geographic differences in policies can be used in difference-in-difference analysis, and people who live on one side of a political or geographic boundary may be able to be used as treatments and counterfactuals in regression discontinuity studies. Each of these three examples are discussed in more detail below, with the central point being that geography extends far beyond maps and can be a useful tool for gaining insights into causal mechanisms driving educational outcomes.

Instrumental Variables

When researchers use observational data to answer causal questions, their models may suffer from endogeneity problems where the key variable of interest is correlated with the model's error term. This can result in omitted variable bias, where we cannot determine whether the outcome is caused by the key variable or whether something missing from the model is driving the results. Instrumental variables can be a useful tool for addressing this problem, but only when researchers can find instruments that minimally meet the following two conditions. The instrument must be highly correlated with the problematic endogenous variable. Next, it should only correlate with the outcome through this endogenous variable, after holding all else constant in the (first stage) of the model. When these conditions are met, the instrument can address the endogeneity problem thereby allow researchers to make more rigorous statements about cause and effect [instrumental variable techniques are discussed in more detail in Bielby, House, Flaster, and DesJardins' (2013) Handbook chapter].

An example will illustrate how distance has been used as an instrument in higher education research. Researchers and policymakers often want to know if students who transfer from community college to four-year institutions are as likely to graduate as those who start at four-year colleges (Reynolds & DesJardins, 2009). If they are, then the transfer function of community colleges could provide students with a viable pathway to the bachelor's degree. Because students self-select into college (i.e., are not randomly assigned), instruments can be used to address this endogeneity problem. Long and Kurlaender (2009) use distance as an instrument because a student‘s decision to start at a community college is highly correlated with this variable – the closer one lives to a community college, the more likely they are to attend such an institution. This correlation should have no bearing on the student‘s eventual bachelor's degree attainment, except through their decision to start at a community college.

By using distance as an instrument, Long and Kurlaender (2009) found that starting at a community college causes students to take longer to earn their bachelor's degree. This "community college penalty" is likely driven by the transfer function of colleges more than students' self-selection into college, suggesting policy and practices related to transfer articulation may actually slow down or even discourage bachelor's degree completion. But distance can be used in many other contexts where self-selection can bias the study's results. For example, the economic returns to a college education may be both a function of the student‘s self-selection into a particular college (i.e., motivation, ability, preferences, etc.) and/or due to the college itself (Carneiro, Heckman, & Vytlacil, 2011; Doyle & Skinner, 2016; Long, 2008). A similar challenge presents itself when measuring the civic returns to college, where the outcome could be driven by either self-selection or by the way colleges prepare students for civic life (Dee, 2004). Researchers interested in using distance as an instrument should become familiar with the longstanding debate about the strengths and weakness of using geography, including Card's (2001) review and analysis.

Difference-in-Differences

Often times in higher education research it is either impossible or impractical to randomly assign subjects into treatment and control groups. For example, it is impossible to randomly assign states to adopt certain policies and it is impractical for states to randomly choose counties when deciding where to build new colleges or universities. Because of this, researchers often look for “natural experiments” where a state or locality adopts a new policy that serves as a plausible source of exogenous variation. This variation is useful for causal inference because it allows researchers to classify states and localities that adopted the policy into a treatment group, whereas those not adopting the policy are included in the control group. By following the trends in both groups before and after the policy change, researchers are able to detect whether there are differences in outcomes for these two groups after the policy change. Such natural experiments can replicate random assignment under certain circumstances, which will be briefly discussed below. Those interested in a more complete discussion of this research technique should see Angrist and Pischke (2015) or St. Clair and Cook (2015) for helpful introductions.

Studies have exploited differences in geography and policy adoption in a number of ways. One of the most common applications is in studies estimating the enrollment effects of state merit-based financial aid (Dynarski, 2002, 2008, Sjoquist & Winters, 2012, 2015; Zhang, 2011; Zhang & Ness, 2010). Merit-based aid programs became popular in the early 1990s and expanded rapidly throughout that decade and into the 2000s (Doyle, 2006). In these studies, states that adopted merit-based aid policies are the “treatment” group and those not adopting are the “control.” By following both groups over time, researchers are able to detect whether outcomes for the treated group change after policy implementation. This research design rests on the assumption that the control group would have followed the same patterns as the treatment group after the policy change occurs. Because the control group never adopted the policy, it serves as the counterfactual for the treatment group – it represents what would have likely happened, after controlling for observables, had a state never adopted the policy [see Flaster & DesJardins (2014) for an accessible introduction to the concept of counterfactuals].

By comparing treated states/localities to non-treated ones, researchers can exploit geographic differences as an identification strategy. When states adopt new policies, like performance-based funding (Hillman et al., 2014; Hillman, Tandberg, & Fryar, 2015) or affirmative action bans (Backes, 2012; Garces, 2013; Hinrichs, 2012), this creates natural experiments based on geographic boundaries where difference-in-differences can be employed. But states do not have to serve as the only unit of analysis – policy changes often occur at the local level where counties within states experience significant policy changes that can be used to draw causal inference.

For example, Denning (2016) exploits variation in community college taxing districts, where 22 Texas municipalities joined community college taxing districts since 1995. Individuals who live outside a district are charged tuition that on average is 63 % higher than in-district rates; the annexations brought new municipalities into the district, resulting in an increase in the number of individuals now eligible for in-district tuition. The study found that reducing tuition in annexed districts expanded college access: a $1000 decrease in tuition induced a 5.1%age point increase in enrollment. Similarly, Carruthers and Fox (2016) examine the effect of the Knox Achievers programs, which provides targeted advising and free in-state community college tuition and fees to all Knox County high school graduates starting in 2009. They compare educational outcomes (e.g., high school graduation, college enrollment, and college credits completed) of Knox County high school graduates against similar high school graduates from nearby counties. The study finds increased college enrollment rates for Knox County students (and for lower-income students in particular), though it may have shifted high-achieving and higher-income Knox County students away from four-year colleges and into community colleges. The Knox Achievers program is one of many "promise programs" existing throughout the country, so similar studies could be replicated in other geographic regions (LeGower & Walsh, 2014). These have already occurred in at least 25 other communities across the country, notably in Kalamazoo (Andrews, DesJardins, & Ranchhod, 2010; Bartik, Hershbein, & Lachowska, 2015) where the program has yielded positive effects and in Pittsburgh where the effects are more mixed (Bozick, Gonzalez, & Engberg, 2015). In each of these studies, geographic and political boundaries are use as ways to identify treatment and control groups.

Geographic Regression Discontinuity

Regression discontinuity (RD) is an increasingly popular quasi-experimental technique because it, like the others, can replicate random assignment with observational data (McCall & Bielby, 2012; Flaster & DesJardins, 2014). Under this design, researchers take advantage of pre-determined eligibility thresholds that assign people into either treatment or control groups (U.S. Department of Education, 2010). For example, a state may require students to earn at least a 3.0 high school grade point average or reach a certain standardized test score in order to be eligible for student financial aid. A student who meets these criteria will receive the aid and is assigned to the "treatment" groups; those who fall just short do not receive the aid and are in the "control" groups. Because the state (and not the student) determined the eligibility threshold to assign students, those just below and above the cut-point are as good as random right at the cut point. Students on either side of the threshold are likely to share similar observable (e.g., test scores, grades) and unobservable characteristics (e.g., motivation, ability, etc.) that in turn make it possible to replicate random assignment.

Instead of using test scores or grade point averages as cut-points, researchers can use geographic borders and political boundaries to assign observations into treatment and control groups. People who live on one side of a border are likely to be similar to those living just on the other side, but the treated group will be eligible for benefits (or interventions) that are unavailable to the control group. For example, someone living on the Indiana side of Elkhart County can pay in-state tuition to the state's public colleges whereas people just a few miles away on the Michigan side of the state line pay out-of-state rates. The state line serves as a forcing variable, allowing researchers to estimate the causal effect a particular intervention like how enrollment changes when tuition rises.

To my knowledge, only one published study employs geographic regression discontinuity in higher education. Heaton, Hunt, MacDonald, and Saunders (2015) use the jurisdictional boundaries of University of Chicago's campus police department to estimate whether campus police reduce crime and violence. Campus police have full police powers granted by state governments although they are employed by the campus they serve; therefore, campus police jurisdiction and patrol zones are limited to the boundaries of a particular campus. In practice, however, campus police often work closely with the city's police department to patrol zones next to campus (Peak, Barthe, & Garcia, 2008). This results in some city blocks being patrolled solely by campus police, whereas other blocks are patrolled jointly with the Chicago Police Department, (CPD) creating geographic variation drawn along jurisdictional boundaries. In this study, researchers compare crime rates just outside the campus police patrol zone (i.e., in CPD's jurisdiction) to those just within the campus patrol zone and found CPD blocks have significantly higher crime rates than campus. Using 8 years of crime data, the authors conclude campus police are more effective at law enforcement because they patrolled similar blocks and had lower crime rates (Heaton et al., 2015).

Distance Elasticity

A longstanding area of research in higher education focuses on how students respond to changes in tuition prices. In this literature, researchers consistently find that a $1000 increase in price is associated with about a 5 % decline in enrollment (Deming & Dynarski, 2010; Heller, 1997; Leslie & Brinkman, 1987). The negative relationship between enrollment and price is consistent with microeconomic theory, though we should not constrain our thinking to only monetary prices when applying this model. After all, demand curves are downward sloping for human capital investment due to the law of diminishing returns and substitution effects (Paulsen & Toutkoushian, 2008). Instead of thinking of consumption being determined by price, we can think of consumption also being determined by distance. That is, the further away one is from college, the less likely they are to enroll, similar to the way high prices discourage students from enrolling.

Whereas most students' enrollment demand falls as the distance to college rises, some students actually prefer to travel far distances to attend college. From this perspective, demand is actually "u-shaped," where demand first falls with distance but then begins to rise as the distance between home and college increases. This pattern is consistent with patterns found with Veblen goods, or luxury items whose consumption actually increases when the expense (in this case, distance) rises. Students who have strong preferences to move away from home or those who choose to enroll out-of-state are likely to be less elastic with respect to distance. Traditional-aged students from more privileged backgrounds are likely to travel the furthest to attend college, whereas students of color, working class students, and those who have familial obligations tent to stay closer to home (Niu, 2014; Desmond

& Turley, 2009). For example, Niu found that White students were consistently more likely than Black and Hispanic students to send SAT scores to more colleges (a proxy for applying to college) and to enroll out of state. Similar differences emerged with respect to income, where higher income students were more likely than lower-income students to send more scores and eventually enroll out of state. Even after controlling for socioeconomic and academic characteristics, Desmond and Turley (2009) find that Hispanic students in Texas are less likely to leave home to attend college due in large part to familial obligations. Further research is necessary to understand how family obligations, academic preparation, and socioeconomic status interact with the decision to stay close to home, but geography of opportunity likely bears on these choices (Ovink & Kalogrides, 2015).

Geostatistics

This section reviews some basic statistical approaches that can be applied to georeferenced data, including: gravity models, dissimilarity index, entropy index, and spatial error regression models. This is not an exhaustive list of geostatistics, yet it should contain promising avenues for higher education researchers interested in applying basic geospatial statistics in their research. Researchers may be interested in other important geostatistical techniques including point pattern analysis to detect whether points are randomly disbursed in a specific geographic area. This technique would allow researchers to identify the extent to which other (i.e., non-geographic) variables might predict the dispersion of various points. Researchers may also find more advanced topics like spatial interpolation and kriging useful if they are interested in using spatially-weighted data to predict unobserved outcomes. There are many statistical techniques researchers might find interesting and useful, but this chapter focuses on a few key introductory techniques. For more information on advanced geostatistical methods, see Baddeley, Rubak, and Turner (2015), Bivand, Pebesma, and Gómez-Rubio (2013) and Chun and Griffith (2013).

Gravity Model

Researchers are often interested in studying the migration patterns of students or knowing the extent to which a college attracts students from a particular catchment area. Some colleges will attract students from a larger catchment area whereas others are relatively small; similarly, some states are net importers of students whereas others are net exporters. This relationship depends on students' own elasticity of demand, but is also a function of the potential pool of students who are located nearby. Gravity models allow us to think about enrollment patterns in the aggregate (rather than at the individual level) where some colleges or states/regions exert a force that pulls students in. In most cases, this force will diminish with distance, where further away colleges or states/region exert weaker force than those that are nearby to students. This is analogous to Newton's law of gravity, where larger bodies exert more force on smaller ones and the distance between them mediates their influence.

When applying gravity models to higher education, we can examine the movement of people from two places and observe the bond (i.e., gravitational force) between these places. One of the earliest applications of the gravity model was Stewart's (1941) examination into the catchment area of Harvard, Yale, Princeton, and Vassar. Wanting to know whether these colleges truly drew from a national pool of applicants, Stewart identified the home state of each undergraduate and measured both the total state population and the state's distance from each of the four colleges. Stewart found that enrollments were proportional to the state's population (i.e., smaller states sent fewer students) but that there was distance decay, where states from further away had lower representation at each of the study institutions. This was one of the first studies to show that distance and enrollment are dependent interrelated.

The basic formula is below where the force of a particular college or geographic region (F) is a function of both the population of two areas and the distance between these places:

$$\mathrm{F}_{ij} = \mathrm{k}\frac{P_i P_j}{d_{ij}} \tag{11.1}$$

Here, P represents the total population within two geographic areas, $_i$ and $_j$. The denominator in Eq. 11.1, d_{ij}, is the linear distance between the two geographic areas; it is possible to change the functional form of the denominator if we believe distance decay is nonlinear. Similar to Newton's law of gravity, k is the unobserved gravitational constant. This equation helps us see that the number of students enrolled in a particular college is directly proportional to the college's own population area and the population area of another institution, but inversely proportional to the distance between them (Kariel, 1968).

The gravity model has been applied and extended in higher education over time (Schofer, 1975). McConnell (1965) identifies the population center of each Ohio county and measures its distance to Bowling Green State University in order to find how closely geography is correlated to enrollment decisions. McConnell found similar results as Stewart (1941) where larger counties were associated with larger enrollments and there was an inverse relationship between distance and enrollments for more distant counties. Gravitation models have been employed for several other colleges including: Western Washington State College (Kariel, 1968), West Virginia University (Ali, 2003), University System of Georgia (Alm & Winters, 2009), Widener University (Leppel, 1993), and Washington State public four-year colleges (Ullis & Knowles, 1975). Each study extends the basic gravity model and applies different measures of geographic distance, yet each study finds enrollment declines as distance rises, though the strength of this relationship varies from case to case.

These studies also show that simply comparing the raw number of students from a particular state or region does not tell the full picture of geographic diversity of student enrollments. The full picture that gravity models help explain is that every region has its own "potential" with respect to the number of prospective students who might be inclined to enroll in a given institution. Gravity models can reveal that a particular county may have several potential students nearby, but few enroll, meaning the college is not meeting its enrollment potential.

Dissimilarity and Entropy Indices

In addition to assessing the extent to which a college draws students from nearby areas, researchers may also be interested in knowing whether the demographic profile of a given college is representative of the college's local or regional demographics. A college that enrolls a very homogenous group of students may have a racial/ethnic mix of students that is dissimilar than its metropolitan or even state demographics. There are several ways to quantify and map these differences that draw largely from sociologists interest in residential segregation (Massey & Denton, 1988; Reardon & Firebaugh, 2002).

This line of inquiry is uncommon in higher education research, though it could become increasingly important considering the changing demographics of the United States where the majority of children born today are non-white (U.S. Census Bureau, 2012). It is important for colleges to reflect these changing demographic patterns in their geographic region: a college located in a community that is not diverse would likely struggle to diversify its enrollments; alternatively, those located in highly diverse areas may want to have a student body reflective of this diversity. Researchers have long examined the geography of school segregation in K-12 education (Card & Rothstein, 2007; Dobbie & Fryer, 2011; Theil & Finizza, 1971). However, this line of inquiry has not been extended into higher education to the same degree, but holds promise with respect to identifying the causes and consequences of segregation among colleges and universities.

One way to examine geographic segregation is by using an index of dissimilarity, which is a standardized measure of evenness ranging from 0 (perfect evenness) to 1.0 (maximum separation) that is commonly used in measuring residential segregation (Massey, Rothwell, & Domina, 2009). When geographic areas become increasingly segregated, the dissimilarity index can tell us what proportion of a particular group would need to change in order to achieve perfect evenness. Massey et al. (2009) give an example of black-white segregation, where a dissimilarity index of 0.60 indicates that 60 % of the geographic area's black population would have to exchange places with the area's white population in order to achieve no segregation. The following formula is used to create the index:

$$\mathrm{D} = \frac{1}{2}\sum_{i=1}^{n}\left[\left| \frac{w_i}{w_T} - \frac{b_i}{b_T} \right| \right] \tag{11.2}$$

Table 11.3 Index of dissimilarity illustration Dane County, Wisconsin

	White	Black	$\frac{w_i}{w_T}$	$\frac{b_i}{b_T}$	$\left\lvert \frac{w_i}{w_T} - \frac{b_i}{b_T} \right\rvert$	D
University of Wisconsin-Madison	23,301	662	0.28	0.12	0.16	
Dane County (18–24 year olds)	58,516	4856	0.72	0.88	0.16	
Total	81,817	5518			0.33	0.16

Source: Enrollment data is from the U.S. Department of Education's IPEDS 2014 fall enrollment survey and county population estimates are from U.S. Census Bureau Bridged-Race Population Estimates http://wonder.cdc.gov/bridged-race-v2014.html

where w and b in Eq. 11.2 represent the White and Black population, respectively, within a geographic subunit (i) and across the geographic area's total population (T). The absolute value of this difference is summed for the total number of subunits and then multiplied by half in order to standardize the score. Table 11.3 illustrates how this could be used by researchers, where it compares enrollment data for White and Black undergraduates against Dane County's White and Black population between the ages of 18 and 24. The purpose is to illustrate the concept, so this example should be viewed as a proof of concept where the sum of each group's absolute values equal 0.33. This is then multiplied by half, resulting in a dissimilarity index of 0.16, suggesting that 16 % of Dane County's Black young adult population would have to exchange places with the University of Wisconsin's White undergraduate students in order to have the university be reflective of the local demographic profile of Dane County.

A different measure of local population distributions is the entropy index, which measures the extent to which multiple racial/ethnic groups are evenly distributed within a particular geographic area (Iceland, 2004). This technique does not allow us to examine segregation in the way a dissimilarity index does, though it can be a useful unifying measure of diversity to assess how diverse a college is relative to the diversity of a given geographic area. The entropy index (E) for a geographic unit (i) is defined in Eq. 11.3:

$$\mathrm{E}_i = \sum_{k=1}^{k} P_i \log\left(\frac{1}{P_i}\right) \tag{11.3}$$

where P_i measures the proportion of a geographic unit's population that is of a particular racial/ethnic group (k).

Using the Dane County example, the maximum entropy score is log(k), so in this case the maximum is 0.69 because there are only two groups being compared (log(2) = 0.69). Higher entropy scores signify greater diversity whereas lower scores signify less diversity: the university's entropy score (0.13) is lower than the county's (0.27), indicating the university is less diverse than the county. It is possible to add more than two racial/ethnic groups, but for simplicity this table only compares Black and White populations. By adding more racial/ethnic groups to the example, the entropy index will change and will provide a unifying variable for total diversity.

The entropy index can also be weighted and standardized to fall between the values of 0 and 1, where 1 indicates totally equal representation between all geographic areas and 0 indicates totally unequal representation. A different index called Thiel's H is often implemented to compare multiple geographic areas, though researchers have developed several extensions to this and other indices [see Reardon and Firebaugh (2002) and Massey and Denton (1988) for examples].

In Franklin's (2013) study, each college was assigned an entropy index based on the equation above. The index was then used in a regression model as the dependent variable, where a series of controls were added to assess the extent to which institutional and regional variables were associated with having a higher or lower entropy index. Results showed that religiously-affiliated, highly selective, and smaller colleges tend to have lower entropy indexes (i.e., were less diverse). Alternatively, public colleges and those located in counties with more diversity among its young population tend to be more diverse. This study suggests institutional policies and practices such as admissions and enrollment management are stronger predictors of a college's diversity than are local demographics. Unless colleges can "import" diversity from other geographic areas, institutions that are located in racially/ethnically homogenous geographic areas will likely struggle to become more diverse.

Spatial Error Regression Models

Ordinary least squares regression assumes that the error terms in a given statistical model are independent and identically distributed ("iid") from one another. Consequently, residual values should be uncorrelated with one another and have constant variance (Baum, 2006). When this independence assumption holds, efficient standard errors result; but violating this assumption can result in artificially low standard errors that increase the chances of Type I errors. Researchers are often introduced to this autocorrelation problem when using time-series regression where a previous year's data is correlated with the following year's data or in multilevel modeling where nested groups are correlated with one another. But autocorrelation does not have to be temporal; rather, it can be spatial where clustering patterns occur relative to geographic areas and the proximity of one point to another (Ward & Gleditsch 2008).

The standard ordinary least squares equation below (Eq. 11.4) does not account for this spatial dependence in the error term:

$$y_{\mathrm{i}} = \alpha_0 + \beta_{\mathrm{i}} \mathrm{X}_{\mathrm{i}} + \varepsilon_{\mathrm{i}} \tag{11.4}$$

where y is the outcome for each observation (i), α is the intercept, and X is a variable with β measuring its slope coefficient; the error terms are assumed to be iid and normally distributed.

After running the regression, it is possible to estimate Moran's I statistic to examine whether the residuals in the OLS equation are spatially correlated. The statistic's null hypothesis is that residuals are randomly distributed where each observation is uncorrelated with its nearest neighbors. Moran's I uses a weighted index of each location relative to other locations to determine whether each observation is correlated to its neighbors (Ward & Gleditsch, 2008):

$$\mathrm{I} = \left(\frac{n}{\sum_i \sum_j w_{ij}} \right) \frac{\sum_i \sum_j w_{ij} \left(y_i - \bar{y} \right) \left(y_j - \bar{y} \right)}{\sum_i \left(y_i - \bar{y} \right)^2} \tag{11.5}$$

The formula in Eq. 11.5 is similar to the Pearson's correlation where values range between −1 and 1. Values closer to 1 are geographically located close to other high values and are thus positively correlated, whereas values closer to −1 are located closer to other low values and are negatively correlated. A value of zero indicates no spatial dependence, implying the data does not violate the iid assumption and there is no spatial pattern. Moran's I statistic does not have to be used for detecting spatial correlation in the error term, it can also be used to measure spatial correlation between observations which can be useful in spatially lagged models (Ward & Gleditsch, 2008).

If Moran's I is statistically significant, then researchers might consider implementing a spatial error regression model that extends the OLS model (Eq. 11.4) by decomposing the error term into two parts:

$$y_{\mathrm{i}} = \alpha_0 + \beta_{\mathrm{i}} \mathrm{X}_{\mathrm{i}} + \lambda \mathrm{w}_i \xi_i + \varepsilon_i \tag{11.6}$$

where ε satisfies the iid assumptions discussed above. Equation 11.6 adds λ to Eq. 11.4, which measures how closely each error (ξ) is correlated with other connected locations (*w*) for each observation (*i*). When error terms are spatially correlated or if there is reason for researchers to believe spatial dependence is problematic in the error term, then they can employ the spatial error model. Spatial dependence can also be handled via spatially lagged models discussed in more detail in Ward and Gleditsch (2008); for applications of this technique in higher education, see McMillen, Singell and Waddell (2007) and González Canché (2014).

Georeferenced Data Sources in Higher Education

This section provides an overview of geographically referenced (georeferenced) data that can be useful for conducting higher education research. When a dataset provides such information as latitude and longitude, county and state FIPS codes, or even the metro/micropolitan statistical area, zip code, or location of a high school, then it is possible for researchers to use these variables to merge with other useful datasets. For example, when Hillman and Orians (2013) examined the relationship between local labor market conditions and community college enrollment levels, they merged IPEDS data with U.S. Bureau of Labor Statistics (BLS) and Bureau of

Economic Analysis (BEA) data based on each college's metro/micropolitan statistical area. In so doing, they discovered that as local unemployment rates rose by 1 % point, enrollments rose between 1 and 3 %. Georeferenced data allows researchers to link education datasets with other datasets outside the field, such as political science, economics, sociology, public health, and others to gain a fuller view of the environment in which educational phenomenon occur.

Products like IPEDS and other survey or administrative datasets available from the National Center for Education Statistics (NCES) provide useful georeferenced variables. This section describes several of the more commonly used georeferenced data available from federal agencies and it directs readers to the websites and agencies where this data is made publically available. After providing an overview of the federal data landscape, the section discusses the hierarchical nature of georeferenced data and explains how national, state, and local data can be merged with common federal higher education datasets.

Hierarchy of Georeferenced Data

Figure 11.3 provides an overview of the hierarchy of geographic data, where national data is the highest level of aggregation for which statistics are produced and Census blocks are the lowest. Every 10 years, the U.S. Census Bureau conducts the national census to collect the official population counts for the entire country. In off census years, Census operates the Annual Community Survey (ACS) to provide updated information on a range of social and economic indicators. While the decennial census provides more detailed data than the ACS, both are useful and can provide at a minimum county-level population estimates. The decennial census provides population estimates as far down as each neighborhood block (discussed below). Figure 11.3 and the following discussion focuses on areal measures of geography, or "regional data that is attributed to some geographic area," as opposed to point measures that are "geolocated individual locations for observations" (Ward & Gleditsch 2008). Areal measures include places such as states, counties, and census blocks, whereas point measures include measures such as latitude and longitude.

The areal measures outlined in Fig. 11.3 can be organized into either legal/administrative entities or statistical entities (U.S. Census Bureau, 1994). States and counties are legal/administrative entities that have their own governing authority, taxing and voting districts, and a range of other governmental functions (e.g., post offices, schools, fire departments, etc.) outlined in laws or administrative codes. Cities, tribal lands, and other incorporated communities are examples of legal/administrative entities that are either embedded within or independent from those listed herein. Statistical entities are alternative ways to measure "local" areas, but statistical areas do not have legal or administrative authority as do states, counties, and cities. For example, micropolitan and metropolitan statistical areas, Census regions/divisions, Census tracts/blocks are all statistical entities and offer alternative ways to cluster geographic areas together.

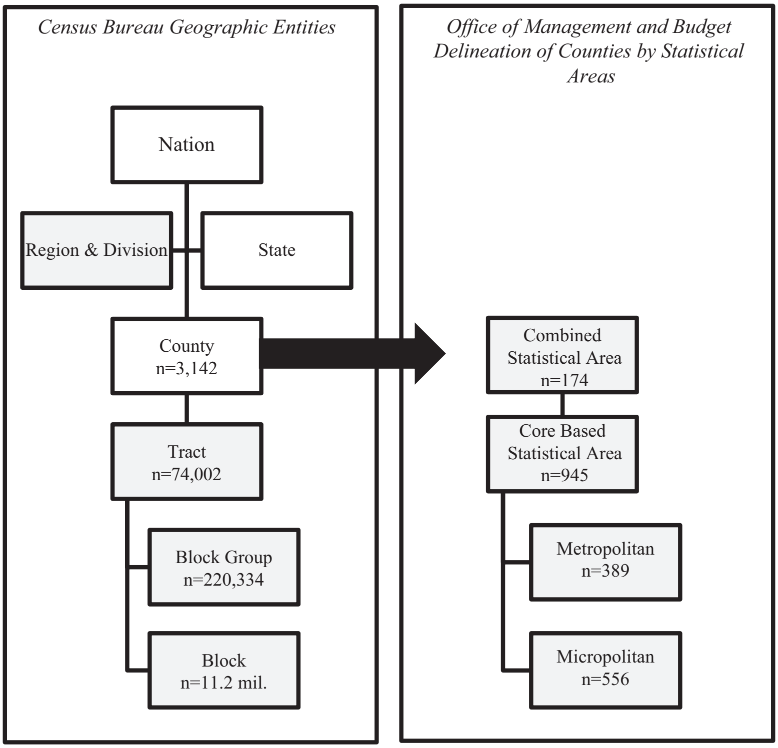

Fig. 11.3 Hierarchy of georeferenced entities

States and Regions

The Census Bureau divides all 50 states into four mutually-exclusive regions: Northeast, Midwest, South, and West. Within each region, states are further divided into at least two smaller "divisions" with the exception of the South region, which has three divisions. Census regions and divisions can provide guidance for researchers interested in state-level studies, where a particular state can be compared to others within the same region or division because they are likely to share similar geographic or demographic features as one another.

An alternative way to categorize states is by linking each to their interstate higher education compact (see Table 11.4). These compacts are clusters of states that collaborate with neighboring states on a variety of initiatives including tuition reciprocity agreements, leadership and policy development, and collective purchasing to gain economies of scale. The four compacts engage in intergovernmental relationships that derive their authority from state statutes, so these compacts are more similar to legal/administrative entities than statistical entities and offer a useful way to classify states in meaningful ways.

Table 11.4 Interstate regional compacts in higher education

Midwestern Higher Education Compact	New England Board of Higher Education	Southern Regional Education Board	Western Interstate Commission for Higher Education
Illinois	Connecticut	Alabama	Alaska
Indiana	Maine	Arkansas	Arizona
Iowa	Massachusetts	Delaware	California
Kansas	New Hampshire	Florida	Colorado
Michigan	Rhode Island	Georgia	Hawaii
Minnesota	Vermont	Kentucky	Idaho
Missouri		Louisiana	Montana
Nebraska		Maryland	Nevada
North Dakota		Mississippi	New Mexico
Ohio		North Carolina	North Dakota
South Dakota		Oklahoma	Oregon
Wisconsin		South Carolina	South Dakota
		Tennessee	Utah
		Texas	Washington
		Virginia	Wyoming
		West Virginia	

Note: n = 49 because Pennsylvania is not included in a regional compact

Counties

There are 3412 counties in the United States, and like states they have their own legal and administrative authority. Counties have their own governing bodies and taxing authority and they provide a wide array of services to local residents that vary both between and within states. They are often subdivided into smaller incorporated areas, consolidated cities, school districts, and other special purpose districts that operate their own services and have legal and administrative authority within and sometimes across a given county (U.S. Census Bureau, 1994). In higher education, 25 states support their community colleges through local property taxes that are assessed at the county or district level (Martorell, McCall, & McFarlin, 2014; Mullin, Baime, & Honeyman, 2015). Counties may also be used to determine a community colleges' service region, in-district and out-of-district tuition rates, or eligibility for educational benefits via local "promise programs" (LeGower & Walsh, 2014).

Blocks and Tracts

When the Census Bureau samples individuals, the smallest geographic entity assigned to an individual is their block. Blocks are defined according to visible features such as roads, railroad tracks, or rivers and by invisible features such as townships, school districts, or county lines. Due to being the smallest geographic entity, there are 11.2 million blocks across the country. These blocks are aggregated into larger block groups (n = 220,334) and then block groups are aggregated into census tracts (n = 74,002). Blocks and tracts offer more "localized" measures of geography than counties or states, though sometimes they may offer too small of an area depending on the research questions at hand. Census tracts optimally have populations of 4000 people, though their sizes range between 1200 and 8000 (U.S. Census Bureau, 2016b).

Combined Statistical Areas

Because people live and work in places that often cross over county or state lines, it can be helpful to cluster counties into larger areas that share common economic activity. The U.S. Office of Management and Budget (2015) creates these county clusters based on commuting patterns between people's homes and employment addresses. These clusters are organized into micropolitan (n = 556) and metropolitan (n = 389) statistical areas that, together, account for 1825 of the nation's counties and 94 % of the total U.S. population (U.S. Census Bureau, 2016c). A metropolitan statistical area must have "at least one urbanized area of 50,000 or more population, plus adjacent territory that has a high degree of social and economic integration with the core as measured by commuting ties" (U.S. Office of Management and Budget, 2015, p. 2). Micropolitan areas follow the same definition but the core area's population must be fewer than 50,000 and greater than 10,000. In both cases, these statistical areas incorporate whole counties and a county cannot belong to both a metro and a micro area; they are mutually exclusive.

Micropolitan and metropolitan areas are collectively referred to as Core Based Statistical Areas (CBSAs), and can even be clustered into larger geographic entities called Combined Statistical Areas (CSAs). There are 945 total CBSAs, but many of these micro/metropolitan areas share common borders and cluster with each other, resulting in 174 CSAs nationwide. A CSA must include multiple CBSAs, resulting in larger geographic areas that include multiple metropolitan and micropolitan statistical areas. Researchers in higher education may prefer to use CBSAs over counties because they account for local commuting patterns; CBSAs may also be preferable over CSAs because they offer a more localized measure of labor markets.

Commuting Zones and PUMAs

Not included in Fig. 11.3, but still relevant for measuring statistical areas, is the U.S. Department of Agriculture's Commuting Zone (CZ) measure. By definition, CBSAs exclude rural counties and the CZ includes them in their definition as an alternative to micro/metropolitan areas. First created in 1980, the U.S. Department of Agriculture created CZs as a way to measure rural economic activity by using Census journey-to-work data to measure county-to-county flows of commuters

(Tolbert & Sizer, 1996). Similar to CBSAs, the rationale behind creating CZs is to account for commuting patterns across state and county lines in order to measure local economic activity. Unlike CBSAs, CZs are not defined by population size but by the extent to which counties cluster together according to their amount of shared economic activity. Higher education researchers have used CZs to measure local higher education and labor markets (Hillman, 2016; Kienzl, Alfonso, & Melguizo, 2007), and they are increasingly used in economic research on intergenerational mobility (Autor & Dorn, 2013; Chetty, Hendren, Kline, & Saez, 2014).

Similar to commuting zones, the Census Bureau provides Public Use Micro Area (PUMA) data that covers the entirety of the United States (U.S. Census Bureau, 2016d). However, PUMAs do not cross over state lines, they must contain at least 100,000 people, and they are delineated by census tract as well as county. These criteria make PUMAs distinct from both commuting zones and combined statistical areas because they can include tract-level data. This delineation allows researchers to construct alternative measures of local geography that map closely with metropolitan statistical areas. The U.S. Census Bureau administers its American Community Survey (ACS) to households within each PUMA, so researchers may be able to link PUMA-level ACS data with other georeferenced census tract or county level data.

Federal Data Sources

U.S. Department of Education

Perhaps the most commonly used federal higher education datasets include NCES' IPEDS data and the suite of nationally-representative student surveys including National Postsecondary Student Aid Study (NPSAS), Beginning Postsecondary Students (BPS), and Baccalaureate and Beyond (B&B). The primary difference between these data sources is the unit of analysis. IPEDS provides aggregate campus-level data on such factors as admissions, enrollment profile, finances, financial aid, and completions. IPEDS provides useful information on the performance of colleges and universities over time; for more information on which colleges report to IPEDS and how this data is used in other products (e.g., Delta Cost Project) see Jaquette and Parra (2014). Unlike IPEDS, the national surveys use students as the unit of analysis where cross-sections of undergraduate and graduate students are sampled every 4 years in the NPSAS survey. The BPS survey follows NPSAS freshmen for 6 years after their initial enrollment, whereas B&B follows NPSAS graduating seniors for 10 years after earning their baccalaureate degree. Other NCES student surveys include High School and Beyond (HSB) and its predecessor longitudinal surveys, National Educational Longitudinal Survey (NELS) and Education Longitudinal Survey (ELS). The Common Core of Data (CCD) provides school-level data on the enrollment and finances of public K-12 schools. Across each data source, it is possible to use georeferenced data such as ZIP codes, state or county codes, and latitude/longitude to link students, their high school, or individual colleges together. Doing so can capture a more complete measure of local economic, social, and demographic environments.

U.S. Census Bureau

The Census Bureau provides a full list of all 50 states and 3142 counties according to their unique Federal Information Processing Standard (FIPS) codes. These codes are fundamental for conducting geospatial analysis and for merging georeferenced data at either the state or county level. States have two-digit FIPS codes and counties have three-digit codes, so the combination of state-county FIPS codes provides unique identifiers for each county in the country. For example, Indiana's state FIPS code is 18 and Elkhart County's county FIPS is 039, so the state-county FIPS for Elkhart County, Indiana, is 18,039. IPEDS provides these five-digit state-county FIPS codes that can then be merged with additional county-level data. However, only 1315 counties have degree-granting colleges receiving federal student aid, meaning IPEDS accounts for only about 40 % of the total number of counties in the U.S., so it is important to have the full universe of counties available for merging with other federal datasets so as to avoid sampling bias.

The Census Bureau also provides free and publically accessible base maps (called "shapefiles") that allow researchers to overlay georeferenced data onto customizable maps. These maps include counties, like Figs. 11.1 and 11.2, though they also include a number of other legal/administrative and statistical entities including: Congressional and state legislative districts, metro/micropolitan statistical areas, American Indian Areas, as well as ZIP codes and even Census tracts and blocks. Appendix A provides Stata commands and a brief narrative for using shapefiles to produce maps used in this chapter. Researchers using other statistical packages can also use these shapefiles and some are built into software programs so users may not need to download these files. In addition to providing identification codes and maps, the Census includes local population estimates, Small Area Income and Poverty Estimates (SAIPE), and county-level data on a range of topics that can be found in the American Fact Finder database.

Other Federal Data

The U.S. Department of Education and Census Bureau provide a wide array of georeferenced data, yet researchers may find the following sources useful for merging additional county-level data or latitude/longitude data with higher education data. The U.S. Department of Labor's Bureau of Labor Statistics (BLS), in collaboration with the Census Bureau, produces monthly and annual Local Area Unemployment Statistics (LAUS) that provide data on both the size of the county's civilian labor force and the number unemployed. The U.S. Department of Commerce's Bureau of Economic Analysis provides CBSA and county-level data on a range of economic indicators including: gross domestic product, personal income, earnings and employment by industry, and total population.

To access annual county-level population estimates according to race/ethnicity, age, and gender, the Centers for Disease Control and Prevention's (CDC) National Center for Health Statistics provides "Bridge-Race Population Estimates" tables. The U.S. Department of Agriculture's (USDA) Economic Research Service office provides county-level data on a number of environmental measuring that could serve as instrumental variables or controls in statistical models, including: oil and gas production, farmland, The USDA also uses county FIPS codes to classify rural-urban commuting areas and commuting zones that can provide additional geographic context to the surrounding community of colleges or places prospective college students live. Rather than accessing each individual administrative agency to download county-level data, the Robert Wood Johnson Foundation and University of Wisconsin Population Health Institute produce the County Health Rankings publicly accessible database that includes a wide range of county-level social, economic, and public health indicators. These draw from both federal and private data sources that link to county-level data on such topics as income inequality, housing and transit, poverty rates, and public health.

Latitude and Longitude

Figure 11.3 focused on areal measures of geography, though georeferenced data also comes in latitude and longitude. Every point location on a map has a corresponding latitude and longitude, so knowing these two points allows researchers to locate particular observations and link them to legal/administrative and statistical entities. It also allows researchers to measure the linear or networked distance between two points. Linear, or geodesic, data is commonly applied in research, where the distance between one point (e.g., student‘s home address) is measured "as the crow flies" to another point (e.g., student‘s college). Networked distance would account for the roads and traffic flows between the two points, where distance is routed through a series of networks. Measuring networked distance opens up new ways of thinking about proximity that incorporate time and traffic patterns and can give a more nuanced view of the proximity between two points. Regardless of whether distance is measured linearly or via networks, latitude and longitude are useful georeferenced data points to facilitate the integration of geography even further into higher education research. If a researcher only has the latitude and longitude of a particular college, they can use Roth's crosswalk to link it to counties or ZIP codes (Roth, 2016a, 2016b). In Stata, the command "geodist" can also measure the linear and curved (because the earth is spherical) distance between points.

Summary

There are several data sources in higher education and across the federal government that can be merged together according to areal or point measures of geography. Knowing a college or student‘s zip code, county code, CBSA, or latitude and longitude enables researchers to link higher education databases to a wide array of other government and private data. By merging datasets on georeferenced data outlined in this section, it is possible to implement some of the analytical techniques described in the earlier part of the chapter. Researchers can use georeferenced data to produce data visualizations such as choropleth, dot, and heat maps to communicate research findings. Similarly, georeferenced data can be useful for conducting quasi-experimental research where the distance between points and areal locations can be used as either instruments, thresholds, or counterfactuals for improving causal estimation in the field of higher education. Distance measures can even be used in elasticity studies to assess how the friction of place might shape students' enrollment decisions. And of course, geostatistics require researchers to utilize georeferenced data sources that may only be available by merging datasets.

There are a number of software packages available for creating maps and conducting geospatial analysis. The most common is Arc/GIS, a platform that provides built-in data and analytical options specifically designed for geographic analysis. There are also free software programs publicly available such as R, Tableau Public, and QGIS that provide a robust range of options for geospatial analysis. These software packages, and R in particular, have the capability of not only producing maps and conducting geospatial analysis, they are also well-suited to creating interactive data tools.

Arc/GIS and R stand out as the leading software programs in terms of the range of geospatial functionality they provide. Researchers can build static and interactive maps, conduct geostatistical analysis, and manage georeferenced data systems in these software programs. Tableau Public and QGIS are more useful as data visualization than as tools to conduct data analysis tools, so researchers may find their functionality to be sufficient if the goal is to communicate findings through visual displays. Stata may be a middle ground between all of these, where it provides a robust set of geospatial commands via *spatreg*, *spmap*, and *spatwmat*; however, to date the software does not have a function for making maps interactive online like R or Arc/GIS does. These packages are useful tools but the list should not be treated as comprehensive in terms of this type of functionality. Arnold and Tilton (2015), Pisati (2004, 2008), Bivand, Pebesma, and Gómez-Rubio (2013), and De Smith, Goodchild, and Longley (2015) provide helpful introductions to some of these as well as other software programs. Regardless of which software program one uses, the information contained in this section should help researchers identify and conceptualize ways to integrate existing georeferenced data sources that hold promise in higher education research.

Extending the Use of Geography in Higher Education Research

Even during a period of great technological advances with distance learning and online delivery of education, "place" still matters. In fact, place may matter even more today for two reasons. First, early evidence suggests that enrolling exclusively online produces negative educational effects for students, yet face-to-face or blended learning hold more promise (Alpert et al., 2016; Bettinger et al., 2015; Bulman & Fairlie, 2016; Figlio et al., 2013). Even if high-quality education were available online, many rural and economically depressed communities do not have high speed internet or have limited access to computers, making it even more challenging to believe distance education is a solution to the geography problem (Pick, Sarkar, & Johnson, 2015; Strover, 2014). Second, as the cost of attending college continues to rise, students may find it more affordable to stay close to home as a cost-saving strategy. According to the Higher Education Research Institute's (2015) annual survey of college freshmen, nearly 20 % of the 2015 incoming class said attending college close to home is "very important." This figure is higher for students attending public colleges and universities and Historically Black College and Universities. While we do not know the extent to which students choose colleges close to home solely because of rising costs, it is telling that the share of incoming freshmen reporting living near home was "very important" has actually risen since the 1980s (Chronicle of Higher Education, 2016). For these reasons, Turley's (2009, p. 126) advice that we "should stop treating the college-choice process as though it were independent of location and start situating this process within the geographic context in which it occurs."

Geography is an enduring topic and one of growing interest among researchers and policymakers interested in affecting educational opportunity and outcomes. Outside of higher education, research on income inequality and social mobility demonstrates that geography has a strong effect on shaping children's life chances (Chetty et al., 2014). Research in public health, nutrition, social services, transportation, and a number of other public services has long examined how the physical and geographic features of communities (i.e., the "built environment") shapes behaviors and social interactions (Kennedy, 2004; Massey, 2005; Soja, 2010; Walker, Keane, & Burke, 2010). And within education, researchers have examined the role of neighborhoods, the location of schools, and how environmental contexts shape students' educational trajectories (Burke, Greene, & McKenna, 2016; Miller, 2012; Tate, 2008). This chapter highlights several examples where higher education research applies geospatial techniques, and it hopefully opens avenues for new areas of research to build upon, extend, and critique the way in which higher education scholarship engages with place. Given the advances in data and software systems that make mapping and geospatial analysis more accessible, researchers may find this chapter to be a helpful introduction for conducting geospatial analysis.

Further Research

Geography can afford researchers a framework from which to conceptualize and analyze various phenomena within the field of higher education. This chapter presented a number of studies that utilize geospatial analysis, showing how geospatial analysis can be useful in visualizing data, improving internal validity of statistical models, conceptualizing the role that place plays in college opportunity, and considering colleges as part of a community's built environment. While there are a number of other ways researchers can engage with geography, below are three general areas that are of great promise for advancing theory, generating new knowledge, and informing public policy.

Geography of Opportunity

Traditional college choice theories focus on temporal stages, a sequence of events that must be achieved prior to advancing to the next sequence of events in the college-going process. Students must first develop predispositions toward college, where they develop aspirations, take courses that prepare them for college, talk with counselors about what is needed for entry into college, and several other milestones. Next, they will search for potential colleges that offer an academic or social environment where they perceive they will fit in; after narrowing the list of schools, students will gather information about these schools from a variety of sources. Finally, after advancing through these processes, students will apply, be admitted to, consider the costs and benefits of attending, and then register and enroll in a postsecondary institution. This is an oversimplification of the traditional college choice process that is commonly used today when discussing student's "match" and the process of transitioning from high school to college (Bowen, Chingos, & McPherson, 2009; Cabrera & La Nasa, 2000; Hossler & Gallagher, 1987).

While instructive, particularly for high school students seeking to attend four-year colleges full-time, if we focus solely on the *process of opportunity* then we will believe the reason students do not attend college is presumably due to informational barriers or failures to make the "right steps" along the way. For a student who needs to stay close to home for familial, cultural, or economic reasons, they may choose a college close to home regardless of whether they have achieved all the milestones outlined under the traditional choice-making process. Because of this reason, it is important to also consider the *geography of opportunity* that prioritizes the importance of place while describing the variations that occur across the country in terms of the accessibility of local colleges. Doing so can extend theories of college choice in new and useful ways.

The number, location, and selectivity of colleges across the country are very uneven where areas with low educational attainment levels and those with large Hispanic populations tend to have the fewest colleges nearby (Hillman, 2016). The example of Elkhart County stands out as an "education desert" where residents of the county have few public options nearby; they must travel to their neighboring county or across state lines to attend a public institution. Although local community colleges offer some satellite campuses not reported in the federal IPEDS database, colleges are concentrated in St. Joseph County. The only public four-year institution in the metropolitan statistical area is Indiana University-South Bend, a moderately selective institution that admitted 73 % of applicants in the fall of 2014, suggesting it may not be broadly accessible to members of the metro area. More research on the ways local areas vary with respect to the structure of higher education marketplaces is a promising area for further research that can help build new theories of college choice that are more applicable to today's "post-traditional" college student. In so doing, researchers may also employ critical geography in ways that explain and offers solutions for why college opportunity structures differ so widely from place to place.

State Policy and Borders

State borders can provide researchers with sharp discontinuities for designing geographic regression discontinuities and even for determining counterfactuals in the difference-in-difference framework. In either case, further research should continue to examine areas located just on opposite sides of state or national borders to help make stronger inferences with respect to a number of important policy topics. For example, it could be possible to examine students living just on one side of the Michigan border to observably similar students on the Indiana side to see whether changes in tuition have differential effects on their enrollment decisions. These effects would likely differ in areas that have tuition reciprocity agreements where residents on either side of the state line can pay in-state rates. Similarly, if a college on one side of the state line was subject to a statewide policy such as performance based funding, perhaps colleges just on the other side of the line not subjected to the policy would serve as a useful counterfactual. These examples show how borders can be useful for examining state policy effects while contributing to the generation of new knowledge around student and organizational behavior.

Michiana also illustrates a cross-border area where the five counties shown in Fig. 11.1 cluster into a common statistical area that spans both Michigan and Indiana. There are many commuting zones and CBSAs that cut across state lines, as shown in Fig. 11.4. Further research could extend the research of Sponsler, Kienzl, and Wesaw (2010) that found state financial aid programs, credit transfer agreements, tuition policy, and student mobility were greatly affected in metropolitan areas that cut across state lines. Figure 11.4 shows counties (clustered by commuting zones) that cross over state lines. Approximately 48 million adults over the age of 25 live in these counties, and there are 429 public four-year and two-year colleges operating in these communities. These numbers are non-trivial as they account for approximately 25 % of the nation's adult population and the same share of the nation's public colleges and universities.

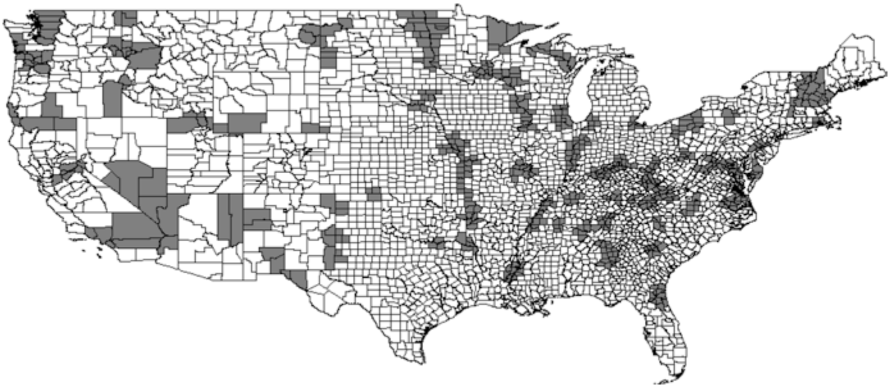

Fig. 11.4 Counties located in commuting zones that cross state borders

Local Higher Education Markets

Colleges and universities are located near each other and that proximity likely affects each institution's behavior. Neighboring colleges may be more inclined to have transfer articulation agreements, partnerships with similar local high schools or businesses, or even compete with each other over pricing and enrollments. Outside of a few examples [e.g., González Canché (2014); McMillen, Singell, & Waddell (2007); Rephann (2007)], these local market contexts are not commonly studied yet they hold promise for advancing new knowledge about the role geography can play in better understanding and improving these marketplaces. For example, among the 17 colleges operating in the Michiana market, what academic programs are available at the various institutions? How representative of the local demographic profile is each institution? Are they competing for the same pool of local high school graduates or adult students? There are many questions that could be answered by examining local higher education markets, and a geospatial lens can help advance understanding in that area.

In addition to understanding how colleges operate within local higher education marketplaces, further research could examine the extent to which students are mobile across various local areas. When examining student mobility, research often focuses on interstate movement and the predictors of attending college out-of-state (González Canché, 2014; Jaquette & Curs, 2015; Niu, 2014). Intrastate mobility is also an important research question for understanding student mobility and for policymakers interested in the extent to which students are attending their nearby institutions. It is plausible that college choice theory differs for intrastate versus interstate mobility, and there could even be important lessons to be learned about how the built environment (e.g., transportation, K-12 schools, etc.) and local socioeconomic indicators (e.g., poverty, tax base, etc.) explain variation in terms of where students attend college and what resources are available in their local public institutions. There are many avenues one could take to explore the local higher education marketplace and a geospatial lens will likely be a helpful way to frame further research in this area.

Conclusion

Geography affects many aspects of higher education, ranging from how students make choices to the site selection of new colleges. Place has been a longstanding topic of interest in higher education, yet is also a topic that has received less attention in the higher education research community than in other disciplines. This chapter presented several examples where geography and a geospatial lens has informed higher education research, and it hopefully offers points of departure for researchers to expand, extend, and critique how geography is currently used in the field. The vignette of Elkhart County and the surrounding area can be applied anywhere across the country, from Tribal lands to the Black Belt of Alabama, because the location of colleges matters for educational opportunity. Even in areas that have many colleges nearby, a student who needs to stay close to home may not have options due to the college's selectivity, program offerings, or affordability. Considering the number of students enrolled today who are working adults and the future of higher education that will be much more diverse, the need to understand the geography of opportunity is becoming even more pressing.

Geography can be destiny for students, where opportunities may be widely available in some communities and few (or nonexistent) in others. When this occurs, it only makes existing educational inequalities more difficult to reverse, so applying a geographic lens to higher education research can help us not only arrive at new understandings about how colleges and students respond to their local opportunity structures, but also to finding innovative ways to reverse inequalities. Designing more rigorous studies and advancing new theories that attend to the issues presented in this paper can help our scholarly community understand the causes and consequences of geographic inequality. But more importantly, applying a geospatial lens to higher education may help us find new solutions to longstanding problems that only now can become clear thanks to new ways of thinking, research strategies, and data sources available for conducting research in this area.

Appendix A

Stata commands to create maps found in this chapter.

```
// Download shape files to working directory from the following
   Census
   website https://www.census.gov/geo/maps-data/data/cbf/cbf_counties.html

// Install geospatial commands: ssc install spmap; ssc install shp2dta

// Set working directory
      cd "C:\Users\Maps"
```

```
// Convert shape file to Stata using shp2dta command
      shp2dta using cb_2015_us_county_500k, data(US_data)
      coor(US_coordinates) genid(id)
      use US_data.dta, clear
      destring GEOID STATEFP, replace

// Merge additional county-level data
      merge 1:1 GEOID using “C:\Users\Maps\ed_attain.dta”
      drop _merge
      replace hs_less = hs_less/100
      format %5.0g hs_less

      merge 1:1 GEOID using “C:\Users\Maps\county_center.dta”
      drop _merge

// Creating “Michiana” dummy variable to shade state map
      gen michiana = 0
      replace michiana=1 if inlist(GEOID,18039,18141,18099,26021,26027)
      label define michiana_label 0 “Not Michiana” 1 “Michiana”
      label values michiana michiana_label

// Base map of Michigan and Indiana with “Michiana” coded (Fig. 11.1)
      spmap michiana if (STATEFP==18|STATEFP==26) using US_coordi-
      nates.dta,
      id(id) clnumber(2) legenda(off) fcolor(white gray)

// County map of Michiana with point location of colleges (Fig. 11.1)
      spmap if michiana==1 using US_coordinates.dta, id(id)
      clnumber(3)
      legstyle(2) point(data(“IPEDS.dta”)select(keep if inlist(GE
      OID,18039,18141,18099,26021,26027)) x(lon) y(lat) by(sector2)
      fcolor(white gray black) ocolor(black ..) size(*0.8) legenda(on)
      legcount) label(x(intptlong)y(intptlat) label(NAME)select(keep
      if inlist(GEOID,18039,18141,18099,26021,26027)))

// County map of Michigan with educational attainment (Fig. 11.2)
      spmap hs_less if michiana==1 using US_coordinates.dta, id(id)
      clnumber(3) legstyle(2) label(x(intptlong)y(intptlat) label(NAME)
      select(keep if michiana==1)) legtitle(“Pct H.S. diploma or less”)

// County map of cross-border commuting zones (Fig. 11.4). Not
shown are steps
    to identify cross-border commuting zones and the variable
    “singlestate,” which are available upon request.
    spmap singlestate if (state~=“AK” & state~=“HI”) using US_coor-
    dinates.dta, fcolor(white gray) id(id) clnumber(3) legenda(off)
```

References

Ali, K. (2003). Analysis of enrollment: A spatial-interaction model. *The Journal of Economics, 29*(2), 67–86.

Alm, J., & Winters, J. V. (2009). Distance and intrastate college student migration. *Economics of Education Review, 28*(6), 728–738.

Alpert, W. T., Couch, K. A., & Harmon, O. R. (2016). A randomized assessment of online learning. *American Economic Review*, *106*(5), 378–382. http://doi.org/10.1257/aer.p20161057

Andrews, R. J., DesJardins, S., & Ranchhod, V. (2010). The effects of the Kalamazoo Promise on college choice. *Economics of Education Review*, *29*(5), 722–737. http://doi.org/10.1016/j.econedurev.2010.05.004

Angrist, J. D., & Pischke, J.-S. (2015). *Mastering metrics: The path from cause to effect*. Princeton, NJ: Princeton University Press.

Arnold, T., & Tilton, L. (2015). *Humanities data in R: Exploring networks, geospatial data, images, and text*. Dordrecht, The Netherlands: Springer.

Astin, A. W. (1980). *The impact of student financial aid programs on student choice*. Los Angeles, CA: University of California-Los Angeles. Retrieved from http://eric.ed.gov/?id=ED187268

Autor, D. H., & Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the US labor market. *American Economic Review*, *103*(5), 1553–1597. http://doi.org/10.1257/aer.103.5.1553

Avery, C., Howell, J., & Page, L. (2014). *A review of the role of college counseling, coaching, and mentoring on students' postsecondary outcomes*. Washington, DC: College Board.

Backes, B. (2012). Do affirmative action bans lower minority college enrollment and attainment? Evidence from statewide bans. *Journal of Human Resources, 47*(2), 435–455.

Baddeley, A., Rubak, E., & Turner, R. (2015). *Spatial point patterns: Methodology and applications with R*. London, UK: CRC Press.

Bartik, T. J., Hershbein, B. J., & Lachowska, M. (2015). *The effects of the Kalamazoo promise scholarship on college enrollment, persistence, and completion* (Working Paper No. 15–229). Kalamazoo, MI: Upjohn Institute.

Baum, C. F. (2006). *An introduction to modern econometrics using stata*. College Station, TX: Stata Press.

Bettinger, E., Fox, L., Loeb, S., & Taylor, E. (2015). *Changing distributions: How online college classes alter student and professor performance* (Working Paper No. 15–10). Stanford University Center for Education Policy Analysis: Stanford University. Retrieved from https://cepa.stanford.edu/sites/default/files/WP15-10.pdf

Bettinger, E. P., & Baker, R. B. (2014). The effects of student coaching: An evaluation of a randomized experiment in student advising. *Educational Evaluation and Policy Analysis*, *36*(1), 3–19. http://doi.org/10.3102/0162373713500523

Bielby, R. M., House, E., Flaster, A., & DesJardins, S. L. (2013). Instrumental variables: Conceptual issues and an application considering high school course taking. In M. B. Paulsen (Ed.), *Higher education: Handbook of theory and research* (Vol. 28, pp. 263–321). Dordrecht, The Netherlands: Springer.

Bivand, R. S., Pebesma, E., & Gómez-Rubio, V. (2013). *Applied spatial data analysis with R*. New York, NY: Springer.

Bowen, W. G., Chingos, M. M., & McPherson, M. S. (2009). *Crossing the finish line*. Princeton, NJ: Princeton University Press.

Bozick, R., Gonzalez, G., & Engberg, J. (2015). Using a merit-based scholarship program to increase rates of college enrollment in an urban school district: The case of the Pittsburgh Promise. *Journal of Student Financial Aid, 45*(2), 2.

Brewer, J. P., Hiller, J. G., Burke, S., & Teegerstrom, T. (2016). A primer: Extension, Indian land tenure, and rangeland limitations. *Rangelands*, *38*(1), 16–22. http://doi.org/10.1016/j.rala.2015.12.002

Briscoe, F. M., & De Oliver, M. (2006). Access to higher education: A conflict between landed interests and democratic ideals. *Education and Urban Society, 38*(2), 204–227.

Brown, D. K. (1995). *Degrees of control: A sociology of educational expansion and occupational credentialism*. New York, NY: Teachers College Press.

Bulman, G., & Fairlie, R. (2016). *Technology and education: Computers, software, and the internet (No. w22237)*. Cambridge, MA: National Bureau of Economic Research.

Bureau of Labor Statistics. (2016). *Unemployment rates by county*. Retrieved April 27, 2016, from http://data.bls.gov/map/MapToolServlet

Burke, K. J., Greene, S., & McKenna, M. K. (2016). A critical geographic approach to youth civic engagement: Reframing educational opportunity zones and the use of public spaces. *Urban Education*, *51*(2), 143–169. http://doi.org/10.1177/0042085914543670

Cabrera, A. F., & La Nasa, S. M. (2000). Understanding the college-choice process. *New Directions for Institutional Research, 2000*(107), 5–22.

Cantoni, D., & Yuchtman, N. (2014). Medieval universities, legal institutions, and the commercial revolution. *The Quarterly Journal of Economics*, *129*(2), 823–887. http://doi.org/10.1093/qje/qju007

Card, D. (1993). *Using geographic variation in college proximity to estimate the return to schooling* (No. 4483). Cambridge, MA: National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w4483

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica, 69*(5), 1127–1160.

Card, D., & Krueger, A. B. (2000). Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania: Reply. *The American Economic Review, 90*(5), 1397–1420.

Card, D., & Rothstein, J. (2007). Racial segregation and the black–white test score gap. *Journal of Public Economics, 91*(11), 2158–2184.

Carneiro, P., Heckman, J. J., & Vytlacil, E. J. (2011). Estimating marginal returns to education. *American Economic Review*, *101*(6), 2754–2781. http://doi.org/10.1257/aer.101.6.2754

Carruthers, C. K., & Fox, W. F. (2016). Aid for all: College coaching, financial aid, and post-secondary persistence in Tennessee. *Economics of Education Review*, *51*(1), 97–112. http://doi.org/10.1016/j.econedurev.2015.06.001

Castleman, B. L., Page, L. C., & Schooley, K. (2014). The forgotten summer: Does the offer of college counseling after high school mitigate summer melt among college-intending, low-income high school graduates? *Journal of Policy Analysis and Management*, *33*(2), 320–344. http://doi.org/10.1002/pam.21743

Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *The Quarterly Journal of Economics, 129*(4), 1553–1623.

Chronicle of Higher Education. (2016). *Backgrounds and beliefs of college freshmen*. Retrieved from http://chronicle.com/interactives/freshmen-survey

Chun, Y., & Griffith, D. A. (2013). *Spatial statistics and geostatistics: Theory and applications for geographic information science and technology*. Thousand Oaks, CA: Sage.

Cohen-Vogel, L., Ingle, W. K., Levine, A. A., & Spence, M. (2007). The "spread" of merit-based college aid: Politics, policy consortia, and interstate competition. *Educational Policy*, *22*(3), 339–362. http://doi.org/10.1177/0895904807307059

Cortes, K. E. (2010). Do bans on affirmative action hurt minority students? Evidence from the Texas Top 10 % Plan. *Economics of Education Review*, *29*(6), 1110–1124. http://doi.org/10.1016/j.econedurev.2010.06.004

De Long, J. B., & Shleifer, A. (1993). Princes and merchants: European city growth before the Industrial Revolution. *The Journal of Law & Economics, 36*(2), 671–702.

De Oliver, M. (1998). Geography, race, and class: A case study of the role of geography at an urban public university. *American Journal of Education, 106*, 273–301.

De Smith, M. J., Goodchild, M. F., & Longley, P. (2015). *Geospatial analysis* (5th ed.). Leicester, UK: Matador Press.

Dee, T. S. (2004). Are there civic returns to education? *Journal of Public Economics*, *88*(9–10), 1697–1720. http://doi.org/10.1016/j.jpubeco.2003.11.002

Deming, D., & Dynarski, S. (2010). College aid. In *Targeting investments in children: Fighting poverty when resources are limited* (pp. 283–302). University of Chicago Press. Retrieved from http://www.nber.org.ezproxy.library.wisc.edu/chapters/c11730.pdf

Denning, J. (2016). *College on the cheap: Consequences of community college tuition reductions* (Working Paper).

DesJardins, S. L., & Flaster, A. (2013). Non-experimental designs and causal analyses of college access, persistence, and completion. In L. Perna & A. Jones (Eds.), *The state of college access and completion: Improving college success for students from underrepresented groups*. New York, NY: Routledge Press.

Desmond, M., & Turley, R. N. L. (2009). The role of familism in explaining the Hispanic-White college application gap. *Social Problems, 56*(2), 311–334.

Dobbie, W., & Fryer, R. G. (2011). Are high-quality schools enough to increase achievement among the poor? Evidence from the Harlem Children's Zone. *American Economic Journal: Applied Economics, 3*, 158–187.

Doyle, W. R. (2006). Adoption of merit-based student grant programs: An event history analysis. *Educational Evaluation and Policy Analysis*, *28*(3), 259–285. http://doi.org/10.3102/01623737028003259

Doyle, W. R., & Gorbunov, A. V. (2011). The growth of community colleges in the American states: An application of count models to institutional growth. *Teachers College Record, 113*(8), 1794–1826.

Doyle, W. R., McLendon, M. K., & Hearn, J. C. (2010). The adoption of prepaid tuition and savings plans in the American states: An event history analysis. *Research in Higher Education, 51*(7), 659–686.

Doyle, W. R., & Skinner, B. T. (2016). Estimating the education-earnings equation using geographic variation. *Economics of Education Review*. http://doi.org/10.1016/j.econedurev.2016.03.010

Dynarski, S. (2002). The behavioral and distributional implications of aid for college. *American Economic Review, 92*(2), 279–285.

Dynarski, S. (2008). Building the stock of college-educated labor. *Journal of Human Resources, 43*(3), 576–610.

Eagan, K., Stolzenberg, E., Bates, A., Aragon, M., Suchard, M., & Rios-Aguilar, C. (2015). *The American freshman: National norms fall 2015*. Los Angeles, CA: Higher Education Research Institute at UCLA. Retrieved from http://www.heri.ucla.edu/monographs/TheAmericanFreshman2015.pdf

Elliott, P. (2000). *Spatial epidemiology: Methods and applications*. London, UK: Oxford University Press.

Figlio, D., Rush, M., & Yin, L. (2013). Is it live or is it internet? Experimental estimates of the effects of online instruction on student learning. *Journal of Labor Economics*, *31*(4), 763–784. http://doi.org/10.1086/669930

Flaster, A., & DesJardins, S. L. (2014). Applying regression discontinuity design in institutional Research. In N. Bowman & S. Herzog (Eds.), *Methodological advances in studying college impacts* (New directions for institutional research). San Francisco, CA: Jossey-Bass.

Franklin, R. S. (2013). The roles of population, place, and institution in student diversity in American Higher Education. *Growth and Change, 44*(1), 30–53.

Fyfe, N. R., & Kenny, J. T. (2005). *The urban geography reader*. Bristol, UK: Psychology Press.

Garces, L. M. (2013). Understanding the impact of affirmative action bans in different graduate fields of study. *American Educational Research Journal, 50*(2), 251–284.

Gasman, M., & Hilton, A. (2012). Mixed motivations, mixed results: A history of law, legislation, historically Black colleges and universities, and interest convergence. *Teachers College Record, 114*(7), 1–20.

Goldin, C., & Katz, L. F. (1998). The origins of state-level differences in the public provision of higher education: 1890–1940. *The American Economic Review, 88*(2), 303–308.

González Canché, M. S. (2014). Localized competition in the non-resident student market. *Economics of Education Review*, *43*, 21–35. http://doi.org/10.1016/j.econedurev.2014.09.001

Green, P. (2010). Racial politics, litigation and Mississippi's public Historically Black Colleges and Universities. *Race, Gender & Class, 17*, 241–269.

Griffith, A. L., & Rothstein, D. S. (2009). Can't get there from here: The decision to apply to a selective college. *Economics of Education Review, 28*(5), 620–628.

Heaton, P., Hunt, P., MacDonald, J., & Saunders, J. (2015). *The short- and long-run effects of private law enforcement: Evidence from university police* (SSRN Scholarly Paper No. ID 2564933). Rochester, NY: Social Science Research Network. Retrieved from http://papers.ssrn.com.ezproxy.library.wisc.edu/abstract=2564933

Heller, D. E. (1997). Student price response in higher education: An update to Leslie and Brinkman. *Journal of Higher Education, 68*, 624–659.

Hillman, N. (2016). Geography of college opportunity: The case of education deserts. *American Educational Research Journal*.

Hillman, N., Tandberg, D., & Gross, J. (2014). Performance funding in higher education: Do financial incentives impact college completions? *Journal of Higher Education, 85*(6), 826–857.

Hillman, N. W., & Orians, E. L. (2013). Community colleges and labor market conditions: How does enrollment demand change relative to local unemployment rates? *Research in Higher Education, 54*(7), 765–780.

Hillman, N. W., Tandberg, D. A., & Fryar, A. H. (2015). Evaluating the impacts of "new" performance funding in higher education. *Educational Evaluation and Policy Analysis*, 37(4), 501–519. http://doi.org/10.3102/0162373714560224

Hinrichs, P. (2012). The effects of affirmative action bans on college enrollment, educational attainment, and the demographic composition of universities. *The Review of Economics and Statistics, 94*(3), 712–722.

Hites, L. S., Fifolt, M., Beck, H., Su, W., Kerbawy, S., Wakelee, J., & Nassel, A. (2013). A geospatial mixed methods approach to assessing campus safety. *Evaluation Review*, 0193841X13509815.

Hossler, D., & Gallagher, K. S. (1987). Studying student college choice: A three-phase model and the implications for policymakers. *College and University, 62*(3), 207–221.

Howell, J. S., & Pender, M. (2016). The costs and benefits of enrolling in an academically matched college. *Economics of Education Review*, *51*, 152–168. http://doi.org/10.1016/j.econedurev.2015.06.008

Hoxby, C., & Avery, C. (2013). *The missing "One-Offs": The hidden supply of high-achieving, low-income students*. Washington, DC: Brookings Institution. Retrieved from http://www.brookings.edu/~/media/Projects/BPEA/Spring%202013/2013a_hoxby.pdf

Hoxby, C. M. (1997). *How the changing market structure of U.S. higher education explains college tuition* (Working Paper No. 6323). Cambridge, MA: National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w6323

Iceland, J. (2004). *The multigroup entropy index*. US Census Bureau. Retrieved July, 31.

Jaquette, O., & Curs, B. R. (2015). Creating the out-of-state university: Do public universities increase nonresident freshman enrollment in response to declining state appropriations? *Research in Higher Education*, *56*(6), 535–565. http://doi.org/10.1007/s11162-015-9362-2

Jaquette, O., & Parra, E. E. (2014). Using IPEDS for panel analyses: Core concepts, data challenges, and empirical applications. In M. Paulsen (Ed.), *Higher education: Handbook of theory and research* (Vol. 29, pp. 467–533). Dordrecht, The Netherlands: Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-94-017-8005-6_11

Jepsen, C., & Montgomery, M. (2009). Miles to go before I learn: The effect of travel distance on the mature person's choice of a community college. *Journal of Urban Economics, 65*(1), 64–73.

Kariel, H. G. (1968). Student enrollment and spatial interaction. *The Annals of Regional Science*, *2*(1), 114–127. http://doi.org/10.1007/BF02096181

Katz, L. (2014). *Long-term unemployment in the Great Recession (Congressional testimony)*. Washington, DC: Joint Economic Committee, U.S. Congress.

Kennedy, L. G. (2004). *Transport and environmental justice*. Bristol, UK: Policy Press.

Key, S. (1996). Economics or education: The establishment of American land-grant universities. *The Journal of Higher Education*, *67*(2), 196. http://doi.org/10.2307/2943980

Kienzl, G. S., Alfonso, M., & Melguizo, T. (2007). The effect of local labor market conditions in the 1990s on the likelihood of community college students' persistence and attainment. *Research in Higher Education, 48*(7), 751–774.

Kling, J. R. (2001). Interpreting instrumental variables estimates of the returns to schooling. *Journal of Business & Economic Statistics, 19*(3), 358–364.

Lacy, T. A., & Tandberg, D. A. (2014). Rethinking policy diffusion: The interstate spread of "finance innovations." *Research in Higher Education*, 1–23.

LeGower, M., & Walsh, R. (2014). *Promise scholarship programs as place-making policy: Evidence from school enrollment and housing prices* (No. w20056). Cambridge, MA: National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w20056.pdf

Leppel, K. (1993). Logit estimation of a gravity model of the college enrollment decision. *Research in Higher Education*, *34*(3), 387–398. http://doi.org/10.1007/BF00991851

Leslie, L. L., & Brinkman, P. T. (1987). Student price response in higher education: The student demand studies. *The Journal of Higher Education*, *58*(2), 181–204. http://doi.org/10.2307/1981241

Liu, S. (2015). Spillovers from universities: Evidence from the land-grant program. *Journal of Urban Economics, 87*, 25–41.

Long, B. (2004). How have college decisions changed over time? An application of the conditional logistic choice model. *Journal of Econometrics, 121*(1), 271–296.

Long, B., & Kurlaender, M. (2009). Do community colleges provide a viable pathway to a baccalaureate degree? *Educational Evaluation and Policy Analysis*, *31*(1), 30–53. http://doi.org/10.3102/0162373708327756

Long, M. C. (2008). College quality and early adult outcomes. *Economics of Education Review*, *27*(5), 588–602. http://doi.org/10.1016/j.econedurev.2007.04.004

Manski, C. F. (1983). *College choice in America*. Cambridge, MA: Harvard University Press.

Martorell, P., McCall, B. P., & McFarlin, I. (2014). Do public tuition subsidies promote college enrollment? evidence from community college taxing districts in texas. *US Census Bureau Center for Economic Studies Paper No. CES-WP-14-32*.

Massey, D. (2005). *For space*. Thousand Oaks, CA: Sage.

Massey, D. S., & Denton, N. A. (1988). The dimensions of residential segregation. *Social Forces*, *67*(2), 281–315. http://doi.org/10.1093/sf/67.2.281

Massey, D. S., Rothwell, J., & Domina, T. (2009). The changing bases of segregation in the United States. *The Annals of the American Academy of Political and Social Science, 626*(1), 74–90.

McCall, B. P., & Bielby, R. M. (2012). Regression discontinuity design: Recent developments and a guide to practice for researchers in higher education. In J. Smart & M. Paulsen (Eds.), *Higher education: Handbook of theory and research* (pp. 249–290). Dordrecht, The Netherlands: Springer.

McConnell, H. (1965). Spatial variability of college enrollment as a function of migration potential. *The Professional Geographer, 17*(6), 29–37.

McMillen, D. P., Singell, L. D., & Waddell, G. R. (2007). Spatial competition and the price of college. *Economic Inquiry*, *45*(4), 817–833. http://doi.org/10.1111/j.1465-7295.2007.00049.

Meade, M. S. (2010). *Medical geography* (3rd ed.). New York, NY: Guilford Press.

Miller, P. M. (2012). Mapping educational opportunity zones: A geospatial analysis of neighborhood block groups. *The Urban Review*, *44*(2), 189–218. http://doi.org/10.1007/s11256-011-0189-7

Molloy, R., Smith, C. L., & Wozniak, A. (2011). Internal migration in the United States. *Journal of Economic Perspectives, 25*(3), 173–196. https://doi.org/10.1257/jep.25.3.173.

Mullin, C. M., Baime, D. S., & Honeyman, D. S. (2015). *Community college finance: A guide for institutional leaders*. San Francisco, CA: Wiley.

Murayama, Y. (2012). Introduction: Geospatial analysis. In Y. Murayama (Ed.), *Progress in geospatial analysis* (pp. 1–12). Dordrecht, The Netherlands: Springer.

Niu, S. (2014). Leaving home state for college: Differences by race/ethnicity and parental education. *Research in Higher Education*, *Online first*, 1–35.

North, D. C., & Thomas, R. P. (1973). *The rise of the western world: A new economic history*. Cambridge University Press.

Ovink, S. M., & Kalogrides, D. (2015). No place like home? Familism and Latino/a–white differences in college pathways. *Social Science Research, 52*, 219–235.

Paulsen, M., & Toutkoushian, R. (2008). Economic models and policy analysis in higher education: A diagrammatic exposition. In John Smart (Ed.), *Higher education: Handbook of theory and research* (Vol. 23, pp. 1–48). Dordrecht, The Netherlands: Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-1-4020-6959-8_1

Peak, K. J., Barthe, E. P., & Garcia, A. (2008). Campus policing in America: A twenty-year perspective. *Police Quarterly, 11*(2), 239–260.

Perna, L. (2006). Studying college access and choice: A proposed conceptual model. In *Higher education: Handbook of theory and research* (Vol. XXI, pp. 99–157). Dordrecht, The Netherlands: Springer. Retrieved from http://link.springer.com/chapter/10.1007/1-4020-4512-3_3

Pick, J. B., Sarkar, A., & Johnson, J. (2015). United States digital divide: State level analysis of spatial clustering and multivariate determinants of ICT utilization. *Socio-Economic Planning Sciences*, *49*, 16–32. http://doi.org/10.1016/j.seps.2014.09.001

Pisati, M. (2004). Simple thematic mapping. *The Stata Journal, 4*, 361–378.

Pisati, M. (2008). *SPMAP: Stata module to visualize spatial data*. Boston College Department of Economics. Retrieved from https://ideas.repec.org/c/boc/bocode/s456812.html

Reardon, S. F., & Firebaugh, G. (2002). Measures of multigroup segregation. *Sociological Methodology, 32*(1), 33–67.

Rephann, T. J. (2007). Community college growth opportunities: Untapped potential in America's heartland? *Growth and Change*, *38*(3), 443–459. http://doi.org/10.1111/j.1468-2257.2007.00378.x

Reynolds, C. L., & DesJardins, S. L. (2009). The use of matching methods in higher education research: Answering whether attendance at a two-year institution results in differences in educational attainment. In J. Smart (Ed.), *Higher education: Handbook of theory and research* (Vol. 23, pp. 47–104). Dordrecht, The Netherlands: Springer.

Rosenbaum, J. E., Reynolds, L., & DeLuca, S. (2002). How do places matter? The geography of opportunity, self-efficacy and a look inside the black box of residential mobility. *Housing Studies, 17*(1), 71–82.

Roth, J. (2016a). *County distance database*. Retrieved May 25, 2016, from http://www.nber.org/data/county-distance-database.html

Roth, J. (2016b). ZIP code distance database – ZIP Code Tabulation Area (ZCTA) distance database. Retrieved May 25, 2016, from http://www.nber.org/data/zip-code-distance-database.html

Rouse, C. E. (1995). Democratization or diversion? The effect of community colleges on educational attainment. *Journal of Business & Economic Statistics, 13*(2), 217–224.

Schofer, J. P. (1975). Determining optimal college locations. *Higher Education*, *4*(2), 227–232. http://doi.org/10.1007/BF01569171

Shipan, C. R., & Volden, C. (2008). The mechanisms of policy diffusion. *American Journal of Political Science*, *52*(4), 840–857. http://doi.org/10.1111/j.1540-5907.2008.00346.x

Sjoquist, D. L., & Winters, J. V. (2012). Building the stock of college-educated labor revisited. *Journal of Human Resources*, *47*(1), 270–285. http://doi.org/10.3368/jhr.47.1.270

Sjoquist, D. L., & Winters, J. V. (2015). State merit-based financial aid programs and college attainment. *Journal of Regional Science, 55*(3), 364–390.

Smith, J. (2014). The effect of college applications on enrollment. *The BE Journal of Economic Analysis & Policy, 14*(1), 151–188.

Smith, J., Pender, M., & Howell, J. (2013). The full extent of student-college academic undermatch. *Economics of Education Review, 32*, 247–261.

Snow, J. (1855). *On the mode of communication of cholera*. London, UK: John Churchill.

Soja, E. (2010). *Seeking spatial justice*. Minneapolis, MN: University of Minnesota Press.

Sponsler, B., & Hillman, N. (2016). *Where you live rather than what you know? The problem with education deserts* (The Brown Center Chalkboard). Washington, DC: Brookings Institution. Retrieved from http://www.brookings.edu/blogs/brown-center-chalkboard/posts/2016/04/11-education-deserts-sponsler-hillman

Sponsler, B. A., Kienzl, G. S., & Wesaw, A. J. (2010). *Easy come, EZ-GO*. Washington, DC: Center for American Progress. Retrieved from https://www.americanprogress.org/issues/labor/report/2010/10/28/8503/easy-come-ez-go/

St. Clair, T., & Cook, T. D. (2015). Difference-in-differences methods in public finance. *National Tax Journal, 68*(2), 319–338.

Stewart, J. Q. (1941). The "gravitation," or geographical drawing power, of a college. *Bulletin of the American Association of University Professors, 27*(1), 70–75.

Strover, S. (2014). The US digital divide: A call for a new philosophy. *Critical Studies in Media Communication*, *31*(2), 114–122. http://doi.org/10.1080/15295036.2014.922207

Tate, W. F. (2008). "Geography of opportunity": Poverty, place, and educational outcomes. *Educational Researcher, 37*(7), 397–411.

Theil, H., & Finizza, A. J. (1971). A note on the measurement of racial integration of schools by means of informational concepts. *Journal of Mathematical Sociology, 1*(2), 187–193.

Thomas, J., Hughes, R. M., & McConaughy, J. (1921). The preliminary report of the Association Commission on the Distribution of Colleges. *Association of American Colleges Bulletin, 7*(4), 5–28.

Tolbert, C. M., & Sizer, M. (1996). *US commuting zones and labor market areas: A 1990 update*. Washington, DC: US Department of Agriculture. Retrieved from http://trid.trb.org/view.aspx?id=471923

Turley, R. N. L. (2009). College proximity: Mapping access to opportunity. *Sociology of Education, 82*(2), 126–146.

Ullis, J. J., & Knowles, P. L. (1975). A study of the intrastate migration of Washington college freshmen: A further test of the gravity model. *The Annals of Regional Science*, *9*(1), 112–121. http://doi.org/10.1007/BF01284992

U.S. Census Bureau. (1994). *Geographic areas reference manual*. Washington, DC: United States Census Bureau. Retrieved from https://www.census.gov/geo/reference/garm.html

U.S. Census Bureau. (2012). *U.S. census bureau projections show a slower growing, older, more diverse nation a half century from now*. Washington, DC: U.S. Census Bureau. Retrieved from https://www.census.gov/newsroom/releases/archives/population/cb12-243.html

U.S. Department of Education. (2015). *Total 12-month enrollment in degree-granting postsecondary institutions*. Retrieved April 29, 2014, from https://nces.ed.gov/programs/digest/d15/tables/dt15_308.10.asp?current=yes

U.S. Census Bureau. (2016a). *Annual estimates of the resident population*. Retrieved April 27, 2016, from http://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk

U.S. Census Bureau. (2016b). *Census tracts*. Retrieved May 23, 2016, from http://www.census.gov/geo/reference/gtc/gtc_ct.html

U.S. Census Bureau. (2016c). *Annual estimates of the resident population: April 1, 2010 to July 1, 2015 – United States – Metropolitan and Micropolitan Statistical Area; and for Puerto Rico*. Retrieved May 25, 2016, from http://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk

U.S. Census Bureau. (2016d). *Public Use Microdata Areas (PUMAs)*. Retrieved June 23, 2016, from https://www.census.gov/geo/reference/puma.html

U.S. Department of Education. (2010). *Standards for regression discontinuity designs*. Washington, DC: Institute for Education Sciences. Retrieved from http://ies.ed.gov/ncee/wwc/documentsum.aspx?sid=231

U.S. Department of Education. (2016). *Table 317–10: Degree-granting postsecondary institutions, by control and level of institution*. Retrieved June 1, 2014, from http://nces.ed.gov/programs/digest/d13/tables/dt13_317.10.asp

U.S. Office of Management and Budget. (2015). *Revised delineations of metropolitan statistical areas, micropolitan statistical areas, and combined statistical areas, guidance on uses of the delineations of these areas* (No. OMB Bulletin No. 15–01). Washington, DC: The White House. Retrieved from https://www.whitehouse.gov/sites/default/files/omb/bulletins/2015/15-01.pdf

Walker, R. E., Keane, C. R., & Burke, J. G. (2010). Disparities and access to healthy food in the United States: A review of food deserts literature. *Health & Place, 16*(5), 876–884.

Ward, M. D., & Gleditsch, K. S. (2008). *Spatial regression models*. Thousand Oaks, CA: Sage.

Zhang, L. (2011). Does merit-based aid affect degree production in STEM fields?: Evidence from Georgia and Florida. *The Journal of Higher Education, 82*(4), 389–415.

Zhang, L., & Ness, E. C. (2010). Does state merit-based aid stem brain drain? *Educational Evaluation and Policy Analysis, 32*(2), 143–165.

Nicholas W. Hillman is an Associate Professor of Educational Leadership and Policy Analysis at the University of Wisconsin-Madison. Dr. Hillman's research focuses on postsecondary finance and financial aid policy, where he has studied such topics as student loan debt, performance-based funding, and college affordability. His work on "education deserts" examines the geography of college opportunity and the role of geography in college choice. Dr. Hillman is also a member of the Scholars Strategy Network and a faculty affiliate with the LaFollette School of Public Affairs and Wisconsin Center for the Advancement of Postsecondary Education (WISCAPE).