Title
Investigating Place Attitudes in Santa Barbara, CA

Author
Davis, Adam Wilkinson

Publication Date
2015

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
Santa Barbara
Investigating Place Attitudes in Santa Barbara, CA

A Thesis submitted in partial satisfaction of the Requirements for the degree
Master of Arts in Geography

by
Adam Wilkinson Davis

Committee:
Professor Konstadinos Goulias, Chair
Professor Susan Cassels
Professor Phaedon Kyriakidis

December 2015

The thesis of Adam Wilkinson Davis is approved.

Susan Cassels

Phaedon Kyriakidis

Konstadinos Goulias, Committee Chair

December 2015

Acknowledgements

Thank you Kostas for all your help with this thesis and my first two-and-a-half years of grad school! Thank you Susie and Phaedon for providing valuable feedback in the later stages of my writing process. I'd also like to thank all of my friends both at UCSB and elsewhere for making my life enjoyable and meaningful; particularly Kelsey Bisson, Crystal Bae, and Heather Sparks, who provided helpful feedback on one section or another of this thesis, and Jay Lee and Elizabeth McBride, the wonderful denizens of the Geotrans lab.

I dedicate this thesis to my parents, who've raised me with endless kindness and patience. I've been a geographer from day one, and I'm honored to continue our odd family business.

Investigating Place Attitudes in Santa Barbara, CA

Adam Wilkinson Davis

Abstract

This thesis investigates the relationship between place attitudes and measurable place attributes in Santa Barbara, CA. People's relationships to places form a key component of travel behavior and decision making, but they have been the subject of limited empirical study. In addition to improving our understanding of place attitudes, I investigate appropriate ways to model a cross-classified dataset with spatially autocorrelated ordinal responses. Data sources include a spatially-constrained place attitudes survey, parcel-based land use, business establishment records, and a collection of local geotagged Tweets, from which a spatial measure of happiness was extracted. The relationships among these variables are investigated using a cross-classified multilevel ordinal regression model, which finds several significant relationships between place attitudes and available opportunities, land use, and natural amenities. I also find that while Twitter data may provide another way to measure place attitudes, more work is needed to develop useful variables from it.

Table of Contents

Figures

Tables

Chapter 1 – Introduction

In this thesis, I hope to improve our understanding of how place attitudes vary and how well individuals' opinions about places reflect measurable attributes of these places. I model responses to a coarse-scale place attitudes survey as a function of a range of related measurable attributes of the region, accounting for the interaction among place attitudes and variation between people and among spatial units. This thesis will address the ways place is studied in both quantitative geography and more theory-driven branches of the discipline. I will describe, justify, and present a cross-classified multilevel structural equations model for ordered responses that can be used to study spatially-stratified place attitude data.

Though the importance of place to human spatial decisions relating to transportation is the main justification for this research, my interest in place stems originally from my background in human geography. While quantitative place geography has so far focused entirely on the relationship between individual people and individual places or on the ways a given place is important to all people, it is also important to understand the social dimensions of place. Geographic research into place has shown that Place (and specific places) is neither fully social nor uniquely personal. Though most quantitative studies of place (including the data and models I present here) center on personal understandings of place, it seems especially important to develop further research into aggregate and social understandings of place. Much as social networks and social structures are vital to understanding people, the connections between places and the different ways groups of people interact with places is vital to a complete understanding of place.

Though the place attitude metrics included in this study address only a small section of the geographic understanding of place, this study is important because it seeks to link people's attitudes about places with some consistently measurable reality. The structure of the GeoTrips survey limits my ability to extract definitive findings from a model, but this thesis investigates methods, data sources, and key variables that can be used today to draw inferences and could be used for future related work.

The structure of this thesis is as follows:

- In Chapter 2, I review the literature of place. I start by providing a brief overview of Sense of Place. Though it originates from phenomenology and a deeply personal understanding of place, most attempts to quantify place have focused on specific aspects of sense of place. I discuss the efforts of the UCSB GeoTrans lab to study sense of place and place attitudes as they effect transportation. Next, I review the broader geographic literature on place, with a special focus on researchers who seek to understand the links between places and who tie place to broader theoretical structures. The work discussed here suggests that much as people are best understood within the context of their social networks and society as a whole, places should be understood through the social processes that shape them, not as atomic units that can be studied in isolation.

- In Chapter 3, I describe the data used to estimate the model. First, I review the GeoTrips survey, which is the main data source for my research. This section describes the spatial structure of the place attitude responses and maps general trends in the data. A main focus of this thesis is to see what consistently measurable attributes of the built and natural environment can be tied to place attitudes. The second half of the chapter describes the various external data sources brought together for this purpose.
- In Chapter 4, I discuss the models I estimated. First, I provide a detailed overview of several key aspects of the model structure, with special focus on the ways this model addresses the multiple categories of repeated measures present in the dataset. By comparing the results of simplified versions of this model, I hope to justify the use of a relatively complex cross-classified design. Next, I review model limitations in addressing the structure of the dataset and propose potential solutions for handling this type of data better in the future.
- In Chapter 5, I interpret the outputs of the full cross-classified model and a simpler model that does not account for the spatial structure of the data, as well as a providing an additional cross-classified model that incorporates data from Twitter. I then discuss the significance of these results in several ways: with respect to our understanding of the Santa Barbara region, with respect to how a fully developed version of this model could improve urban planning, and in terms of the geographic understanding of place.

- A final chapter provides an overall summary of the research and discusses future work that could build on the GeoTrips hexagon survey design and other data products developed for this study.

Chapter 2 – Place Literature

Sense of Place in Quantitative Geography

Space and place are two key concepts in geography that frame the world in which we live. Our understanding of space, a realm of measurable variables and clear boundaries, lends itself well to study, but place is fuzzy and difficult to pin down. Though human behavior can be modeled in purely spatial terms, any attempt to bring attitudes and other subjective aspects of human spatial behavior into the equation necessarily raises the issue of place. Understanding how people feel about places is a key to predicting where people will go to shop, eat, and socialize, but these emotions are inherently subjective and difficult to quantify.

Sense of Place (the powerful connections between people and places) has a long history in human geography, but because the theory operates in realms that are either deeply personal or social but invisible, it has only recently been considered as a quantifiable aspect of human society that might be included in behavioral models. Among the key theorists responsible for the contemporary geographic understanding of sense of place are Yi Fu Tuan, who defined sense of place as individual humans' "affective ties with the material environment" (Tuan, 1974), and Doreen Massey, who incorporated a Marxist conception of power and difference into her definition of place (Massey, 1991). Though the geographic understanding of place has a fairly broad reach, most attempts to quantify place have focused on specific aspects of individuals' sense of place. Kathleen Deutsch's dissertation contains an excellent literature review of the history and development of research into sense of place (Deutsch, 2013), and her work led me to many of the sources discussed in this section.

The gradual resurgence of quantitative human geography and increased interest in place within the fields like environmental psychology and sociology has led to many attempts to quantify certain aspects of sense of place. Among the earliest research in this field was carried out by David Canter, who sought to bring the phenomenological aspects of place "into a form that is amenable to empirical examination" (Canter, 1983), and Reg Golledge who sought to apply these efforts to understanding measurable human behavior (Bolton, 1989; Golledge & Stimson, 1997). Specific aspects of sense of place that researchers have focused on include place identity, which is "a person's identity with relation to the physical environment" (Proshansky, 1978), place attachment, which is defined as "the positive bond that develops between a person and their environment" (Low & Altman, 1992), place dependence, which is defined as the "perceived strength of association between a person and a place" (N. Stokols, 1981) and place satisfaction "a person's level of satisfaction with the services, environment and needs provided for by a specific place" (Stedman, 2003).

In addition to the difficulty of knowing what questions to ask about place, it is also unclear where one should ask. Places often have unclear boundaries, making it difficult to include place-based metrics in spatially explicit models. One typically geographic response has been to conceive of places hierarchically (both in terms of spatial scale and scale of interaction), with smaller personal places nesting within larger shared places (Rapoport, 1977). Uncertainties in spatial scale are coupled with the difficulty of pinning down places' constantly changing meanings. Early sense of place research in environmental psychology suggested that understanding the temporal components of person-place interaction is at least as important as pinning down space, since people's interactions with places change constantly in interaction with the built and natural environment (Canter, 1977, 1983), but few place researchers have explicitly sought to study change in sense of place over time. As Deutsch notes (2013), spatial measurement scale has been a major question in the development of Sense of Place metrics, with no single scale being strongly preferred. Recent quantitative place work has investigated people's sense of place with respect to their homes and vacation homes (Jorgensen & Stedman, 2001, 2006), their changing relationship to changing neighborhoods (Brown & Werner, 2008), their urge to feel connected to nature in parks (Davenport & Anderson, 2005; Smaldone, Harris, & Sanyal, 2005), and their lingering attachment to specific ethnic meanings of cities that changed hands in war (Lewicka, 2008).

One way to make place tractable is to collect place attitudes data about mapped areas and aggregate measures of certain elements of place. Deutsch's GeoTrips Survey asked respondents about several aspects of their attitudes about parts of the Santa Barbara area, using a grid of 23 hexagons (each roughly two mile across) to constrain their responses (Deutsch, 2013). By tying responses directly to a consistent spatial frame, this study should eliminate much of the uncertainty caused by differences in the perceived boundaries of places. The use of a single, fairly coarse, spatial scale limits the aspects of place included to ones that are meaningful at this scale, namely perceived attractiveness, opportunity, and danger, and the respondent's familiarity with the region.

Much of the work in UCSB's Geotrans Lab has focused around measuring Sense of Place and applying the results to substantive questions in transportation geography. This work has shown that studying place at relatively broader scales can greatly expand our understanding of people's day-to-day activities outside the home (Deutsch & Goulias, 2010; Deutsch, Yoon, & Goulias, 2013). Along these lines, a recent study in Santa Barbara, CA, asked respondents to link their destination choices to a variety of location attributes (Deutsch, Ravualaparthy, & Goulias, 2013). This work has shown that the affective and emotional aspects of a destination exert a strong influence on decisions of destination selection. In this same research, we were able to create an individualized index of attraction for four different aspects of place attitudes (Deutsch, 2013; Deutsch & Goulias, 2013). Recent work has also sought to link sense of place directly to travel behavior by university students (Deutsch, 2013; Lee, Davis, & Goulias, 2015).

The work done by GeoTrans ties into a larger push by transportation researchers to directly investigate human spatial decision making (Ferguson & Kanaroglou, 1995; Hunt, Boots, & Kanaroglou, 2004; Paleti, Bhat, Pendyala, & Goulias, 2013; Pellegrini & Fotheringham, 2002). Accounting for attitudes and perception in spatial discrete choice models can allow us to construct simulated decision makers that are more heterogeneous and realistic. This can be accomplished via latent factors within an integrated system of structural equations with discrete choice models (Ben-Akiva et al., 2002, 2002; Bhat & Dubey, 2014).

This thesis extends Deutsch's work by relating subjective place attitudes responses to measurable human and natural attributes of the region. Because the dataset contains multiple interrelated response variables, I use a structural equations model (Kuppam & Pendyala, 2001). Because of repeated measures at the levels of both the 561 individual respondents and the 23 response hexagons, I employ a cross-classified multilevel structural model (Bhat, 2000; Fielding & Goldstein, 2006), estimated in Mplus. Because the added complexity of the cross-classified model may make it seem daunting, I include a simpler model that accounts for person-person (but not hexagon-hexagon) variability at a second level. Though the two models produce very similar estimated coefficients, the simpler model returns standard errors that are unrealistically small because it does not account for the non-independence of hexagon-level responses (Fielding & Goldstein, 2006).

An Overview of Place Geography

Even when they do not explicitly study sense of place, most human geographers interact with the concept of place. Work on place has come from a broad range of perspectives in human geography, from the origins of Sense of Place in phenomenology to cultural landscape research to the various branches of critical human geography. Some researchers make more concrete attempts to understand all the processes at work in a specific place while others engage in more abstract consideration of the relative importance of place in human life. In addition to being quite diverse in approach, these studies vary considerably in the types of places they prefer to study and the scales at which they study them. Cultural landscape geographers love everyday “vernacular” places and will sing rhapsodies to the many eras of American barn building; Foucault was obsessed with the oppressive power of institutional places.

The various branches of human geographic work on place all point to one clear shortcoming of recent quantitative work on sense of place: places and people do not exist in isolation, instead they operate within the context of powerful economic, social, and environmental processes. The processes linking places to each other are as important as an individual’s sense of place. To understand how this literature might be brought into conversation with quantitative sense of place research, in this section I will cast a somewhat broader net than Deutsch did, providing an overview of how geographers have addressed place as a product and producer of human processes.

It is somewhat surprising that so much quantitative place research emerged from Tuan’s work, given that phenomenology seems at first glance to be at least as averse to empirical study as structural Marxism. Though his later work acknowledges the links between people and social processes much more directly (notably *Dominance and Affection*, in which he traces how people created pets to suit our need for happiness (Tuan, 1984)), Tuan’s early work constructs places as somewhat unique and deeply personal, albeit with the power to shape an individual’s interaction with the outside world (Tuan, 1974). In part because Tuan’s *Topophilia* focused on the direct links between an individual person and a personally meaningful place, it provided a valuable stepping stone for quantitative place research.

Though cultural landscape studies emerged more or less in parallel with work on sense of place, it owes more to architectural theory than phenomenology. The work of J.B. Jackson sought to understand the past and present of a particular scene; though Jackson used the word "landscape" and focused almost exclusively on visual sensory information and everyday "vernacular" landscapes, cultural landscapes as they have come to be studied broadly match most concepts of place – finite spaces into which people build meaning (Wilson & Groth, 2003). Great attention is paid to how the visual aspects of a place reflect its history; for instance cultural landscape geographers have looked at how migration changes the look and feel of small towns (Ghose, 2004), how capital investment in extractive industries produced and destroyed company towns (Buckley, 1997), how improved infrastructure and a taste for suburbia created the modern office park (Mozingo, 2011), or how residential hotels created a unique and fragile landscape for lower-class urbanites (Groth, 1994). All of these works describe unique places that exist at the intersection of human processes and the environment.

Landscape geographers argue that vernacular landscapes are important because they represent the sorts of places with which most people usually interact. By linking their work to other bodies of theory, writers in this field have shown how everyday places produce and are produced by social and economic interaction. Ultimately, cultural landscape studies provides a full understanding of neither individual people's relationships to specific places nor the links between place and global processes – and it seems fair to ask whether it really matters that small town cemeteries change over time to reflect the changing tastes of its residents (Francaviglia, 1971). Though cultural landscape research does not always produce satisfying results, the tools it developed – generally coupled with archival research, analysis of local sources, and interviews – have become widespread in human geography. These methods can provide valuable insights into the workings of a region, making it possible to engage in more meaningful quantitative research. Fortunately, though landscape geographers formalize visual methods to a degree, "reading the landscape" is something that many people do naturally.

Marxist geography is generally much less place-focused than other branches of human geography, but the field has produced significant work on the economic relationships between places. Structure-oriented Marxist geographers like David Harvey are often skeptical of place, which they consider much less important than the spatial processes of global capitalism. To Harvey, places exist primarily as loci of the class-based power struggles that govern society (Harvey, 1993), and he critiques place-centered research both for its focus on the local and for the unclear boundaries of places (Harvey, 2006). Harvey's student Richard Walker, another prominent Marxist economic geographer, takes a more favorable view of place. Much of his early work addressed the significance of place to economic geography through innovation and location decisions (Walker, 1989), and Walker's later work focuses largely on the significance of place in the San Francisco Bay Area ("Richard Walker," 2015). In his book The Country in the City, he studies the ways in which open-space preservation and other environmental efforts are so central to the meanings of the San Francisco Bay Area as a place (Walker, 2009). While countless cultural landscape geographers have focused on either small rural towns centered around extractive industries or big cities centered on finance, it took Marxist critical geographers like Walker and Gray Brechin to see the links between the two, arguing that economic surplus flows from the former to help produce the grandeur of the latter (Brechin, 2007).

Other Marxist geographers explicitly address place in their work. Doreen Massey's "A Global Sense of Place" was one of the first pieces to call into question assumptions made by early place geographers that places were either completely personal (and thus had no shared meanings) or had a single universal meaning (Massey, 1991). She argues that different people truly understand and experience places differently and that these differences often reflect differences in class, gender, and ethnic background. Massey also expanded this view of place to consider the ways these differences in meaning are a product of differing place histories, resulting in groups contesting the "authentic" meaning of a place (Massey, 1995).

Recent discussion of place in critical urban geography unites Massey's understanding of class-based contested place meanings and authenticity with the work of mid-20th century social theorists to study urban change, gentrification, and access. Michel Foucault's critique of separation of history from geography, Pierre Bourdieu's socio-spatial concepts of field and habitus, and Henri Lefebvre's extension of Marxism to describe the social production of space have profoundly influenced the development of human geography. These researchers ask whether gentrification can only destroy place authenticity or whether it can create new authentic meanings (Zukin, 2011), how new communities relate with changing places (Ley, 2003; Douglas, 2012), and whether planners can produce meaningful places (Clarke, 2012). The intense personal significance of home is widely recognized, but for nomadic groups and homeless people, home exists as a place without a fixed location or the sense of privacy and acceptance that less marginalized people experience (Johnsen, May, & Cloke, 2008; Sparks, 2010; Convery & O'Brien, 2012).

Efforts to protect places in order to preserve local culture and promote tourism (which seem to be conflicting goals) often encounter fierce debates over the true meaning of place. Work at the intersection of landscape and critical geography demonstrates how various actors contested their own conceptions of a place's "authentic" cultural meaning when rebuilding after an earthquake (Puleo, 2010) and collecting support for a UNESCO world heritage site (Puleo, 2013). Other researchers have found protection of place authenticity to be the main goal of preservation movements, which they find problematic due to its denial of the heterogeneous (Certoma, 2009) and changing (Horlings, 2015) meanings of places.

While place can be divisive, it can also unite. Not all preservation efforts run into fierce disagreements, and successful efforts to present a single apparent identity can make places that are attractive to residents and tourists alike – note the faux-Mission style architecture that makes Santa Barbara's downtown instantly identifiable. Additionally, geographers recognize that place shapes people – much of political geographer John Agnew's work has focused on the role of place in politics (Agnew, 2002; Agnew & Duncan, 2014). He investigates how spatial differences in political activity arise from people's attachment to the history and meaning of their respective home region.

Variability in place meanings among different groups of people interacts with individual variability. Deutsch notes the unique significance of familiarity (as opposed to the other measures included in GeoTrips), since it indicates "both the level of exposure to the region, and the attachment of meaning … that are integral to patterns of movement and decision making for activities" (Deutsch, 2013) and notes that Golledge and Spector's work (1978) provides some evidence of this.

Place is soft, but important. The rich and diverse geographic literature of place incorporates a wide range of theoretical perspectives, but limited empirical work. Efforts to quantify place have focused almost exclusively on a humanistic concept of sense of place that largely ignores the ways human difference and the interaction among places shapes the way a given individual interacts with a place. Qualitative and theory-driven human geographers who investigate place provide a set of important questions that have been somewhat absent from quantitative geography and suggest tools that could be valuable in answering these questions.

People contest the meanings of all sorts of places, whether spectacular and culturally significant or vernacular and personal. These differences in meaning spring from differences in access, history, language, and acceptance that affect the way people interact with places. Just like people should be studied in the context of their social network, places are not units that can be studied in isolation. Studying people's behavior with respect to places requires understanding the ways difference shapes places and the ways difference is enacted through behavior.

Further research on the effects of race, class, ethnicity, gender, personal biographies and other forms of difference on activity spaces and attempts to measure differences in place over space (which this work feeds into) could help bridge the gap between geographic theory and quantitative research. Additionally, it is important to note that not all elements of place are humanistic – economic geographers often ask questions about the significance of place, though often using different language.

Chapter 3 – Data Sources

The bulk of the data used in this thesis comes from Deutsch's GeoTrips Survey, which collected information about 561 residents of southern Santa Barbara County, California from May to July, 2012. Respondents were asked to report their attitudes about the region using a grid of 23 hexagons. In addition to data from the GeoTrips survey, this study draws from a wide array of sources to determine which factors affect people's opinions of places at the scale of our hexagons. All of these variables are aggregated to the hexagon level.

GeoTrips Data

A general description of the people who responded to the survey is shown in Table 1. Of the 561 respondents, 238 (42.4%) were male and 323 (57.6%) were female. The mean age of all respondents was 48 years and the median was 49. Though the survey attempted to gather a random sample, the resulting sample does not match the population as a whole in terms of gender and age. The largest cities in this region are Santa Barbara and Goleta, but respondents were also drawn from smaller communities such as Montecito, Isla Vista, and Summerland; respondent home locations are shown in Figure 1 and broadly match the spatial distribution of people in the study area. The GeoTrips survey was conducted online, with respondents recruited by mail and email. The survey design is further described in several papers by Deutsch (Deutsch, 2013; Deutsch, Ravualaparthy, et al., 2013).

In addition to collecting demographic data and other information on decision making preferences, the survey included an interactive mapping exercise that asked respondents to report their attitudes about different parts of the Santa Barbara area. Though people's attitudes likely vary continuously over space, it was necessary to constrain these responses to specific bounded spatial units. To this end, the survey provided a tessellated grid of 23 hexagons, each 4 km across (Figure 1). Hexagons were chosen for the survey because they have the lowest edge-effects of any shape that can completely cover a region (Aitken & Prosser, 1990; Montello, Friedman, & Phillips, 2014). For each of the 23 hexagons, each respondent provided their agreement on a seven-point Likert scale for each of the following statements (Deutsch, 2013):

- This is an attractive area of Santa Barbara.
- I am familiar with this area of Santa Barbara.
- This area provides me with a lot of opportunities to do things I like to do.
- This is a dangerous area of Santa Barbara.

Table 1 GeoTrips Survey Respondent Age Breakdown

Age Group	Count	Percent
Age 18-25	56	10.0%
Age 26-39	132	23.5%
Age 40-64	257	45.8%
Age 65+	105	18.7%

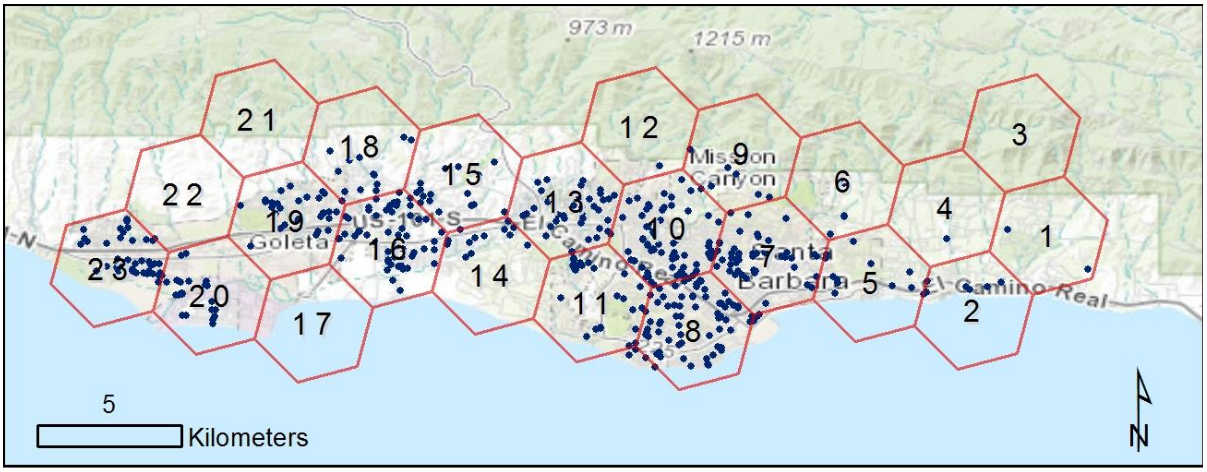

Figure 1 GeoTrips Hexagons and Respondent Home Location

Table 2 Place-Attitudes Response Totals. Values correspond to percent of total ratings for each place attitude. 561 respondents rated all 23 hexagons for each variable, so n=12,903 in each column.

Response	**Attractive**	**Familiar**	**Opportunity**	**Danger**
Strongly Agree	31.2%	28.6%	18.7%	2.7%
Agree	24.2%	16.9%	14.4%	3.3%
Slightly Agree	16.8%	17.7%	16.9%	7.1%
Neutral	18.5%	16.0%	30.3%	27.1%
Slightly Disagree	4.2%	7.4%	6.8%	11.7%
Disagree	2.9%	6.3%	6.0%	16.8%
Strongly Disagree	2.1%	7.1%	6.8%	31.2%

The breakdown of responses in each category is shown in Table 2. Each of the hexagon response variables skews fairly positive, which means they should generally not be interpreted as numeric data (Fielding, 1997; Grilli & Rampichini, 2011). The final model will account for the ordinal structure of the data using a probit link function, but it is useful to look at their distribution in space as well, which necessitates treating the ordered responses as continuous variables. Figures 2, 3, and 4 compare each hexagon's average score (with Strongly Disagree scored 1 and Strongly Agree scored 7) for a pair of variables. Hexagons with an above-average score for each variable are colored blue, a below-average score for each red, and mismatched hexagons are green or yellow. Plotted bubbles are scaled by number of business establishments in the hexagon.

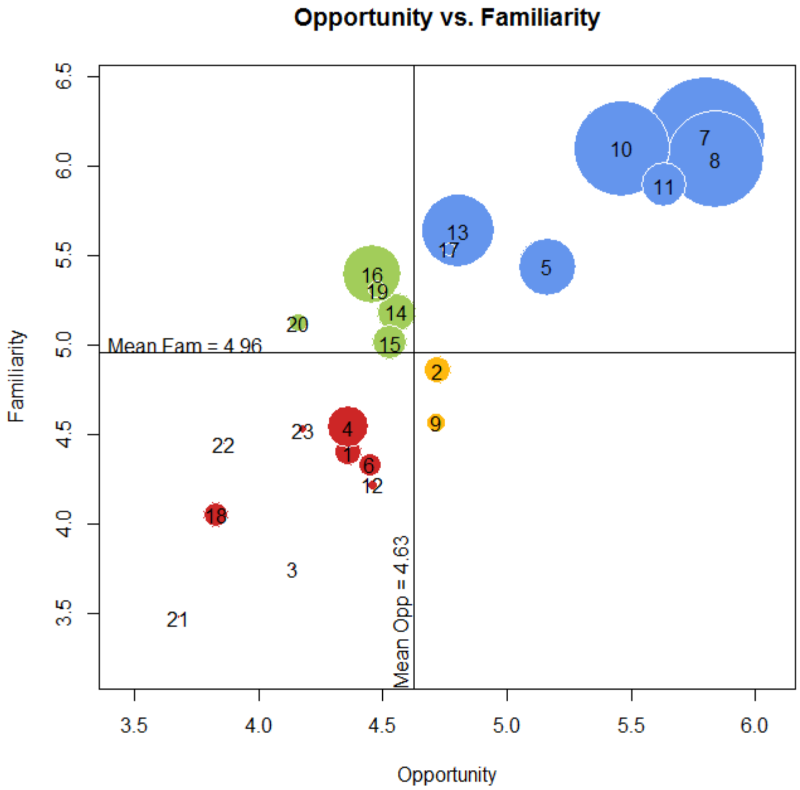

Figure 2a GeoTrips Opportunity vs Familiarity bubble plot. Each circle corresponds to one hexagon's mean response for Familiarity and Opportunity. Circles are scaled by total number of businesses in each hexagon. Circle color determined by hexagon's relationship to overall mean score for each question.

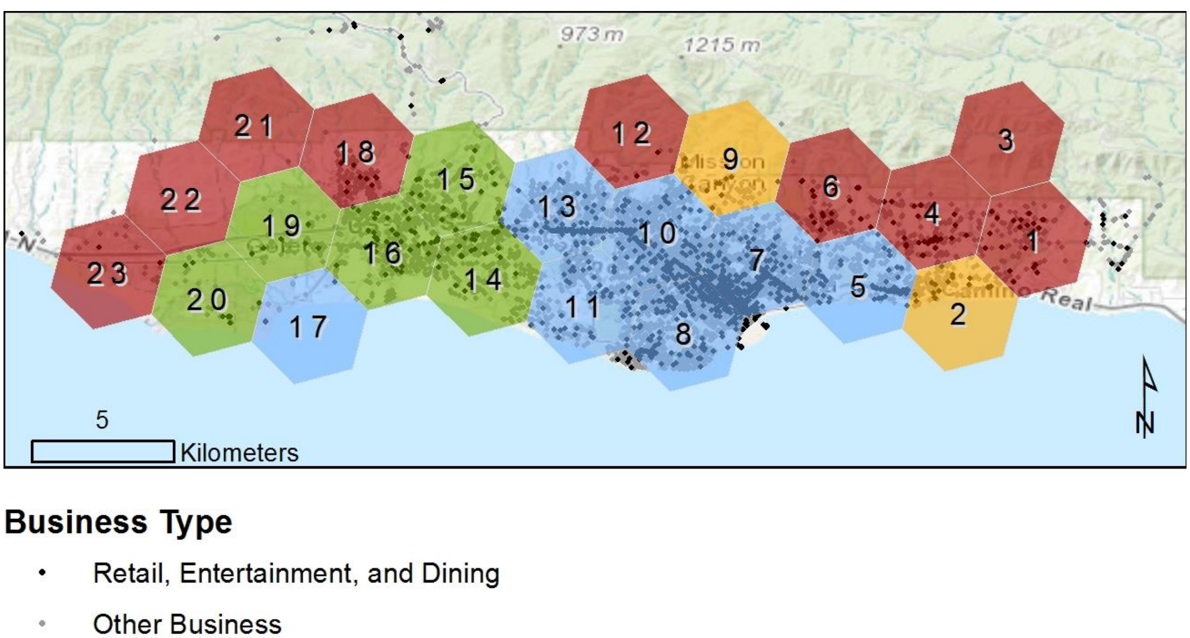

Figure 2b Opportunity vs Familiarity, with Business Locations. Hexagon colors and labels match figure 2a.

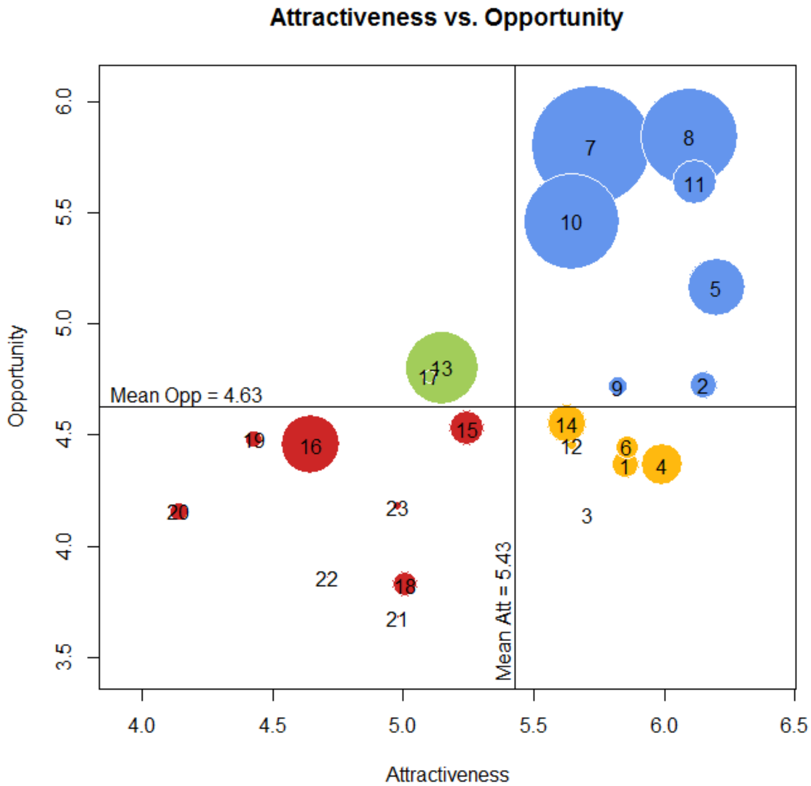

Figure 3a GeoTrips Attractiveness vs Opportunity bubble plot. Each circle corresponds to one hexagon's mean response for Attractiveness and Opportunity. Circles are scaled by total number of businesses in each hexagon. Circle color determined by hexagon's relationship to overall mean score for each question.

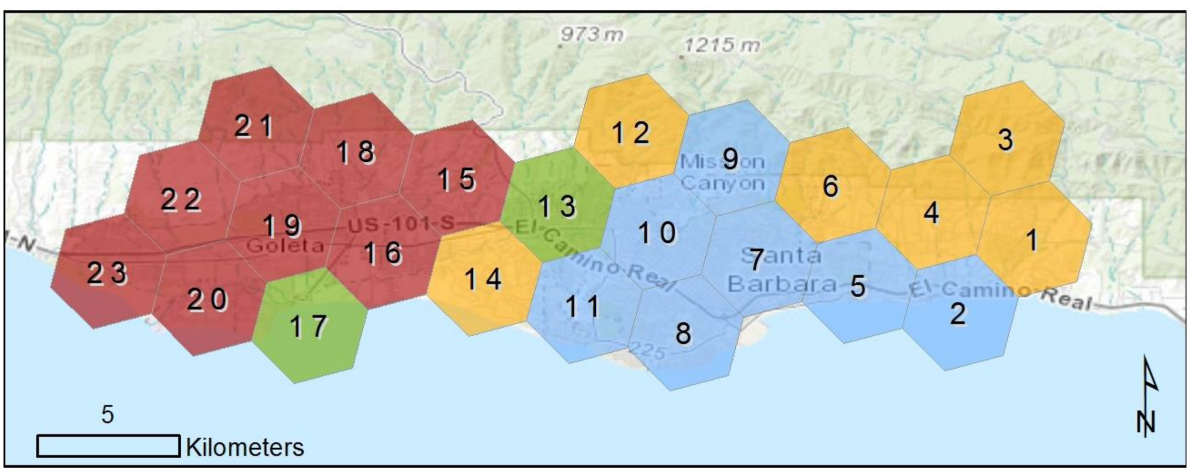

Figure 3 Attractiveness vs Opportunity. Hexagon colors and labels match figure 3a.

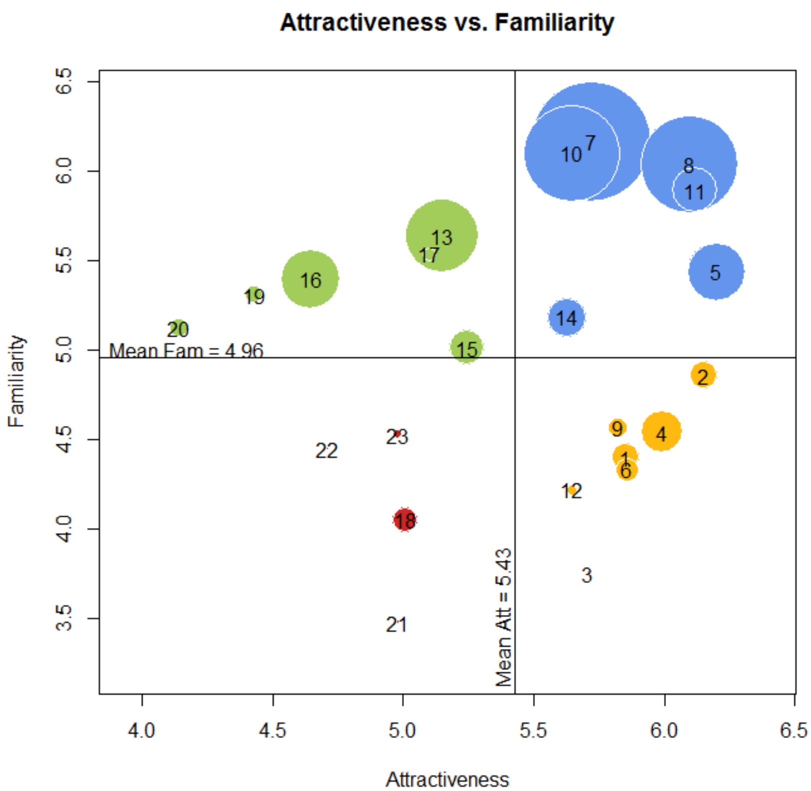

Figure 4a GeoTrips Attractiveness vs Familiarity bubble plot. Each circle corresponds to one hexagon's mean response for Attractiveness and Familiarity. Circles are scaled by total number of businesses in each hexagon. Circle color determined by hexagon's relationship to overall mean score for each question.

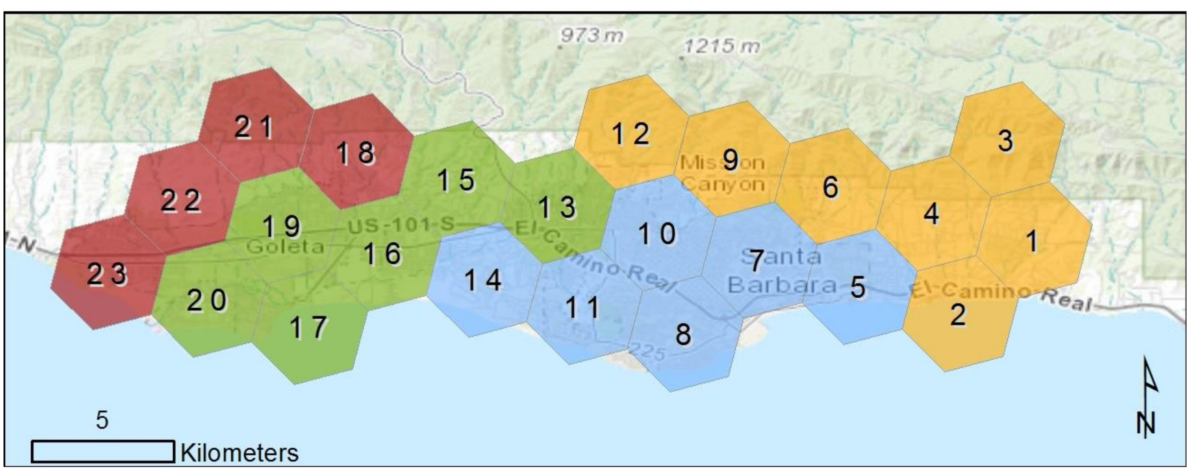

Figure 4 Opportunity vs Familiarity. Hexagon colors and labels match figure 4a

These plots and maps show that there is a positive relationship between hexagon-level attractiveness and opportunity, and that people are generally more familiar with hexagons that are attractive and opportunity-rich. Opportunity and familiarity appear to have a substantially stronger relationship than the other variable pairings. Additionally, there is a strong but inconsistent relationship between perceived opportunity and number of businesses; hexagons 7, 8, and 10 all contain large numbers of businesses and are perceived as presenting large numbers of opportunities, but hexagons 13 and 16, which cover a lower-rent area of Santa Barbara and the heart of working-class Goleta, respectively, are perceived as presenting few opportunities despite being relatively business-rich. The differences in location between businesses that customers must visit and other businesses, such as consulting firms and construction companies (shown in Figure 2) may account for some of this inconsistency. The strong relationship among familiarity, attractiveness, and opportunity suggests that this dataset would be well suited for structural equations modeling, which is designed to address the relationships among multiple dependent variables.

Business Establishments

Of the place attitudes examined in GeoTrips, opportunity and familiarity seem to have the strongest direct link to travel behavior. Opportunity is a major cause of travel and Familiarity is likely an effect. Though people occasionally travel only for the enjoyment of travel – as has been noted both by landscape geographer J.B. Jackson (Jackson, 1958) and transportation researchers investigating the positive utility of travel (Arentze & Timmermans, 2005; Mokhtarian, 2005) – most trips are intended to allow the person to fill some need that they cannot meet at home. Trips to work and school are essentially fixed in time and place (Golledge & Stimson, 1997; Miller, 2005), but trips for shopping or entertainment can be much more flexible. Places that provide a more numerous and wide range of opportunities should generally attract more trips, and the more people visit a place, the more familiar with it they should become. To this end, we include measures of opportunity density and diversity in our models.

The importance of opportunity density seems fairly clear – places with more opportunities should generally attract more visitors, but diversity also has important impacts on people's interaction with places. Urbanists and critical planners since Jane Jacobs have supported mixed-use development as a way to revitalize cities (Jacobs, 1961; Grant, 2002), and work in travel behavior has provided empirical backing to many of the perceived benefits. Cervero investigated the relative importance of land use density, diversity, and design on vehicle travel, concluding that dense, mixed-use neighborhoods discourage travel by car (Cervero & Kockelman, 1997) and other work has found similar results (Ewing & Cervero, 2001; de Abreu e Silva, Golob, & Goulias, 2006). Because it should help reduce greenhouse gas production, mixed use development is also now a major consideration in regional planning in California under SB 375 (Steinberg, 2008). General land-use diversity plays a substantial role in shaping people's relationships with a region, but the diversity of available opportunities has not been as widely addressed. Areas that present a more diverse set of opportunities should appeal to a wider range of people, which should be reflected in higher familiarity and perceived opportunities.

To find one measure of opportunity density, we extracted the total number of business establishments in each hexagon at each level-2 Standard Industrial Classification (SIC) codes. SIC codes provide a hierarchical classification scheme that can be used to differentiate businesses. The original data comes from the 2010 NETS Dataset (a database that tracks the location, employees, and industry type of every business establishment in the United States). The portion of the Santa Barbara Area covered by our hexagons contains 13,802 business establishments, with hexagon totals ranging from 0 in mountainous hexagon 3 to 3,517 in downtown Santa Barbara hexagon 7. We aggregated several SIC classes into a broad Consumer category that contains SIC classes covering retail (classes 53, 54, 56, 57, and 59), eating and drinking (58), and entertainment (78, 79, and 84). Our model included the total number of consumer establishments in each hexagon. Table 3 contains hexagon-level descriptive statistics of the aggregate Consumer Establishment counts. The mean is substantially higher than the median, which indicates that the number of consumer-serving businesses in each hexagon has a substantial positive skew – essentially that these businesses are heavily concentrated in a few hexagons. Because of this, it may be worthwhile to use the logarithm of business establishments instead.

To investigate the effects of opportunity diversity, we calculated the Shannon Entropy (Equation 1) of businesses in each hexagon based on level-2 SIC codes. This measure quantifies the uncertainty in the class membership of a randomly selected entity and provides a relatively easy-to-use measure of diversity (DeJong, 1975). Diversity of opportunities may play an important role in attracting visitors to a neighborhood, though diversity of business establishments by SIC-category may be an imperfect measure of opportunity diversity, particularly because many level-2 SIC codes represent similar or closely related industries, so a diversity metric that treats these groups as distinct may be misleading. Nevertheless, it is included in this model because it has a significant effect in the model.

To calculate diversity, we determine proportion of a hexagons businesses in each of the 80 level-2 SIC categories ($p_i \; where \; i \in R$). For each class present in a hexagon ($p_i > 0$), we multiply class proportion by its natural log, then sum across all level-2 SIC categories (Equation 1). Shannon Entropy is higher when more classes are present in similar proportion and lower when few classes are present or almost all businesses belong to one or two categories. Shannon entropy of business establishments in our hexagons ranged from roughly 1 to 3.5, with almost all hexagons falling between 2.3 and 3.5. Hexagons lacking any business establishments were assigned a Shannon value of 1.

$$Shannon\ Entropy = -\sum_{i=1}^{R} p_i \ln p_i$$

Equation 1 Shannon Entropy Calculation

Table 3 Business Establishment Variable Descriptive Statistics (hexagon totals)

	Consumer Ests.	Est. Diversity
Mean	82.6	2.61
Median	18	2.71
Standard Deviation	146.6	0.697
Minimum	0	*1* (1.10)
Maximum	592	3.47

Classified Parcel Data

Though business establishment counts and diversity are a useful measure of opportunity density, business is not the only reason people leave their houses. Land use likely also has a substantial effect on people's attitudes about an area, in terms of both opportunity (parks and beaches) and attractiveness (open space and ocean views).

Remote sensing data is a common basis for land cover data, but other data sources are often required in order to estimate land use, which is generally not visible from above. Because we had access to another form of land use data for the Santa Barbara region, we opted to use it instead. A remote sensing product would provide a clearer sense of paved area / vegetated open space, which could be closely related to attractiveness, but the classified parcel data is a reasonable solution for most other land use / land cover types. To create hexagon-level aggregate land use data, we retrieved tax assessment parcels for Santa Barbara County ("County GIS Spatial Catalog," 2015). The data source assigned each parcel to one of 82 land use categories, so we sought to reclassify them into a smaller number of categories to make this tractable for our analysis. In order to do this, we developed a "crosswalk" as shown in Table 4.

Table 4 shows the correspondence between the 82 categories present in the original dataset and the 18 used here, along with total areas. Categories were matched manually with a goal of differentiating between activity types experienced by visitors. For example, unnecessary detail was eliminated by reclassifying 14 individual categories into a broader “Commercial” class. For quality control, some manual reclassification was performed for oddly classified parcels, most of which were along the railroad right of way.

The parcel shapefile covered the entirety of each hexagon except for roads and the ocean. To add these, we performed a spatial union between the parcels and a California coast shapefile. All areas not covered by a parcel but on land were classified as road. We then performed a union between this file and our hexagons; all areas covered neither by a parcel nor land area were classified as ocean. We ended up with 20 classification categories, plus the road and ocean categories, with total areas per category for each hexagon. As an input to our model, we further aggregated some of these and converted total areas to percent of land area, to account for differences in land area among the hexagons. Total area (in hectares) in each category (as well as area covered by roads and the ocean) within one of the GeoTrips survey hexagons are also shown in Table 4.

Table 4 Parcel Land Use Categories Crosswalk and Total Areas

Output	**Input**	**Hectares (% of Tot.)**
Single Family Housing	Single Family Residence	8,580 (29.6%)
Multi-Family Housing	Apartments, 5 Or More Units; Condos, Community Apt Projects; Mobile Home Parks; Mobile Homes; Residential Income, 2-4 Units; Rest Homes	1,176 (4.1%)
Mixed Use	Mixed Use-Commercial/Residential	16 (0.1%)
Commercial	Banks, S&Ls; Bed And Breakfast; Commercial (Misc); Commercial And Office Condos (including commercial planned unit developments) Day Care; Department Stores; Other Food Processing, Bakeries; Restaurants, Bars; Retail Stores, Single Story; Shopping Centers (Neighborhood); Shopping Centers (Regional); Store And Office Combination; Supermarkets; Wholesale Laundry	454 (1.6%)
Lot Commercial	Auto Sales, Repair, Storage, Car Wash, etc; Drive-In Theatres; Parking Lots; Petroleum And Gas; Service Stations	161 (0.6%)
Hotels	Hotels	98 (0.3%)
Office	Office Buildings, Multi-Story; Office Buildings, Single Story; Professional Buildings	240 (0.8%)
Public	Hospitals; Public Buildings, Firehouses, Museums, Post Offices, etc.	733 (2.5%)
School	Colleges; Schools	841 (2.9%)
Religious	Churches, Rectory	232 (0.8%)
Indoor Recreation	Auditoriums, Stadiums; Bowling Alleys; Clubs, Lodge Halls; Dance Halls; Recreation	96 (0.3%)
Golf & Riding Ranges	Golf Courses; Horses; Race Tracks, Riding Stables	637 (2.2%)
Open Space	Beaches, Sand Dunes; Institutional (Misc); Miscellaneous; Mortuaries, Cemeteries, Mausoleums; Parks; Pasture Of Grazing, Dry; Pipelines, Canals; Rancho Estates (Rural Home Sites); Recreational Open (Misc); Rights Of Way, Sewer, Land Fills, etc; Rivers And Lakes; Waste	3,027 (10.4%)
Agriculture	Dry Farms (Misc); Field Crops-Irrigated; Flowers; Irrigated Farms, Misc; Nurseries, Greenhouses; Orchards; Orchards, Irrigated; Pasture-Irrigated; Tree Farms; Truck Crops-Irrigated; Vines And Bush Fruit-Irrigated	2,801 (9.7%)
Industry	Heavy Industry; Industrial Condos (including industrial planned unit developments); Industrial, Misc; Light Manufacturing; Lumber Yards, Mills; Mineral Processing; Open Storage, Bulk Plant; Packing Plants; Warehousing	378 (1.3%)
Utilities	Utility, Water Company; Water Rights, Pumps	388 (1.3%)
Unclassified	Highways And Streets; Poultry	327 (1.1%)
Vacant	Vacant	2,479 (8.6%)
Roads	excluded from parcels	2,839 (9.8%)
Ocean	excluded from parcels	3,480 (12.0%)

Network Centrality Data

Road network geometry influences the spatial distribution of economic activities in urban areas (Ravualaparthy & Goulias, 2014), and we assume it may also affect people's attitudes about a region. To test this, we calculated several network centrality measures for the Santa Barbara area road network, using methods described in Ravulaparthy & Goulias (2014) and averaged the values over each hexagon. Network centrality types considered include:

- *Closeness*, which measures the shortest-path distance between a given link and other links. More central links should generally be closer to other links than should more peripheral links.
- B*etweenness*, which measures the number of shortest paths between other links that pass through a given link. Freeways will often have the highest values for this measure, since they provide direct, rapid movement between different parts of the region.
- *Straightness*, which measures the deviation of links from shortest path distances. Gridded networks tend to have relatively high straightness, which often marks out downtown areas from newer suburbs with curvilinear streets.
- *Reach*, which measures the number of other links that can be reached from a given link by traveling up to a given distance. Places well-served by the network should be able to reach the rest of the network more easily than peripheral areas.

Though network geometry has been shown to affect business success indirectly by making places more or less easily accessible by potential customers, we generally found network centrality to have no significant effect on any of the attitudes variables we measured. The difference between the findings in this paper and Ravulaparthy's work is likely at least in part due to a change in analysis scale – the network controls access to specific businesses, which exist at a single (point) location, but these differences may flatten out across a larger spatial area, diminishing the effects of centrality when compared to other more direct measures of opportunity.

Geotagged Tweets

Surveys are likely the most rigorous way to collect specific information about places, but newer data sources may provide complementary information while incorporating the opinions of far more people at far smaller expense than traditional methods. While Twitter is hardly the only social media platform to collect geotagged information, few can match its sheer volume or the relative ease with which data can be collected from Twitter. Because of this, many researchers have focused on tweets as carriers of social information, particularly to do with happiness. One caveat for this work is that while roughly 500 million tweets are made worldwide per day ("About Twitter, Inc.," 2015), less than 1% are geotagged (Solon, 2014), and not all of these contain much real information beyond the mere fact that the poster is located at a certain place and a certain time (Hiruta, Yonezawa, Jurmu, & Tokuda, 2012).

Twitter is used by roughly 20% of the American internet-using population, with usage concentrated among people under the age of 30, who traditional travel behavior surveys often miss, but Twitter is becoming increasingly popular across all age groups (Duggan et al., 2015). Many tweets are geotagged, which provides a reasonable degree of certainty about a person's location, addressing a problem of measurement error that often plagues crowd-sourced spatial data and Volunteered Geographic Information (VGI) (Goodchild, 2007). Unlike standard VGI, the geographic component of tweets may be incidental: people use twitter to share information, but their location is shared only with their tacit acknowledgement – as little as a one-time choice to share their location. Geotagged tweets fall somewhere between intentionally volunteered geographic information and "coerced" geographic information (McKenzie & Janowicz, 2014), and care must be taken to respect the privacy of Twitter users. Because of the precision and wide availability of Twitter geotags, most applications to transportation research use Twitter as a source of real-time locations to infer information about disruptions of the transportation system (Chan & Schofer, 2014; Pender, Currie, Delbosc, & Shiwakoti, 2014; Ukkusuri, Zhan, Sadri, & Ye, 2014) or study travel through a region (Lee, Gao, & Goulias, 2015).

It is considerably more difficult to determine how much meaningful information can be gleaned from a tweet's text. Humans can easily understand and interpret the text of a few tweets, but the task becomes much more difficult and uncertain when this process must be automated in order to process large samples. An early attempt to automatically classify tweets into five broad categories achieved 82% accuracy compared to human cross-checkers based on a set of broad thematic categories (Hiruta et al., 2012), but it is unclear to what extent this is a product of the specific categories used. Efforts to extract small pieces of information from the text of tweets have been somewhat more successful. Researchers produced a map of happiness in New York City based on the spatial distribution of geotagged tweets containing certain emotionally-coded words, phrases, and emoticons (Bertrand, Bialik, Virdee, Gros, & Bar-Yam, 2013). A coarser-scale US-wide study compared the happiness and expression of different states and cities based on the relative frequencies of a wider range of words in local tweets; this work also made a compelling case for the significance of this data by linking word use on twitter to sociodemographic information measured by other sources (Dodds, Harris, Kloumann, Bliss, & Danforth, 2011; Mitchell, Frank, Harris, Dodds, & Danforth, 2013). The latter studies employed Amazon's Mechanical Turk crowdsourcing service to establish empirical happiness ratings for over 10,000 words and developed a method to convert these happiness scores into an overall happiness rating for a body of text. This word list and the method they used to convert it to a measure of place happiness are reapplied in this thesis.

Social media data sources can provide a stunning range of social and geographic information, but the first step must be to determine how much of this information is actually useful. Humans can easily extract meaning from a few tweets, but the task becomes much more difficult and uncertain when large samples necessitate computers for the task. A US-wide study compared the happiness and expression of different states and cities based on the rates at which local tweets (including place-tagged tweets, which are much more common than geotags but only provide a city, instead of precise geographic coordinates) included specific emotionally significant words; this work also made a compelling case for the significance of this data by linking word use on twitter to sociodemographic information measured by other sources (Mitchell et al., 2013). Other work has analyzed happiness at a finer spatial scale within smaller regions: researchers produced a map of happiness in New York City based on the spatial distribution of geotagged tweets containing certain emotionally-coded words, phrases, and emoticons (Bertrand et al., 2013). The work done by these researchers serves as a starting point for the developing field.

Though they do not represent an explicit measurement of place characteristics in the same way business establishments and land use do, variables produced from tweet locations and text may provide a free, easy-to-collect measure of people's attitudes towards a place. To test this, we continuously collected geotagged tweets in the study area (using a bounding box of 119.5-120 degrees west longitude and 34.3-34.5 degrees north latitude) nearly continuously from November 23, 2014 to April 6, 2015 using the Twitter API and the Python package Tweepy. After the removal of numerous tweets that contained only imprecise "place" tags rather than mobile-device geotags, this process yielded roughly 150,000 tweets. Analysis was initially hampered by the vast disparities in tweet frequency among different users. Though the dataset included tweets from 8,084 unique users, the 171 (2.1%) most frequent tweeters accounted for fully half of all the tweets in the region, and the ten busiest tweeters accounted for one tenth of all tweets. In contrast 3,239 (roughly 40%) users only tweeted once in the region and 6,536 (over 80%) tweeted ten or fewer times. A quantile-quantile plot of tweets by user shows this discrepancy very clearly (Figure 5).The various studies cited in the literature section did not discuss this issue, and it is possible that it is particularly apparent in long-term collections for a relatively small region. Extracting broadly meaningful information from these tweets must address this imbalance, so tweet counts per hexagon were not included as a variable.

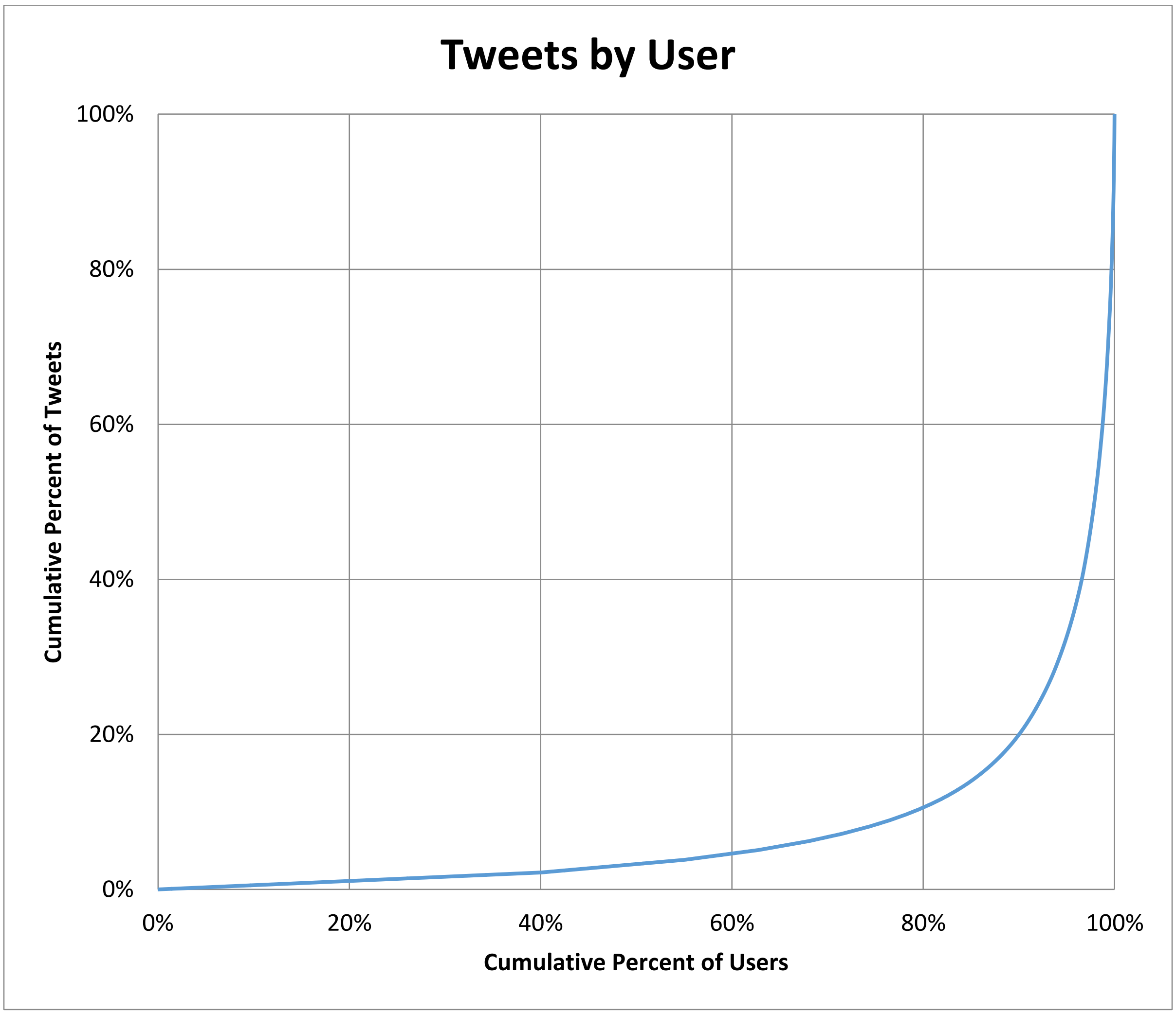

Figure 5 Quantile-Quantile Plot of Number of Tweets by User (Geo-tagged Tweets in the Santa Barbara region, November, 2014 – March, 2015)

Tweet locations provide some information about people and the region, but geotags are largely incidental to the actual purpose of tweeting, which is mainly expressed in the text of tweets. It is currently impossible to automatically extract the full abstract meaning of a block of text, but promising methods based on word frequency have been developed, using spatial word counts. The first step in creating a spatial word count is to extract words from the text of all individual tweets that were posted in the region in question, which we did using a regular expression that split the text into blocks of units of characters containing only letters, numerals, and apostrophes. A simple word count was conducted among all tweets in each GeoTrips hexagon such that all appearances of each word in any tweet in a given hexagon were counted equally. In an attempt to somewhat lessen the impact of the few users who tweeted most frequently and clustered their tweets around a few specific locations (likely home and work), we also computed user-based counts. For these, each word that appeared in the combined text of a given user's tweets made within a given hexagon was counted once, and then total word counts were created by combining the word lists of everyone who tweeted in the region, meaning each word's total count was equal to the number of users who used that word at least once in their tweets in the hexagon. This method does not give all users equal weight, since high-frequency users are likely to provide more unique words from a given hexagon and tweet from more hexagons when compared to lower-frequency users. Both word count processes were repeated for all tweets within the study area (with the user-based one counting each unique word used by each user in all their tweets once).

Once word counts were calculated, we converted them into "hedonometer" happiness scores, using the algorithm and word list provided by Dodds, Mitchell, and their fellow researchers (Dodds et al., 2011; Mitchell et al., 2013). This method calculates the aggregate of happiness of a text by computing the average happiness scores for each word that appears in the text, weighted by its frequency. The averaging process excludes neutral "stop" words that received an average happiness rating of between 4 and 6 from their respondents on Amazon Mechanical Turk, which excludes 6,491 of the 10,222 words they tested.

The two word count methods produced very similar results (correlation coefficient = 0.945 across the 23 hexagons), and region-wide scores of 6.08 for the total word count and 6.04 for the user-based word count. Both of these scores are higher than the reported year 2011 nationwide average of 6.01 but lower than the reported value for Santa Barbara of approximately 6.145, which made it the 14th happiest city of the 190 for which Mitchell calculated a score (Mitchell et al., 2013). Calculating the average happiness for hexagons 5-14, which most closely correspond to the borders of Santa Barbara resulted in slightly higher scores than what was reported in the literature. Among the possible reasons for this discrepancy are changes in Twitter's user base between 2011 and late 2014, differences in duration of sample, and the fact that their sample appears to have included place-tagged tweets in addition to tweets with precise geotags.

Much as Mitchell found a negative correlation between the number of tweets per capita in a city and its happiness score, we find a negative relationship between the number of tweets posted in a hexagon and their overall happiness score, regardless of which word count method is used (correlation coefficients were -0.42 for overall word counts and -0.48 for user-based word counts). The largest number of tweets originated from the student neighborhood of Isla Vista (hexagon 20, generally), adjacent to UC Santa Barbara, so part of the apparent relationship may stem from younger people generally posting more negative tweets. Figure 6 shows the general distribution of tweets throughout the region and the average happiness scores of those tweets. Tweets from the eastern half of the region are substantially happier than those from the eastern half overall. One possible explanation for the negative relationship between tweet frequency and tweet happiness is that frequent tweeters (including students) may use the social media site to share information about all aspects of their lives, whereas less-frequent users may be more inclined to share happy moments.

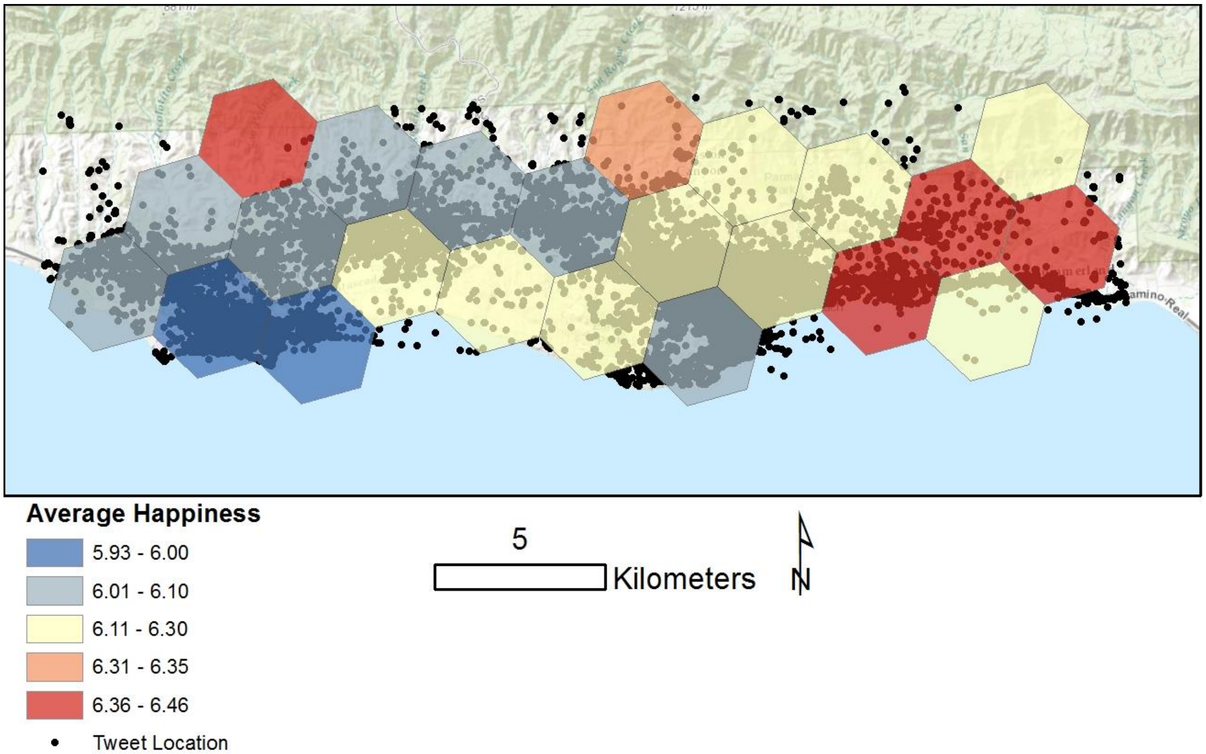

Figure 6 Tweet Happiness and Tweet Locations

A complete list of all variables used in a model presented in this thesis is provided in Table 5.

Table 5 Description of Model Variables

Name	Source	Description	Range (Mean, SD)
			Response Variables
Attractiveness	GeoTrips	This is an attractive area.	7-choice Likert
Familiarity	GeoTrips	I am familiar with this area.	7-choice Likert
Opportunity	GeoTrips	This area provides opportunities to do things I like to do.	7-choice Likert
			Hexagon-Person Pairing
Danger	GeoTrips	This is a dangerous area of Santa Barbara.	1-7 (2.8, 1.6) – Likert treated as numeric
Local Att.	GeoTrips	Respondent's mean Attractiveness rating for hexagon's neighbors.	1-7 (5.4, 1.2) – From Likert as numeric
Local Fam.	GeoTrips	Respondent's mean Familiarity rating for hexagon's neighbors.	1-7 (5.1, 1.4) – From Likert as numeric
Local Opp.	GeoTrips	Respondent's mean Opportunity rating for hexagon's neighbors.	1-7 (4.7, 1.3) – From Likert as numeric
Local Dan.	GeoTrips	Respondent's mean Danger rating for hexagon's neighbors.	1-7 (2.9, 1.3) – From Likert as numeric
			Between Hexagons
Shoreline (km)	ArcGIS online data	Length of shore in hex (km). Likely an imperfect proxy for appreciation for the ocean; performed better than ocean area in hex.	0-5.0 (1.5, 1.8)
% Open Space	SB Co Parcels	Percent of hexagon land area designated for open space.	1.9-49.6 (11.5, 12.0)
% Roads	SB Co Parcels	Percent of hexagon land area covered by roads.	0.5-22.7 (11.7, 6.4)
% Housing	SB Co Parcels	Percent of hexagon land area used for single- or multifamily housing.	2.6-72.4 (37.7, 19.4)
Consumer Ests	NETS	Number of customer-serving businesses (/100). Covers retail (SIC classes 53, 54, 56, 57, and 59), eating and drinking (58), and entertainment (78, 79, and 84).	0-5.9 (0.8, 1.4)
Shannon	NETS	Shannon Entropy of business establishments in hexagon, based on level-2 SIC code.	1-3.5 (2.5, 0.9)
Tweet Happy	Twitter Sample	Average "hedonometer" score from user-based word counts of tweets in hexagon, method from (Dodds et al., 2011; Mitchell et al., 2013).	5.93-6.46 (2.52, 0.91)
			Between People
Female	GeoTrips	Dummy variable for gender = female	57.6% of respondents
Age 18-25	GeoTrips	Dummy variable for age in [18,25]	10.0%
Has Car	GeoTrips	Dummy variable for car ownership	97.7%

Chapter 4 – Methods

The goal of this study is to understand how measurable attributes of place relate to reported place attitudes variables recorded in the GeoTrips survey. The peculiarities of our dataset require a specific structure that will allow us to model the relationships among variables appropriately and account for some potential sources of endogeneity. The model we present can be characterized as follows:

- Ordinal regression makes it possible to model ordinal response data appropriately.
- Cross-classified multilevel modeling is required because the data contains repeated observations grouped along two separate axes: respondents and spatial units.
- Structural Equations Modeling makes it possible to understand the relationship among multiple response variables simultaneously, rather than focusing on one variable at a time.

The following section will discuss aspects of the model structure in detail and then address the model's shortcomings.

Methods Used

Ordinal Choice Modeling

Linear models are straightforward to estimate, but they are designed for continuous response variables. Because all of the place attitudes are measured on an ordinal scale (we know that "Strongly Agree" is higher than "Agree," but we don't know how much higher), we must choose a slightly different model. Ordinal dependent variable models convert between discrete ordinal survey responses and a continuous index function that is assumed to be the basis of the variables (when asked a question with an ordered response, people select the rank that is closest to what they actually feel). The unobserved true value is linked to the observed response using a probit function with each outcome classified into an ordinal rank. The modeling process involves estimating the set of thresholds that separates the ranges for each level of the order.

Place attitude responses can be thought of as ratings of a place, and so the continuous index function that the model uses is analogous to utility as it is used in choice modeling (Greene & Hensher, 2010, Chapters 1–3). Instead of predicting a respondent's first choice from a range of distinct options, ordered choice models predict the respondent's rating for a particular option, relative to other options. In a discrete choice model, a separate utility is calculated for each option, but in ordered choice models, each object being rated has a single utility that is classified into an ordinal rating. Green and Hensher outline multiple applications of this model structure, using Netflix movie ratings and other similar ratings scales. While the variables of interest are not especially geographic, their structure is nearly identical to the place attitudes variables used in this study. Ordinal movie ratings reflect an individual's enjoyment of a particular movie, and ordinal place ratings reflect an individual's enjoyment of a particular place (with respect to its attractiveness or opportunities). In addition, both datasets present repeated measures both of ratings subjects (reviewers/respondents) and objects (movies/hexagons). Rating responses do not have quite the same meaning as yes-no choices in discrete choice models because they do not produce a single top choice, but they do provide more information about the range of a person's feelings. A discrete choice model based on home location could also be developed from this dataset, which would enable us to understand the relationship between the various axes of place attitudes and a measurable choice.

Cross-Classified Multilevel Modeling

It would be attractive to assume that because the dataset contains 12,903 records for each of our place attitudes variables, it has that many independent observations; unfortunately this is not the case. In any study that records multiple values for a given subject, there is likely to be a degree of consistency among that subject's responses, which breaks the assumption that each observation is independent and exogenous. Multilevel modeling is the conventional way to address this source of endogeneity.

In our survey, all of the responses from an individual are likely to be correlated to both their overall interpretation of the hexagon attitudes questions and the measurement scale as well their own overall sense of the region. Likewise, different people are likely to rate a specific hexagon similarly because some inherent characteristics of the region are likely to weigh into their response, even though each person has a different specific relationship with it. Given the structure of this survey, this between-level variability can take two forms:

1) Varied typical response. Some people give more high scores while others give more low scores, and some hexagons receive consistently high or low scores in a given category. Multilevel models are designed to address this issue, and a cross-classified model can address the effect from both sources.
2) Wider or narrower interpretation of each category. Some people assign at least one hexagon to every category while others clump their responses more tightly regardless of whether these reflect actual differences of opinion. Some work has been done on modeling ordinal data with heterogeneous response thresholds (Greene & Hensher, 2010, Chapter 7; Johnson, 2003), but this thesis will not address that issue.

Since GeoTrips asked each respondent a series of similar questions, it is likely that their experiences and biases shape their response in a consistent way. I see two main sources of person-level variability:

a) Each person has a unique perspective of the region. A respondent who has lived in Santa Barbara all their life will likely (and accurately) report a very high degree of familiarity with the whole region, even with the hexagons that a newer arrival would have never visited.
b) All survey instruments are imperfect, and different people may interpret the questions somewhat differently. If one person says they "agree" that a hexagon presents many opportunities and another says that they "neither agree nor disagree" about the same question for the same hexagon, this may

still represent the same fundamental opinion of the hexagon expressed in two different ways.

Any sort of multilevel modeling can address both of these effects (as long as they affect the basic response level, not the response range, as noted above), but a multilevel model that also attempts to determine the sources of the person-to-person variability can provide deeper insight into the first question. For instance, young people (including many college students who grew up outside the region) and people who did not own a car reported universally lower familiarity with the region, which undoubtedly reflects a true difference between them and other residents of the area. While differences in survey interpretation may vary systematically with some unknown variable (e.g. pessimism vs optimism), it should be safe to treat this as a random attribute of each individual, which multilevel models account for; thus, I interpret all significant person-level variables as bearing on their true understanding of the region.

Much as this survey entails repeated measures at the level of an individual respondent, each hexagon is rated by each of the 561 respondents. A multilevel model that only accounted for person-to-person variability would assume that each person was asked about a totally unique set of hexagons. Because each respondent in our survey rated the same set of 23 hexagons, these responses are not independent. One solution is to estimate a cross-classified multilevel model in which each response draws an effect from both the individual respondent who supplied it and the hexagon in question.

As with person-person variability, each hexagon's mean response is modeled as a function of its characteristics. Hopefully, these characteristics approximate the "reality" that forms the basis of people's opinions of a place. Understanding the relationship between these hexagon-level variables and the target place attitudes variables is the primary goal of this study. Most of the external variables in this study are intended to address hexagon-level variability.

Multilevel models are commonly used to address groupings in data, but this dataset contains groupings along two axes (each respondent and each hexagon have multiple responses), which necessitates a cross-classified model. This is a relatively rare case, and little work has addressed it. Though choice experiments can be interpreted as being cross-classified (as long as multiple respondents are offered any of the same options), but choice modelers are generally more interested in how individuals weight the different attributes of their options than in anything essential about a specific option, so the structure is somewhat different. Bhat presents one use of a cross-classified scheme to account for the effects of respondents' home and work locations within a discrete choice model for travel mode, grouping responses by home location and by work location (Bhat, 2000). The need for cross-classified modeling arises more often in Education research (for instance, each student attends a school and lives in a neighborhood, and these groupings may not match up consistently). Fielding and Goldstein explain the structure's necessity and describes several use cases (Fielding & Goldstein, 2006); the authors of Mplus use this paper as the basis for its implementation of cross-classified models (Asparouhov & Muthen, 2014).

Multilevel designs barely impact the estimated coefficients, since they are roughly equivalent to including each grouping as a dummy variable, but models that do not account for true groupings present in the data will underestimate coefficient standard errors, leading to unreasonable claims of variable significance (Fielding & Goldstein, 2006). In their simplest form, multilevel and cross-classified models provide a unique intercept for each group in each grouping scheme, but they can also be used to model the relationship between group-level variables and group means. This allows a model to analyze relationships operating across different levels and between different groupings while still accounting for groupings in the data (Fielding & Goldstein, 2006).

GeoTrips respondents and hexagons are perfectly cross-classified: each possible person-hexagon paring occurs exactly once in our dataset. This means that hexagon mean responses cannot possibly be correlated with individual-level variables (and vice versa), since any relationship would balance out across all hexagons, enabling us to segregate respondent-level variance from hexagon-level variance. This independence does not extend to all aspects of the respondent-hexagon relationship because the survey did not stratify by home location. Some hexagons are home to large numbers of respondents, and some are home to none; because individuals generally rate their home hexagons higher in all respects, hexagons with disproportionately many residents in our sample likely received higher average scores than they would have if the entire population of Santa Barbara responded to the survey. One potential solution would be to exclude each individual's home-hexagon response from the dataset (although the effect is also present for hexagons adjacent to the home hexagon).

The Results section will directly compare the results of a full cross-classified (person-response and hexagon-response) model with the simpler multilevel (person-response) design to show their similarities.

Structural Equations Modeling

The GeoTrips survey asked respondents about each hexagon's level of Opportunities, Attractiveness, and Danger, as well as their Familiarity with it. Though these variables could be modeled separately, they are all aspects of an individual's overall feelings about an area that likely impact decisions about what to do and where to go. Because we are interested in the interaction among these variables, this model will be a relatively simple Structural Equations Model (SEM), with multiple dependent variables but no latent variables.

SEM is particularly valuable for studies of human behavior and attitudes because these cannot often be summarized in a satisfactory way by a single variable. Human processes are complex and multifaceted, and our modeling should reflect that by analyzing the complex relationships among multiple related variables, which is a particularly common problem in transportation research (Kuppam & Pendyala, 2001; Farag, Schwanen, Dijst, & Faber, 2007). Because there is no single variable that can completely measure place, this is an obvious candidate for SEM. Applications to travel behavior research have included attempts to differentiate the effects of neighborhood characteristics on travel behavior from the effects of neighborhood (Cao, Mokhtarian,

& Handy, 2007), and Deutsch's investigation of the relationship between multiple demographic variables and various measures of sense of place (Deutsch, Yoon, et al., 2013).

The output of these models can show us what measurable attributes of a place are reflected by people's attitudes and which aspects of attitudes are related, but it cannot directly address cause. SEM requires the modeler to propose a set of unidirectional relationships among the endogenous variables, but without collecting data at multiple time points to measure change, this sort of model cannot justify claims about cause. Though there are likely mutually causal relationships among the place attitudes – e.g. people visit opportunity-rich places more frequently, and as they become more familiar with a place, they become aware of more of the opportunities it presents – but for the purposes of this model, the variables were arranged in a hierarchy with each relationship modeled as unidirectional (though the estimated size and sign of these relationships does not change significantly if the direction is reversed). In this case, we chose to regress perceived opportunity on both familiarity and attractiveness, and attractiveness on familiarity. Models that included all four variables failed to converge, and danger seems to vary much more idiosyncratically from person to person, so it was chosen as the variable to exclude.

Data and Model Structure

The models presented here are designed to determine which measurable characteristics of a place have a stronger relationship with people's attitudes about it. Since there are multiple separate place attitudes that are important, we use a structural equations model to understand the relationship among these variables and between place attitudes and the measurable "reality" of the places we asked people to rate. Because our measures of place attitudes are recorded on an ordinal scale, we used a probit link function to convert between people's ordered responses and a continuous index function that theoretically reflects their true opinion. The task is further complicated by two sources of non-independence in our data, which we address using a cross-classified multilevel design. A cross-classified design allows us to separate variance at the level of the individual respondent, at the level of an individual hexagon, and for a given person-hexagon pairing. The overall structure of the model is shown in Figure 7.

Once all datasets were processed, they were arranged in a tabular format, with one entry per person per hexagon, for a total of 561 x 23 = 12,903 entries total. Simple two-level and cross-classified effects were also estimated in R to supplement the general results (the results of these are shown in the next section), but the more complex Cross-Classified SEM models were estimated in Mplus.

If the three place attitudes variables were modeled independently rather than as part of an SEM, the cross-classified ordinal model for each would take the form shown below. Equations were adapted from multiple sources identified in the Mplus documentation (Fielding & Goldstein, 2006; Asparouhov & Muthen, 2014).

The structure of a simple multilevel regression model is shown in Equation 1. For each observation i for person j, the reported place attitude is Y_{ij}. Because this variable is ordinal, we also use a continuous index latent variable Y_{ij}^{*} that can be interpreted as the propensity for responding in each category of the order. $Y^*_{1,ij}$ refers to the observation-specific component of the variability (essentially a random variable, once variability between individuals is accounted for) and $Y^*_{2,j}$ refers to the persistent effect (essentially a random intercept) of the respondent across all hexagons for a given attitude variable.

$X_{1,ij}$ contains the vector of predictor variables at the observation level – attributes specific to a given person-hexagon pairing (such as spatially lagged attitudes variables, and potentially home location) as well as the attributes of the hexagon to which this observation corresponds – and β_1 contains the relationship between these variables and the response-level variability of the given place attitude. The random error associated with a specific observation is $\varepsilon_{1,ij}$. The second level of the model estimates the persistent effect of the individual's response style or understanding of the region and takes a generally similar form to the first level. Individual characteristics to be used as predictor variables (e.g. gender and age) are stored in $x_{2,j}$; the coefficients of the individual intercept on these variables are stored in β_2; and a component of random error is represented by $\varepsilon_{2,j}$. One feature unique to the second level of the model is ν_2, the intercept for individual-level model.

$$Y^*_{ij} = Y^*_{1,ij} + Y^*_{2,j}$$
$$Y^*_{1,ij} = \beta_1 x_{1,ij} + \varepsilon_{1,ij}$$
$$Y^*_{2,j} = \nu_2 + \beta_2 x_{2,j} + \varepsilon_{2,j}$$

Equation 2 Multilevel Model Structure

To adjust the model so that it can account for cross-classification in the data (for individual j and hexagon k), a third level is added to the model and the other two levels are adjusted slightly (Equation 2). Note that each hexagon-person pair will have exactly one observation i, but this does not affect the model. The overall model is now split into three components, with $Y^*_{2,j}$ and $Y^*_{3,k}$ containing random intercepts for the individual and hexagon, respectively, and $Y^*_{1,ijk}$ containing the residual portion of the observation not explained by the person or hexagon intercept. Since the model can contain only a single overall intercept, it is moved to the observation level as ν_1. The only other difference in the response level of the model is that $x_{1,ijk}$ now appropriately contains only explanatory variables specific to the hexagon-individual pairing, with hexagon attributes moved to the hexagon level of the model as $x_{3,k}$. Aside from the removal of the overall intercept, the individual level of the model ($Y^*_{2,j}$) remains unchanged and the hexagon level ($Y^*_{3,k}$) works in an identical fashion to the individual level.

$$\begin{aligned} Y^*_{ijk} &= Y^*_{1,ijk} + Y^*_{2,j} + Y^*_{3,k} \\ Y^*_{1,ijk} &= \nu_1 + \beta_1 x_{1,ijk} + \varepsilon_{1,ij} \\ Y^*_{2,j} &= \beta_2 x_{2,j} + \varepsilon_{2,j} \\ Y^*_{3,k} &= \beta_3 x_{3,k} + \varepsilon_{3,k} \end{aligned}$$

Equation 3 Cross-Classified Model Structure

Both the multilevel and cross-classified models use a probit function to link the index variables for place attitudes Y^*_{ijk} to the observed ordinal rankings Y_{ijk} (Equation 3). To achieve this, the model estimates a set of cut points μ_c to classify the continuous index function into ordinal responses

$$\begin{aligned} Y_{ijk} &= Strongly\ Disagree\quad if\ Y^*_{ijk} \le \mu_1, \\ &= Disagree\ if\ \mu_1 < Y^*_{ijk} \le \mu_2, \\ &= Slightly\ Disagree\ if\ \mu_2 < Y^*_{ijk} \le \mu_3, \\ &= Neutral\ if\ \mu_3 < Y^*_{ijk} \le \mu_4, \\ &= Slightly\ Agree\ if\ \mu_4 < Y^*_{ijk} \le \mu_5, \\ &= Agree\ if\ \mu_5 < Y^*_{ijk} \le \mu_6, \end{aligned}$$

$= Strongly\ Agree\ if\ \mu_6 < Y^*_{ijk}.$

Equation 4 Ordinal Classification of Index Variable

A particular wrinkle caused by the cross-classified model structure is that an estimation method must be chosen more carefully (Fielding & Goldstein, 2006). Maximum likelihood estimation is the typical means of fitting structural equations models with ordinal variables, and this method can handle multilevel models with some random effects. Though there is no theoretical reason it should not be used with cross-classified multilevel models, maximum likelihood estimation often becomes infeasible in models containing more than three or four random effects, in which case it often returns non-positive definite variance covariance matrices or negative residual variances (Asparouhov & Muthen, 2014). To avoid these problems, Mplus exclusively estimates cross-classified multilevel models using Bayesian estimation, which is more stable and can reliably handle models with large numbers of groups (23 hexagons and 560 respondents, in this case) with any number of random effects (Asparouhov & Muthen, 2014). The major downside of this estimation method is that Mplus has not yet implemented many fit statistics, which makes it difficult to compare the results of multiple models. In this case, a final model was chosen based on the fit diagnostics of its simple multilevel equivalent and the reasonableness of its coefficients.

Because this is a structural model, the relationship among the three endogenous variables has a major bearing on the interpretation of the relationships between place attitudes and place attributes. The relationship between each pair of variables is equal to the direct effect plus any indirect effects. Indirect effects are the product of the direct effect of one variable on an endogenous variable and the relationship between that variable and the endogenous variable in question. For instance, the total effect of shoreline on perceived opportunities is equal to the direct effect of shoreline on perceived opportunities plus the product of effect of shoreline on attractiveness and the effect of attractiveness on opportunities.

Structural equations provide estimates of regression coefficients that are named direct effects to distinguish from the influence an exogenous variable via a mediating third variable. For instance, Tables 15 and 16 show that Danger has a direct (and negative) effect on perceived Opportunity and on Attractiveness. Because Attractiveness also has a direct effect on Opportunity, the total effect of Danger on Opportunity is the combination of its direct effect and an indirect effect mediated through Attractiveness.

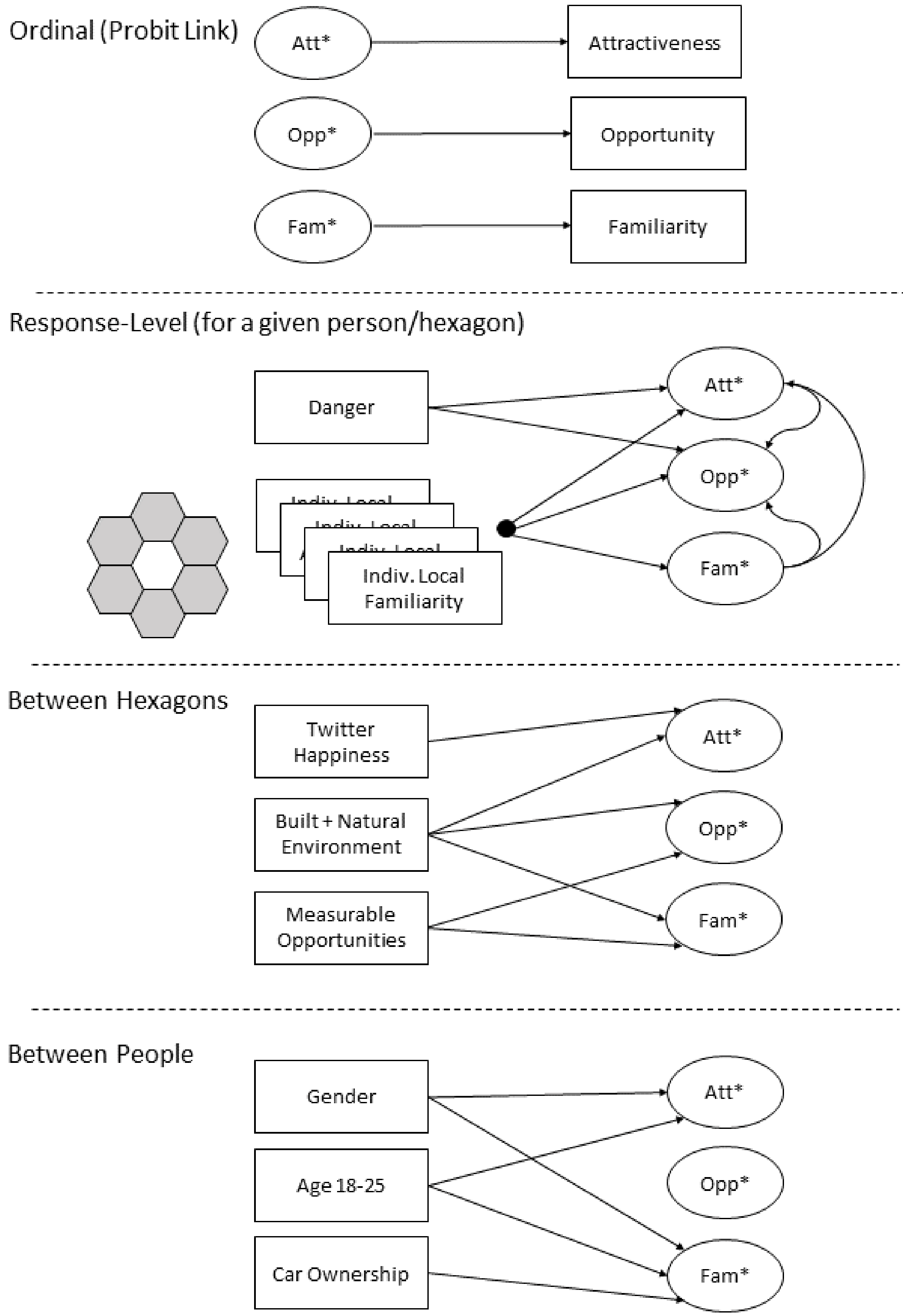

Figure 7 Cross-Classified Multilevel SEM Ordinal Probit Model Structure

Methods – Discussion and Critique

No model is perfect, and it is important to understand the shortcomings of a specific model in order to recognize the limitations of its results. In some cases the model presented in this thesis was limited by the data source and in some cases by difficulties in estimation given available software and expertise. Though this model is clearly imperfect, it can provide a foothold for quantitative place research.

Is a Cross-Classified Structure Necessary?

Given the two types of groupings in our data, it is clear that a cross-classified multilevel model is the most appropriate design. However, this design adds substantially to the model's complexity and increases the time to estimate by more than tenfold. Additionally, Mplus does not provide fit indices for cross-classified models, which limits our ability to fully understand our results. Because of this, a "good-enough" solution may be to dispense with the cross-classified structure of the model if one of the two grouping systems has a substantially smaller effect than the other, especially since the cross-classified design has little effect on estimated coefficients of most variables.

One way to test the differences between modeling strategies is by running a series of simple models for each response variable in question: one simple 2-level model with respondent-level random intercepts, one with hexagon-level random intercepts, and a cross-classified model with random intercepts for both levels. Results of these tests are shown in. One way to assess the relative importance of different grouping schemes is to compare the variances of each group's intercepts in the cross-classified models (Fielding & Goldstein, 2006); because the response variables are ordinal and may be scaled differently between models, it is particularly important to compare the two groupings directly in a single model. In each case, there appears to be more variability among respondents than among hexagons, which indicates that opinions about the region vary more significantly from person to person than over space.

Opportunity and attractiveness show the least difference between hexagon-level and respondent-level variances, which indicates that these responses may be somewhat less subjective or personal than the other variables. Individual respondents have widely varied opinions about the opportunities available in the region, but there is a general consensus that the main commercial areas in Santa Barbara provide more than other hexagons. The differences between hexagon-level and person-level variability are slightly larger for familiarity, since different people likely have substantially different levels of familiarity with the region as a whole and with specific hexagons. Danger, which ostensibly represents a quantifiable hexagon attribute, shows the widest disparity between person-level and hexagon-level variance, suggesting that the overall sense of danger is largely personal and no single area of Santa Barbara is universally agreed to be more dangerous than the rest.

In addition to comparing the intercept variance, we can also compare the models' fit indices, to determine whether it is worthwhile to incorporate the added structure. Models like this can be compared directly using their log likelihoods. Multilevel structures are an all-or-nothing proposition, since each response belongs to at least one grouping in each level. Assuming a multilevel structure is theoretically justified, then including dummy variables only for the groups most different from the rest of the data would not be an acceptable solution.

In each case, the switch from a simple multilevel model to a simple cross-classified model results in a substantial likelihood ratio improvement. Even considering the large number of degrees of freedom exhausted in the switch from a hexagon-based multilevel design to a cross-classified model (561, one per respondent), all of the cross-classified models represented an extremely significant improvement. From this, we can conclude that a cross-classified design clearly improves the model functionally, in addition to being theoretically correct.

We estimated full models using both a 2-level design (grouped by person) and a cross-classified one, as discussed in the results section. While coefficients were consistent between 2-level and cross-classified models, the two-level model vastly underrepresented the standard errors of hexagon-specific coefficients, which were modeled as between-level variables in this case. Because the two-level model implies all the observations are independent once individual respondents are taken into account, it treats hexagon-level variables as independent, so hexagon attributes are likely to be declared highly significant, since they will be assumed to have several thousand degrees of freedom, instead of the 23 they actually have.

Table 6 Simple Cross-Classified vs 2-Level Model Results

Variable	*Level*	*Log Likelihood (2-Level)*	*Log Likelihood (Cross)*	*Intercept Variance (2-Level)*	*Intercept Variance (Cross)*	*Likelihood Ratio*
Att	Hexes	-19957	-18310	0.1880	0.2890	3295
	Resps	-19648		0.4426	0.6014	2676
Opp	Hexes	-22109	-20754	0.1466	0.2078	2711
	Resps	-21761		0.3593	0.4474	2015
Fam	Hexes	-22143	-19989	0.2044	0.3484	4307
	Resps	-21549		0.5916	0.8373	3120
Dan	Hexes	-20650	-16903	0.1178	0.2697	7514
	Resps	-18094		1.3050	1.7194	2383

Spatial and Social Endogeneity

Because the survey was not designed to sample social networks, it is safe to assume that individuals in our sample likely do not know each other and can be modeled independently. Though excluding social ties eliminates one major source of potential error from the model, a complete model would include these effects. Travel and the needs it serves are very often socially driven (Deutsch & Goulias, 2013), so each individual's opinion of the region is shaped by the people they know. Whether it is a matter of a friend taking you to their favorite taco shop in a part of town you had never thought to explore, or a couple going on a date to a secluded beach, people experience an area in large part through other people. Greater understanding of the interaction among social networks, travel, and spatial/placial knowledge is needed.

Respondents can be safely assumed to be independent (at least at the scale of their relationship with the Santa Barbara area), but the survey hexagons share a spatial relationship, and are not independent. Spatial autocorrelation is likely to make many measureable attributes of nearby hexagons similar, but it will also make neighboring hexagons similar in unmeasured ways, likely resulting in correlated error terms for any model estimated from the data.

The spatial relationship among hexagon means can partly be corrected by including appropriate exogenous variables that explain some of this spatial variation, but some of the effect is likely to be due to unmeasured or unmeasurable similarities between nearby hexagons. It may be best to address this by adding an autoregressive term to the model. Bhat's work on spatially correlated logit choice models is an especially relevant example of this sort of model: a simple spatially autoregressive model considers the first-order neighbors of each hexagon (Bhat & Guo, 2004) and a more complex one also attempts to account for the decay of relationships over space (Paleti et al., 2013). Bhat applies these to a discrete choice housing model, but the application would be similar in a cross-classified Structural Equations Models. Unfortunately, this type of spatially correlated models cannot be estimated in Mplus.

A less geographic way to address spatial correlation in our data would be to make each response dependent both on the characteristics of that hexagon and on its neighbors, using a multiple membership model (Fielding & Goldstein, 2006, p. 33). In addition to having an intercept for its specific hexagon, each response would also see an effect (presumably smaller in magnitude) of its neighboring hexagons. However, given the large number of hexagons and hierarchy of relationships present (e.g. adjacency, second-order adjacency, etc.) that would need to be specified individually, it may not be possible or practical to fully implement this sort of model.

Much as there is spatial autocorrelation in terms of the aggregate character of each hexagon, there is likely also a spatial relationship among the ratings of multiple hexagons by any individual respondent. An individual who is particularly familiar with a given hexagon is likely to be familiar with its neighbors too, at least because of the spatial necessity of traveling through one of them to get anywhere else. As an attempt to account for this, we have included spatially-lagged "personal neighborhood average" variables in our model. However, these variables are problematic for a number of reasons: they are numerical means of ordinal data and they reflect both spatial autocorrelation of hexagon mean values and individual responses.

Efforts to Measure Spatial Autocorrelation

Spatial dependency is a relatively well-understood process when applied to phenomena measured numerically, but there are several features of our data that make spatial dependency difficult to measure. The GeoTrips dataset contains multiple ordinal observations in each hexagon, and these are also grouped by individual respondent. As noted above, this likely means there is spatial dependence operating on two different levels: hexagon means are related to the means of nearby hexagons and individual responses about nearby hexagons are also related. Both expressions of spatial dependency are made difficult to measure by the ordinal data structure. This section will explore methods to measure multilevel spatial autocorrelation in an ordinal dataset.

It is unclear whether any published work has ever addressed the issues examined in this section before, but a time series of weather station data could represent a potentially analogous case: the mean readings from nearby stations (climate) would likely be more closely related than those from distant stations, and this would be true for individual readings (weather) as well. In either dataset, it should be possible to investigate the spatial autocorrelation of response means separately from the spatial autocorrelation of individual responses. In the case with weather stations, a multilevel time series model would be one solution; in our case, a relatively straightforward way to do this would be to construct a simple cross-classified multilevel model that groups responses by individual and hexagon, as was done in the first part of this section. Hexagon-level intercepts would be used to investigate overall spatial autocorrelation, and model residuals would form the basis for studying individual-response-level spatial autocorrelation. Because hexagon means can be expressed in relation to the continuous index function, this portion of the analysis would be possible in an ordinal dataset if either Mplus or R made it possible to extract hexagon intercepts from cross-classified ordinal models. Since there is no way to generate a single residual for ordinal data, respondent-level spatial autocorrelation would be more difficult to address.

There is no clear measure of spatial autocorrelation that can be applied to ordinal data, so traditional measures of spatial autocorrelation (e.g. Moran's I and LISA) could not be applied to individual responses in this dataset without falsely treating them as continuous. As noted in the section on GeoTrips Data, the place attitudes variables are ordinal and distributed asymmetrically, which means they cannot accurately be treated as continuous variables.

The repeated measures component of this study provides one potential avenue for investigating spatial relationships. Because each hexagon was rated by all respondents, it is possible to calculate the correlation between all respondents' ratings of any given pair of hexagons. When plotted against distance between hexagon centroids, this scatterplot resembles a correlogram, but includes multiple correlations at each distance, instead of an overall measurement of the entire dataset's spatial autocorrelation at each distance band.

One tool for comparing paired ordinal datasets is the Spearman Rank Correlation Coefficient, which compares whether two variables with paired observations share a monotonic relationship (if an increase in one variable always corresponds to an increase in the other, regardless of the scale of that increase, the two variables have a perfect positive monotonic relationship). A value of -1 corresponds to perfect negative relationship, while +1 corresponds to a perfect positive relationship. The first step is to convert each variable to a partial ranking – the lowest value is ranked 1, the second-lowest is ranked 2, and ties are resolved by assigning each response the average ranking of all tied values. The ranking process is demonstrated for a sample of 10 responses in Table 7.

Once both variables are ranked, the difference in the rankings of each observation between the two datasets is used to calculate the correlation between variables, as shown in Equation 4. If observation i is ranked 10^{th} for variable 1 and 2^{nd} for variable 2, then d_i^2 would be equal to $(10-2)^2=8^2=64$. In this case, the resulting value will reflect the degree to which people who rank one hexagon highly also rank the other hexagon highly, and vice versa.

Equation 5 Spearman's Rho Rank Correlation Coefficient

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

Table 7 Partial Ranking for 10 Ordinal Responses

Respondent	**Response**	**Position in Ascending Order**	**Rank Assigned**
A	Strongly Agree	10	(10+9)/2 = 9.5
B	Strongly Agree	9	(10+9)/2 = 9.5
C	Agree	8	8.0
D	Neutral	7	(6+7)/2 = 6.5
E	Neutral	6	(6+7)/2 = 6.5
F	Somewhat Disagree	5	(3+4+5)/3 = 4.0
G	Somewhat Disagree	4	(3+4+5)/3 = 4.0
H	Somewhat Disagree	3	(3+4+5)/3 = 4.0
I	Disagree	2	2.0
J	Strongly Disagree	1	1.0

The key problem with applying Spearman correlations to our hexagons is that respondents appear to have used different measurement scales from each other. Some people rate most hexagons highly and some rate most hexagons lowly, possibly reflecting their overall opinion of the SB region; when all responses for a given hexagon are ranked, respondents who generally give higher responses will wind up with higher ranks in nearly all hexagons. The small variation of each individual's rankings across multiple hexagons will lead to positive correlations between most pairs of hexagons. A measure of spatial autocorrelation should not capture non-spatial, interpersonal variability.

I will demonstrate this issue and two proposed fixes using a subset of the dataset, shown in Table 8; the effects are more pronounced with larger numbers of hexagons. Hexagons 7 and 8 are adjacent, cover part of Downtown and have the highest average Opportunity ratings of all hexagons. Hexagons 18 and 19 are adjacent low-opportunity hexagons in Goleta, and hexagon 13 is located between the other pairs and covers the opportunity-rich Upper State part of Santa Barbara. Respondents were chosen to provide a mix of homogeneously high (C), mixed high (A and D), and mixed responses (B and E).

Table 8 Subset of Opportunity Ratings for Five Hexagons from Five Respondents

Person	Hexagon 7	Hexagon 8	Hexagon 13	Hexagon 18	Hexagon 19
A	Strongly Agree	Strongly Agree	Strongly Agree	Slightly Agree	Slightly Agree
B	Agree	Strongly Agree	Agree	Strongly Disagree	Disagree
C	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree
D	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Disagree	Strongly Disagree
E	Slightly Disagree	Agree	Slightly Agree	Strongly Agree	Agree

Table 9 Ranking from Untransformed Responses

Person	Hexagon 7	Hexagon 8	Hexagon 13	Hexagon 18	Hexagon 19
A	2	2.5	2	3	3
B	4	2.5	4	4.5	4
C	2	2.5	2	1.5	1
D	2	2.5	2	4.5	5
E	5	5	5	1.5	2

When raw ratings are converted to rankings of responses for a given hexagon (Table 9), respondents with generally higher ratings of all hexagons (A and C in this case) are ranked higher in nearly every hexagon. In this case, respondent C's opportunity ratings are tied for the top position in all five hexagons despite showing no variability in his rankings; his responses thus have bearing on the correlations among this set of hexagons. Some share of the correlation measured between hexagons will thus reflect the consistency of some respondents' ratings rather than spatial autocorrelation.

To address the differences between individual means without losing the data's ordinal character, each individual's median response for a given variable can be mapped to 0, with their other responses adjusted up or down accordingly. The result of the median transformation is shown in Table 10.

Table 10 Median-Transformed Ratings

Person (Median Resp.)	Hexagon 7	Hexagon 8	Hexagon 13	Hexagon 18	Hexagon 19
A (Strongly Agree)	0	0	0	-2	-2
B (Agree)	0	1	0	-5	-4
C (Strongly Agree)	0	0	0	0	0
D (Strongly Agree)	0	0	0	-6	-6
E (Agree)	-3	0	-1	1	0

Table 11 Ranking from Median-Transformed Ratings

Person	Hexagon 7	Hexagon 8	Hexagon 13	Hexagon 18	Hexagon 19
A	2.5	3.5	2.5	3	3
B	2.5	1	2.5	4	4
C	2.5	3.5	2.5	2	1.5
D	2.5	3.5	2.5	5	5
E	5	3.5	5	1	1.5

When median-transformed ratings are used as the basis for rankings of responses for a given hexagon (Table 11), the rankings are less affected by individual respondents' overall rankings. In the simple ratings, all people except for E gave hexagon 8 the highest possible rating, but in all but one case, this was also their median response. In contrast, person C gave only one Strongly Agree and had a median response of Agree, meaning their relative median-transformed rating is highest. Though this method partly eliminates the differences in people's mean response, it does not address differences between individuals' variability of response because not everyone uses the entire range of potential responses, from Strongly Agree to Strongly Disagree. In this set of hexagons, there is no clear difference over space between the responses given by respondents A and D. Both rate the first three hexagons equal and high and the last two equal and low; the difference between the Slightly Agree ratings given by respondent A and the Strongly Disagree ratings from D may reflect their different opinions of the region as a whole or different understandings of the rating system.

A more satisfactory method would account for both differences in average response and differences in response variability (which would be analogous to converting a numeric variable to z-scores, setting its mean to zero and standard deviation to one). One way to achieve this result would be to convert each individual's responses to a partial ranking, and to use these ranks as the input variable for the hexagon rankings of individuals. This method takes advantage of the fact that all respondents rated all 23 hexagons, so all individual rankings are distributed similarly in the range of 1-23. The result of converting each individual's responses to ranks is shown in Table 12.

Table 12 Ratings Ranked for each Individual

Person (Median)	Hexagon 7	Hexagon 8	Hexagon 13	Hexagon 18	Hexagon 19
A (Strongly Agree)	2	2	2	4.5	4.5
B (Agree)	2.5	1	2.5	5	4
C (Strongly Agree)	3	3	3	3	3
D (Strongly Agree)	2	2	2	4.5	4.5
E (Agree)	5	2.5	4	1	2.5

Table 13 Ranking from Individual Rankings

Person	Hexagon 7	Hexagon 8	Hexagon 13	Hexagon 18	Hexagon 19
A	1.5	2.5	1.5	3.5	4.5
B	3	1	3	5	3
C	4	5	4	2	2
D	1.5	2.5	1.5	3.5	4.5
E	5	4	5	1	1

Using individual-based rankings as the input to the hexagon ranking appears to yield a more satisfactory result because each individual respondent now has high rankings in some hexagons and low rankings in others, which means individual variation in overall response has mostly been removed. The result of removing person-person variability out of hexagon correlations is that ranking correlations are lower in every distance band and remain highest for hexagons that are adjacent, which indicates that transforming the data retained the spatial dependency while removing spurious autocorrelation caused by individual response levels. The decrease in mean measured spatial autocorrelation in a series of 5 km distance bands is shown for Opportunity in Figure 8: the red line corresponds to rankings based off raw rankings, the green line off median-transformed rankings, and the blue line off individual rankings. For the rank-based rankings, average correlation between adjacent hexagons is around 0.2, and is negative for all bands beyond 10 km.

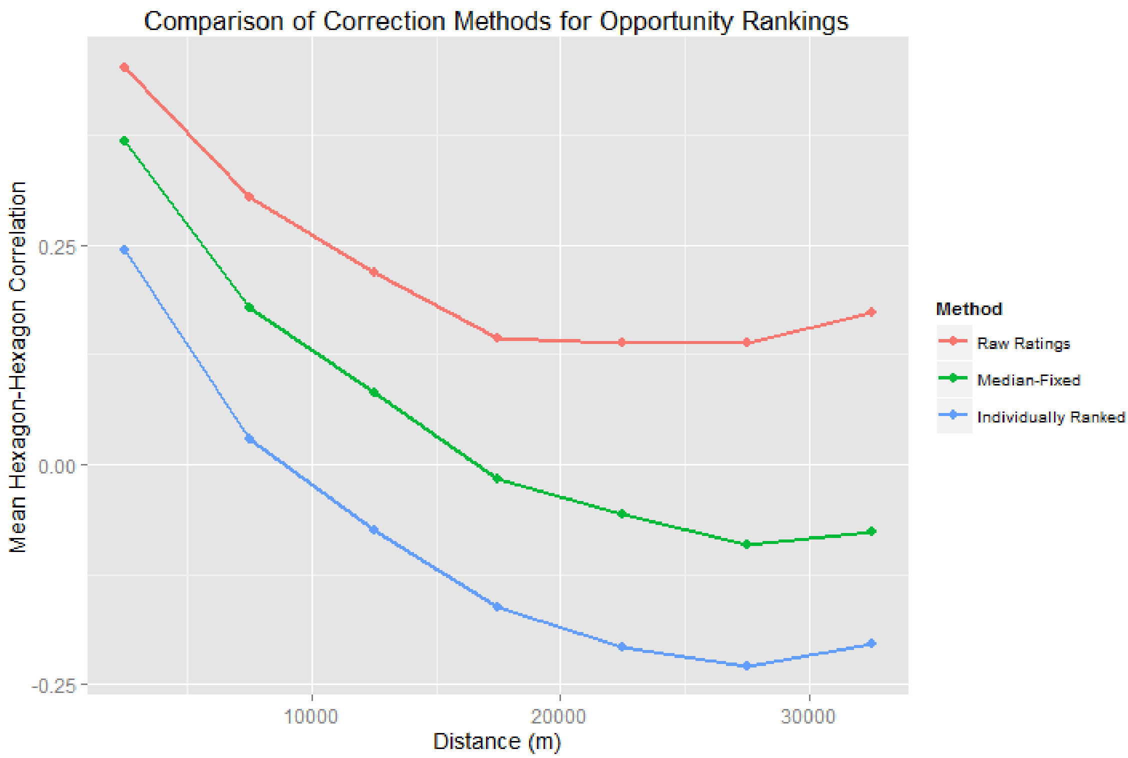

Figure 8 Mean Hexagon-Hexagon Correlation of Reported Opportunity from 3 Methods

Correlation scatterplots provide one way to view the degree of spatial autocorrelation in this dataset. The following figures show the decay of autocorrelation over space for the four place attitudes considered in this study. Each dot represents the correlation between one the rankings of one pair of hexagons with centroids a certain distance apart. Opportunity (Figure 9) and Attractiveness (Figure 10) show generally similar autocorrelation patterns: most pairs of adjacent hexagons have positive correlations between 0 and 0.5, and correlation gradually decreases with distance; near hexagons are somewhat more weakly correlated for Attractiveness than for Opportunity, and distant hexagons are correspondingly less negatively correlated. Almost all hexagons separated by more than 15 kilometers are negatively correlated with each other. Outliers in the familiarity plot at each distance correspond to similar, generally low-opportunity hexagons (typically found in the hills on the extreme north of the region). In these cases, the high correlation is caused by a split between people who consider these hexagons to present very few opportunities and people who presumably like hiking.

Familiarity (Figure 11) shows the steepest dropoff in correlation, as might be expected. People will generally be more familiar with places near other places they're familiar with because of the need to travel continuously over space. At the opposite extreme, Danger (Figure 12) shows less spatial autocorrelation than the other measures, possibly because it's varies more from person to person than from place to place.

This method of investigating spatial autocorrelation of a multilevel ordinal dataset does not address hexagon-level and person-response-level spatial dependency separately, but it removes the non-spatial component of variability between respondents. In all the variable correlation plots shown above, correlation is far higher for adjacent hexagons than at any distance beyond that; one method that could potentially reduce or eliminate the person-level spatial autocorrelation in the dataset by sampling responses. For a grid of tessellated hexagons, splitting the sample into 3 groups could eliminate first-order (shared edge) neighbors from all individuals' responses. If four groups were used, it would eliminate the next-nearest category (adjacent vertices) as well. While this would reduce the number of responses by 2/3 or 3/4, there would still be over 100 measurements for every hexagon, and the resulting hexagon means would be free of respondent-level autocorrelation.

Since it is difficult to separate hexagon means from individual observations, a method that sought to account for spatial autocorrelation at the individual level might not be able to detect it separately at the hexagon level and vice versa. It would be possible to test this for numeric data by generating a dataset with simulated multilevel spatial dependency and running multilevel regression models to see if the parameters are estimated accurately. This analysis would require the following steps: generate a set of spatially autocorrelated random hexagon means, a set of random individual means, a set of spatially autocorrelated individual-response errors, and fully random response errors. Estimate a cross-classified multilevel model and investigate whether spatial autocorrelation among estimated hexagon means and individual respondent residuals match the settings used to generate the data.

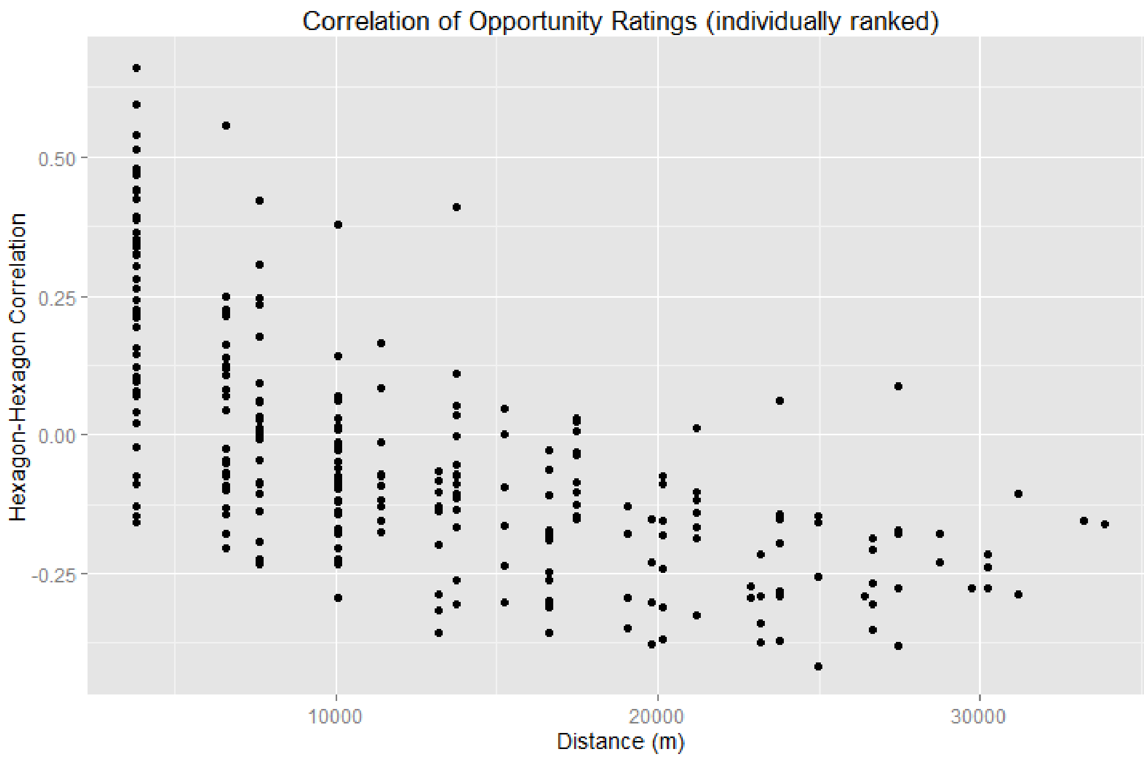

Figure 9 Spatial Correlation Scatterplot for Opportunity

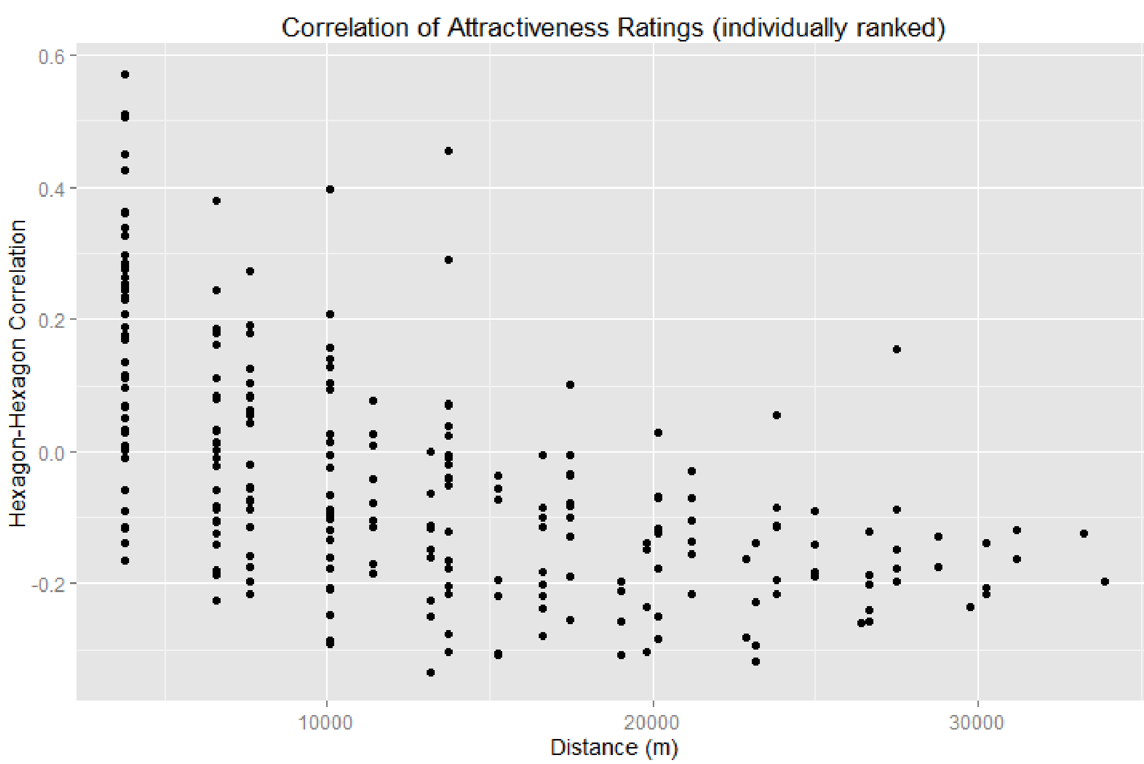

Figure 10 Spatial Correlation Scatterplot for Attractiveness

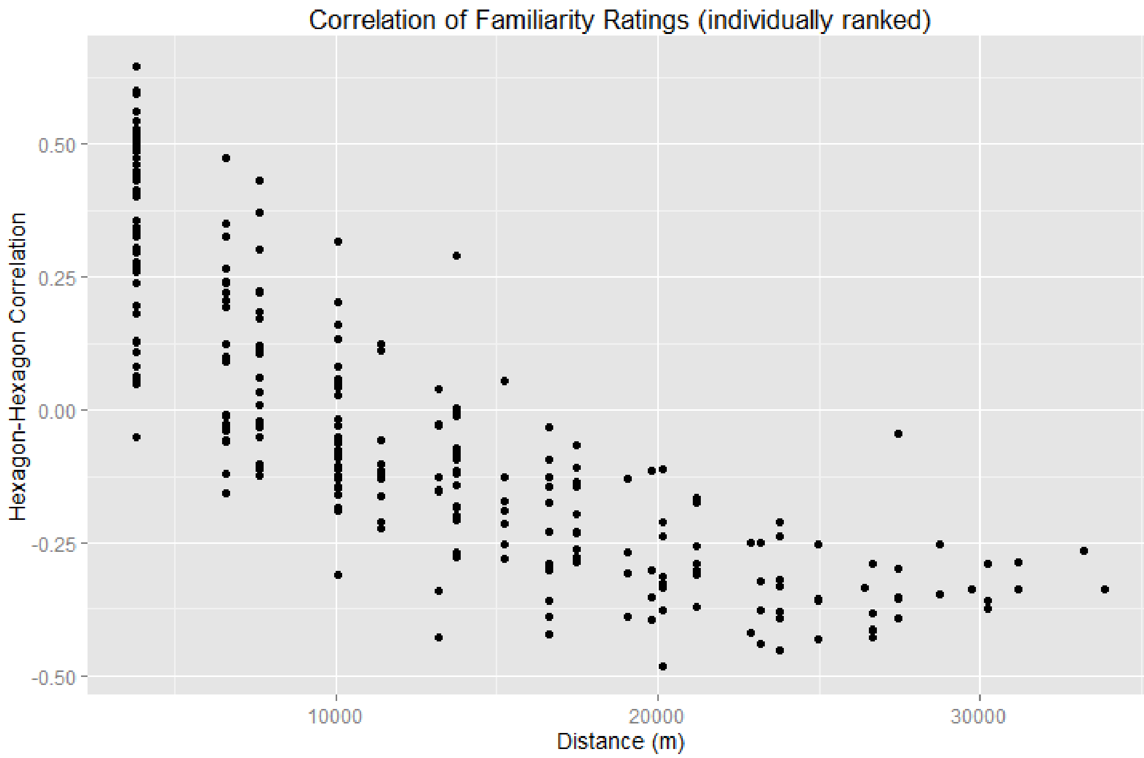

Figure 11 Spatial Correlation Scatterplot for Familiarity

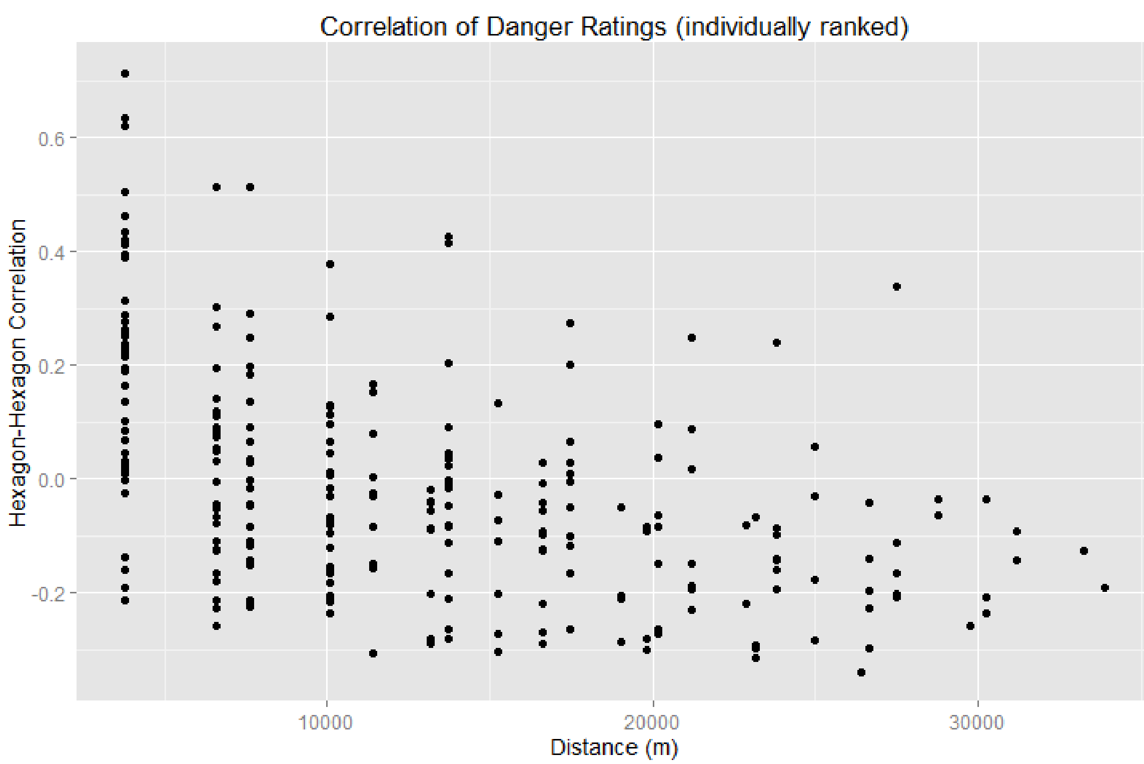

Figure 12 Spatial Correlation Scatterplot for Danger

Taste and Joint Hexagon-Individual Effects

An individual's rating of a hexagon is not merely the sum of their average response, that hexagon's average rating, and random error. There are two general categories of non-spatial person-hexagon interaction that were not addressed in this model, but do not represent threats to the independence of observations: a) interaction between person characteristics and place attributes (interaction between person-level and hexagon-level variables), and b) individuality of preference for place attributes.

Some variables in our conceptual model operate exclusively at the level of the individual and others exclusively at the hexagon level, but the way an individual interacts with a specific area cannot fit into either of those levels of the model. Home location clearly shapes one's attitudes about the surrounding area. Work and school locations are also no doubt important, since people visit these almost daily. Home location was investigated in early stages of the modeling process, but it dwarfed the effects of many of the hexagon-level variables we were interested in, since people generally chose to live in attractive and opportunity-rich hexagons. A future modeling effort with this dataset may seek to model hexagon-scale home location choice directly, and future survey efforts may attempt to measure individuals' travel through hexagons, since this likely shapes (and reflects) opinions of places.

As constructed, this model addresses variability at the level of the individual respondent and individual hexagons, as well as the respondent's interaction with a given hexagon. This structure implies that all individual characteristics have a uniform effect over space and that all people weight all hexagon attributes identically. We know this not to be true. Previous work on the GeoTrips data showed that different groups of people had very different opinions about specific hexagons (with the student town Isla Vista being especially contentious) (Deutsch, 2013). Fortunately, we consider only a few individual-level variables, and of these, only gender had any marked spatial effects – women in general rated all areas less safe than men did, but the difference was greater in Isla Vista and in the relatively remote northern hexagons than in other parts of the region. People belonging to ethnic and linguistic minorities (in Santa Barbara, the likeliest case is that Spanish speakers may like the opportunities present in the Milpas and Old-Town Goleta neighborhoods more than the rest of the sample would). Because of the limited scale of this difference and the fact that many of the clearest examples likely affect the relationship between groups and specific hexagons rather than hexagon variables, we will exclude this axis of variation from our model design for now.

Different people pursue different recreational activities and have different tastes in food and shopping and have different financial abilities, and this no doubt shapes their relationship with specific areas. GeoTrips attempted to answer this question by asking respondents to weight various factors (such as price, distance to home, and quality of good/services) that went into their destination choices, but it is unclear how to include this data in our model. Spatial differences between groups or individuals could be addressed by allowing hexagon-level coefficients to vary by including random effects in the model (Greene & Hensher, 2010, Chapter 7).

Aggregating Subjective Measures Over Space

At the scale of this study, most human activity locations would best be modeled as points. A given home or business takes up a tiny fraction of the area covered by a hexagon 4 km across, but the GeoTrips survey aggregates each respondent's attitudes about places to the scale of these hexagons. This aggregation process means the survey responses represent unknown aggregations of individuals' place attitudes. A basic question is whether the place attitudes that are being aggregated are better represented as a continuous surface or as a web of related places that people compare and travel between when pursuing their daily activities.

In addition to the question of whether hexagons better represent an aggregate assessment of a continuous surface or of a network of points, it is unknown how people weigh different values when rating a hexagon. Does a frightening stretch of mountain road render the entire area unsafe? How does one rate attractiveness of the UC Santa Barbara campus? It features secluded beaches, a gorgeous lagoon, and uninspired architecture, all within a few hundred meters of each other.

Much as individual responses are aggregated to the hexagons, all place attribute variables also represent attempts to use a single number to represent the character of a spatial area without accounting for its internal variation. All hexagon-level variables included in this model are either total values for the hexagon (shoreline length, businesses) or spatial averages (land cover fractions and Shannon entropy of local businesses). A number of dummy variables were tested (notably presence of shoreline), but the model determined that these variables generally had a less significant effect than did related continuous variables, which indicates that it is likely appropriate to include attributes averaged or summed over space. Still, it would be worthwhile to test additional dummy variables and aggregations based on minimum/maximum values for a given variable in a hexagon to see if this has an effect on the results.

Future attempts to measure people's place attitudes over space must also consider what size of hexagon is appropriate for the questions asked. There is likely a range of scales over which a given question could sensibly be answered, but if the hexagons are too small, then most cells will be essentially identical to their neighbors; if hexagons are too large, then responses will miss local variability. On the other hand, an individual's ability to answer questions about a small response region is likely to vary over space in inverse proportion with their familiarity with an area. In the area right around one's home, it might be reasonable to expect block-by-block responses to many questions that could only be answered about whole neighborhoods elsewhere in the city and only in very general terms about cities in other regions. While scale-varied responses are likely to be impractical for this sort of survey, differences in response certainty over space may be a concern. In any event, once scale is determined, respondent burden (in terms of number of hexagons collected per person), sample size (number of people), and desired responses per hexagon will determine how large of an area this sort of survey can cover.

Place Attitudes

Familiarity is clearly a different sort of variable from the others, since it measures a person's relationship to an area rather than their attitudes about it. In some sense, this question was an attempt to measure the filter through which each person perceives the attributes of an area. People less familiar with a region should generally have less predictable opinions about it. Essentially, the latent place attitude variables are likely heteroscedastic, with variance a function of familiarity. Though this does not bias linear regression coefficient estimates, it can be more of a problem for ordered data (Greene & Hensher, 2010, pp. 232–236). Additionally, familiarity is the variable most likely to change over time. People will become more familiar with areas they visit frequently, but will lose familiarity with areas they do not visit.

Though it probably should be included in the model as a dependent variable, danger is excluded. The practical concern was that models that included danger failed to converge, but its exclusion can also be partially justified by an understanding of the study area. Santa Barbara is generally quite safe, so we do not believe perceived safety plays a very large role in many people's destination choice in this region, excluding trips to Isla Vista. Danger is much less spatially stratified than the other variables we measured, only one hexagon on the east side of downtown and Isla Vista had persistently higher perceived danger than the region as a whole, and danger otherwise seems to depend almost exclusively on the individual respondent.

Though they ostensibly represent different aspects of a person's opinion of a place, accounting for place attitudes as distinct variables may not be the right approach. A latent variable model would assume that the various place attitudes represented related attempts to measure a single concept. This model would replicate the cross-classified multilevel structure shown in this thesis and would face the same difficulties with spatial and non-spatial dependence that these models face.

Chapter 5 – Estimation Results and Discussion

This section provides a comparison between two multilevel structural ordered models (the more comprehensive of which is shown in Figure 7, each of which was estimated based on the same variables. One model uses a two-level structure to account for repeated measures of respondents, but (falsely) assumes each individual was asked about a unique set of 23 hexagons. The second model uses a cross-classified design to also account for consistency and autocorrelation at the level of hexagons, because there is presumably some shared truth on which respondents base their ratings. Mplus does not provide overall model fit indices for cross-classified models since they can only be estimated using Bayesian methods. To decide on a final model formulation, I compared numerous models estimated with both a two-level and cross-classified structure and chose the model that returned the most understandable relationships with significant coefficients.

Tables 14-17 show the direct, indirect and total effects that describe the relationship between the three place attitudes and significant attributes of people and places. All effects shown in these tables relate to the latent index functions for a place attitude, not the ordinal responses directly (since they do not have a numerical interpretation), but scales are roughly consistent between response variables, though the specific thresholds vary. In this model, an increase in the index function of 0.5 will correspond to an increase of one response level on the likert scale, but this varies somewhat, as shown in Table 15. If the model finds a positive relationship between some measurable attribute of place and an axis of place attitudes, what this means is that a higher value of this attribute increase the probability that a survey respondent will score that hexagon highly and decrease the probability that they will give it a low score. Coefficients show the effect of a one point increase in a continuous independent variable (ranges shown in Table 5) or a change from false to true for a dummy variable.

The two model structures produce very similar estimates of the relationships between variables, and the discussion that follows will match both models.

Table 14 Direct Effect Standard Error Comparison, 2-Level vs Cross-Classified Model

Direct Effect SEs	2-Level			Cross-Classified		
Variable	Fam	Att	Opp	Fam	Att	Opp
Familiarity		0.013	0.014		0.011	0.012
Attractive			0.014			0.011
Danger		0.016	0.014		0.009	0.010
Local Fam.	0.021	0.014	0.016	0.009	0.012	0.013
Local Att.		0.023	0.015		0.012	0.014
Local Opp.			0.021			0.011
Local Dan.		0.017	0.016		0.012	0.012
Shoreline (km)		0.008	0.007		0.037	0.047
% Open Space	0.001	0.001		0.013	0.006	
% Roads		0.002			0.012	
% Housing	0.000	0.000		0.006	0.003	
Consumer Ests	0.000		0.000	0.001		0.001
Shannon Entropy	0.015		0.011	0.212		0.221
Female	0.017	0.019		0.019	0.020	
Age 18-25	0.018	0.032		0.032	0.032	
Has Car	0.031			0.063		

Table 15 Response Thresholds for Ordinal Link. Numbers show the minimum value that corresponds to a given response (e.g. the first row shows the break point between Strongly Agree and Agree).

	2-Level			Cross-Classified		
Response	Fam	Att	Opp	Fam	Att	Opp
Strongly Agree	5.254	5.253	5.072	5.431	5.573	5.711
Agree	4.533	4.347	4.394	4.678	4.641	5.003
Slightly Agree	3.779	3.668	3.718	3.880	3.944	4.305
Neutral	2.960	2.576	2.412	3.018	2.821	2.979
Slightly Disagree	2.444	2.100	1.966	2.477	2.327	2.529
Disagree	1.824	1.514	1.385	1.830	1.719	1.945

Table 16 2-Level (Person) Direct and Indirect Effect Coefficients

	2-Level	Between Responses (for a given person)						
		Familiarity	Attractiveness			Opportunity		
		Direct	Direct	Indirect	Total	Direct	Indirect	Total
SEM	Familiarity		0.099		0.099	0.282	0.048	0.330
	Attractive					0.170		0.170
Survey/Spatial	Danger		-0.239		-0.239	-0.061	-0.041	-0.102
	Local Fam.	0.820	-0.044	0.081	0.037	-0.209	0.238	0.029
	Local Att.		0.768		0.768	-0.114	0.131	0.017
	Local Opp.					0.738		0.738
	Local Dan.		0.169		0.169	0.048	0.029	0.077
Hexagon Atts.	Shoreline (km)		0.119		0.119	0.039	0.020	0.059
	% Open Space	-0.015	0.007	-0.001	0.006		-0.003	-0.003
	% Roads		-0.015		-0.015		-0.003	-0.003
	% Housing	-0.010	0.006	-0.001	0.005		-0.002	-0.002
	Consumer Ests	0.001		0.000	0.000	0.001	0.000	0.001
	Entropy	0.232		0.023	0.023	-0.007	0.069	0.062

		Between People						
		Familiarity	Attractiveness			Opportunity		
		Direct	Direct	Indirect	Total	Direct	Indirect	Total
Person Atts.	Female	-0.048	0.051	-0.005	0.046		-0.006	-0.006
	Age 18-25	-0.036	-0.030	-0.004	-0.034		-0.016	-0.016
	Has Car	0.052		0.005	0.005		0.016	0.016

Table 17 Cross Classified (Person and Hexagon) Direct and Indirect Effect Coefficients

	Cross-Classified	Specific to Hexagon-Person Pairing						
		Familiarity	Attractiveness			Opportunity		
		Direct	Direct	Indirect	Total	Direct	Indirect	Total
SEM	Familiarity		0.167		0.167	0.342	0.067	0.409
	Attractive					0.196		0.196
Survey/Spatial	Danger		-0.205		-0.205	-0.070	-0.040	-0.110
	Local Fam.	0.870	-0.101	0.145	0.044	-0.244	0.306	0.062
	Local Att.		0.796		0.796	-0.119	0.156	0.037
	Local Opp.					0.756		0.756
	Local Dan.		0.127		0.127	0.059	0.025	0.084

		Between Hexagons						
		Familiarity	Attractiveness			Opportunity		
		Direct	Direct	Indirect	Total	Direct	Indirect	Total
Hexagon Atts.	Shoreline (km)		0.123		0.123	0.072	0.024	0.096
	% Open Space	-0.018	0.007	-0.003	0.004		-0.005	-0.005
	% Roads		-0.008		-0.008		-0.002	-0.002
	% Housing	-0.010	0.007	-0.002	0.005		-0.002	-0.002
	Consumer Ests	0.001		0.000	0.000	0.001	0.000	0.001
	Entropy	0.200		0.033	0.033	0.047	0.075	0.122

		Between People						
		Familiarity	Attractiveness			Opportunity		
		Direct	Direct	Indirect	Total	Direct	Indirect	Total
Person Atts.	Female	-0.044	0.048	-0.007	0.041		-0.007	-0.007
	Age 18-25	-0.031	-0.033	-0.005	-0.038		-0.018	-0.018
	Has Car	0.068		0.011	0.011		0.025	0.025

Main Model

As expected, all three place attitudes are positively correlated; familiarity and opportunity are most strongly related (which matches the apparent relationship shown in Figures 2, 3, and 4), and the weakest correspondence is between familiarity and attractiveness. GeoTrips collected respondents' sense of danger in each hexagon in the same way it collected familiarity, attractiveness, and opportunity, but I excluded it from the list of endogenous variables because it seems least similar to the other variables and because attempts to do so failed to converge. Danger is included as an explanatory variable using the 7 numeric likert scores, which is an approximation but tolerable (Grilli & Rampichini, 2011). Unsurprisingly, danger is negatively correlated with our other variables, particularly attractiveness. A structural equations regression model provides one way to investigate linkages among the place attitudes, but it may not be the only way. As discussed in the Place Attitudes section, familiarity may work very differently from the other variables.

The individual respondents in the survey can be safely assumed to have responses that are independent among survey participants. The same is necessarily not true of the survey hexagons, since they share a spatial relationship. Spatial autocorrelation is likely to make nearby hexagons similar in both measurable and unmeasurable ways as well as to make a given individual's attitudes about neighboring hexagons similar independent of the "truth" about those areas. As discussed in the chapter on methods, the spatial relationship can be partly addressed by including appropriate exogenous variables that are also spatially varied (as many of hexagon-level variables are), but this may not eliminate the spatial autocorrelation of individuals' responses. It may be best to address this by adding an autoregressive term to the model, but this is currently impossible in Mplus. Instead, the model includes personal spatially lagged terms that contain each individual's average response in the adjacent hexagons (e.g. for the entry representing respondent 88 and hexagon 8, the variable Local Attractiveness would equal the average attractiveness score that respondent reported for hexagons 7, 10, and 11). This is an imperfect solution (both since it represents an average of ordinal data and since it could involve "double counting" person-level and hexagon-level variation), but these spatially lagged variables are highly significant coefficients in the models and operate in a positive direction, which makes sense. Spatially lagged variables have a strong positive relationship with the same variable and a weaker positive total effect on other variables. Essentially this means that familiar/attractive/opportunity-rich hexagons are near other similar hexagons and that individuals tend to have similar opinions about adjacent regions.

Hexagon-level relationships are the main target of our model, since they show us which attributes of a place seem to have an important bearing on what people think about that place. Because this level is treated most differently by the two versions of our model, total effects of hexagon-level variables are somewhat less consistent, though in all cases the signs on total effects remain the same.

Several attributes of the natural environment were related to hexagon attitudes. Unsurprisingly, coastal hexagons were perceived as being very attractive. Though length of shoreline may not be the perfect measure with which to understand this effect, it performed better than other proxies, such as total ocean area in hexagon or an ocean dummy variable. Some stretches of the Santa Barbara coast are inaccessible, but coastline generally provides opportunities for recreation, which is reflected by its positive total effect on opportunity. Open space (from parcel data) somewhat increases an area's attractiveness, but decreases people's familiarity with it (possibly because much of the open space in the area is covered by steep mountains) and slightly decreases the perceived level of opportunities. More direct investigation of this could allow us to differentiate between the effects of open space in terms of the opportunities and restrictions it creates.

The built environment also impacts place attitudes. Roads have a slight negative effect on attractiveness that was not significant in the cross-classified model; road area direct effects on familiarity and opportunity were not included in the final model because none of our models found them to be significant. While respondents were generally extremely familiar with their home hexagons, the large number of hexagons with large amounts of housing but no other features to attract visitors may help explain the counterintuitive negative relationship between housing area and familiarity. Endless housing tracts that contain no other destinations are profoundly unattractive destinations, and cause their residents to make longer trips for shopping, socializing, and entertainment.

The presence of customer-serving businesses is positively related to both familiarity and perceived opportunity, which is no surprise – these businesses are opportunities. Though the effect appears small in this model, note that this is measured per business, and many of the hexagons contain hundreds of businesses. People also seem to value diversity of businesses, which can be seen in the strong positive relationship between Shannon Entropy and both familiarity and opportunity (and attractiveness, indirectly).

Some attributes of individuals had a consistent relationship with their responses. Women generally reported lower familiarity throughout the region. Younger people and people without cars were also generally less familiar with the region, which is not surprising. The effect of gender was clearly significant at the 0.05 level in both model designs, but neither of the other variables quite reached that threshold; however, I include them for illustration. Though several different income-related variables were tested, none had a significant effect in any of our models. For person-level variables, it is always somewhat unclear whether the observed relationships represent true differences in the way groups of people relate to the region or with systematic differences in the way people interpreted the survey questions. Because each individual has a different relationship with the region, there are likely to be differences in the spatial patterns shown by individuals' responses; while this model addresses the overall variability of responses as it relates to measurable attributes of the hexagons (e.g., do people generally prefer areas with more stores and restaurants?) as well as differences between individuals' attitudes about the region as a whole (e.g., do women generally rate Santa Barbara more or less safe?), it does not directly address differences in place attitudes that reflect both personal and spatial effects. When these differences are purely individual, our model will adequately capture them in the between-response error term, but our model does not address systematic spatial differences that exist between groups of people (e.g., in addition to rating the whole region less safe than men, women rated the college town Isla Vista as particularly unsafe; Mexican-Americans are likely to rate hexagons that include stretches of Milpas higher than other people due to the large variety of restaurants and groceries that cater to that community in that part of Santa Barbara).

The two models paint very similar pictures of the relationship between attributes of and attitudes about places, but they differ greatly on the significance of these relationships with respect to hexagon-level variables. As shown in Table 14, the 2-level model that correctly addresses the non-independence of responses from an individual person estimates much smaller standard errors for all hexagon-level variables. By failing to account for true groupings in the data, models substantially underestimate standard errors, leading to overstatements of variable significance (Fielding & Goldstein, 2006, p. 23). Essentially, the 2-level model assumes that variables linked to an individual respondent are measured once per respondent (561 times, which is large enough to confirm the effects of gender, if not quite of car ownership, since most people have cars) and all other effects are measured 12,903 times (561 respondents, each with 23 observations). In reality there are only 23 independent sets of observations of hexagon-specific variables (and this only if spatial autocorrelation of neighboring hexagons is ignored). Since the hexagon level of the model operates with only 23 degrees of freedom, it should be no surprise that it only finds the strongest relationships to be significant (namely shoreline and both appearances of consumer establishment counts, though % Housing and % Open Space are close).

Twitter Model and Meaningful Proxy Variables

A model that includes the Twitter Happiness variable is presented in Table 17. Though this variable is generally higher in more attractive and opportunity-rich hexagons, as shown in Figure 6, it takes a highly significant negative sign in the model. In addition, the inclusion of this variable shifts the coefficient estimates for many other variables.

It is reasonable to wonder what sort of a relationship would be expected between a measure happiness of tweets sent from an area and people's subjective assessment of that area. Even if all variables involved were measured without error or bias, there are fundamental differences between what this variable attempts to measure and any of the place attitudes. People's happiness *in* a place is quite different from their attitudes *about* that place. Previous work with this variable has shown that this measure varies in consistent and predictable ways over space, which means it may be a useful proxy measure of some aspect of people's relationships with a place, but this relationship may be incidental. People's moods may be influenced by the attractiveness of their surroundings, but other events in their life likely have an impact as well and may be more likely to show up in their tweets. Additionally, the presence of a few particularly emotionally intense locations within a hexagon may be more likely to generate tweets containing words that the hedonometer picks up, but these sites may not be taken into account when survey respondents rated it for attractiveness, opportunity, and familiarity.

In addition to questions about the true relationship between happiness in a place and attraction to that place, there are two other factors that may limit the hedonometer's usefulness as a variable in general and with respect to this study in particular. This measure of place happiness is based on simple word counts that ignore the context in which these words were used within the tweet and the (uncollected) personal context in which the person posted the given tweet. Additionally, while the survey respondents are older than the region's population as a whole, Twitter users are generally younger, meaning that the biases of these datasets likely run in opposite directions for any measure that varies with age. Addressing Twitter's inherent biases and other sources of error seems key to increasing the value of harvested tweets to transportation research.

The Shannon entropy measure calculated from counts in 80 industry categories presents a different sort of problem. While diversity seems like an obvious variable to use and has a fairly clear positive relationship to any measure of the opportunities provided by a region, this may not be the best way to measure it. Whereas the hedonometer is relatively stable over this region, the summed logarithmic calculation for entropy means that it can behave poorly in edge cases (in this case, when very few businesses or categories are represented). While the hedonometer is based on an empirical ranking of words, this entropy measure is reliant on a specific classification scheme that it unrealistically treats as absolute. In contrast, though the hedonometer has a peculiar effect in this model and may not actually be a proxy for place happiness, it has a number of features that should make it behave better as a proxy variable in general: it is distributed roughly symmetrically and varies consistently both in this case and in broader studies (Dodds et al., 2011).

Table 18 Cross-Classified (Person and Hexagon) Direct and Indirect Effects with Twitter Happiness

	Cross-Classified	Specific to Hexagon-Person Pairing						
		Famil.	Attractiveness			Opportunity		
		Direct	Direct	Indirect	Total	Direct	Indirect	Total
SEM	Familiarity		0.167		0.167	0.342	0.067	0.409
	Attractiveness					0.196		0.196
Survey/Spatial	Danger		-0.205		-0.205	-0.069	-0.040	-0.109
	Local Familiarity	0.871	-0.102	0.145	0.043	-0.244	0.306	0.062
	Local Attractiveness		0.795		0.795	-0.119	0.156	0.037
	Local Opportunity					0.756		0.756
	Local Danger		0.126		0.126	0.059	0.025	0.084

		Between Hexagons						
		Famil.	Attractiveness			Opportunity		
		Direct	Direct	Indirect	Total	Direct	Indirect	Total
Hexagon Atts.	Shoreline (km)		0.125		0.125	0.072	0.025	0.097
	% Open Space	-0.017	0.008	-0.003	0.005		-0.005	-0.005
	% Roads		-0.015		-0.015		-0.003	-0.003
	% Housing	-0.010	0.008	-0.002	0.006		-0.002	-0.002
	Consumer Ests	0.129		0.022	0.022	0.122	0.048	0.170
	Entropy	0.238		0.040	0.040	0.056	0.089	0.145
	Hedonometer		-0.470		-0.470		-0.092	-0.092

		Between People						
		Famil.	Attractiveness			Opportunity		
		Direct	Direct	Indirect	Total	Direct	Indirect	Total
Person Atts.	Female	-0.047	0.044	-0.008	0.036		-0.009	-0.009
	Age 18-25	-0.031	-0.034	-0.005	-0.039		-0.018	-0.018
	Has Car	0.053		0.009	0.009		0.020	0.020

Chapter 6 – Summary and Conclusion

In this thesis, I analyze the relationship between place attitudes and measurable place attributes while accounting for interpersonal and spatial variation in a realistic way. I do this by testing the results of a spatially-constrained place attitudes survey for the southern Santa Barbara County, CA against a variety of place attribute variables. Santa Barbara's spectacular setting provides an excellent study area for the relationship between features of an urban area's natural environment and people's attitudes of it, but this may limit our ability to see the significance of aspects of the built environment. People living in an area in which development is not hemmed in by steep mountain ranges and the ocean may respond very differently to their city's geography. An additional model is developed that examines the utility of a measure of place happiness created from harvested tweets as an explanatory variable for this sort of model. Our model development process points to possible ways that subjective place attitudes and objectively measured spatial attributes can be linked. Though my findings were hampered by the small number of hexagons and thus could not make any firm declarations about what makes people like certain places more than others, this thesis can serve as a starting point for related research by suggesting which variables should be investigated further.

This model indicates that place attitudes measures are strongly related to measurable attributes of places in ways that generally make sense, which both validates their usefulness as measures of people's relationship with places and indicates that they can (and should) be considered when trying to understand people's spatial decision making. Because of concerns about spatial autocorrelation, it is especially important to develop spatial metrics that can explain some of this autocorrelation.

The final model does not account for all aspects of place attitudes and is also imperfect in its treatment of spatial autocorrelation and related issues. We do not know whether the specific hexagon tessellation has especially different effects from any other. Additionally, the hexagons are likely too large to study some significant aspects of place. Because the hexagon structure smooths out differences across space and because the small number of hexagons limits the analysis of hexagon-level variability to 23 degrees of freedom, the results of models made from this dataset are likely to be more limited than they would be if we were able to model at a finer spatial scale. Despite these limitations, this study finds some significant relationships between place attitudes and measurable attributes of place. Coastline and open space boost the attractiveness of parts of an already scenic region. Classified business establishment counts are a reasonable proxy for opportunities experienced by area residents, especially if business establishment diversity is also taken into account. In an area with large swathes of uniform residential development, people know their home neighborhood well, but may be very unfamiliar with similar areas elsewhere in the region.

This model also demonstrates the necessity of using more correct model structures when working with complex datasets like the GeoTrips survey. Failing to account for hexagon-level repeated measures meant that many variables were initially determined to be much more clearly significant than they are. However, the similarity in output between two model structures (one that accounts for the repeated measurement of hexagons and another that does not) suggests that a simpler model accounting only for the more variable grouping (analysis presented in the section under Methods – Discussion and Critique suggested that for most place attitudes measures, person-person differences accounted for about twice as much variability as hexagon-hexagon differences) can be used to test multiple model formulations more quickly, since cross-classified models generally took several hours to estimate.

One key consideration excluded from our model is home location, which clearly relates strongly to familiarity and the other aspects of place attitudes. Unfortunately, home location dwarfed the effects of many of the hexagon-level variables we were interested in, so we excluded it from our final model. A future modeling effort with this dataset may seek to model hexagon-scale home location choice directly and jointly with the rest of the place attitudes variables.

Future Work – Social Media

While Twitter data turned out not to be particularly useful in this model, it has many potential uses in geographic research. While the severe imbalance of Tweet frequency among different people makes it difficult to extract spatially aggregate information, the presence of very heavy users raises the possibility of using Twitter to collect longitudinal data. This could make it possible to model the growth and change in personal action spaces over time, link these to estimated demographic characteristics based on users home locations (which are distressingly easy to guess given a large enough set of geotagged tweets from one person), and investigate people's emotional ties to certain places and activities. Additionally, the high spatial and temporal density that long-term tweet harvesting provides would be an excellent data source with which to investigate differences in activity patterns by time of day and links between specific activity types and specific business locations.

Aggregate happiness variables cannot capture all the variability of place meaning that tweets contain. The frequencies with which specific words are used can provide much more information, and word clouds are a particularly useful way of visualizing this information. Figure 13 contains four hexagon-level word clouds from our study. The two on the left correspond to wealthy, coastal Montecito (hexagon 5, happiness = 6.39) and the two on the right to the student community Isla Vista (hexagon 20, happiness = 5.92); Tagxedo.com, the service that produced these clouds aggregates related words (e.g. "stopped" and "stopping" get grouped under "stop"). In Montecito, happy words largely describe experiences outdoors ("butterfly," "seasons," "beautiful," and "coast"), but Isla Vista's tweets reflect the intensely social atmosphere experienced by the thousands of students who live in the town: a much greater share of the positive words refer to people. The hexagons' sad words are much more similar, though more of Montecito's may relate to relate to travel delays ("wait," "traffic," "slow," and "stop") and the consequences of partying ("drunk") appear much more prominently in Isla Vista.

The primary finding of this study with respect to data produced from harvested tweets is that while the data is not perfect, it could potentially be useful in models. Previously published tweet happiness findings, the initial comparison between Twitter happiness in Santa Barbara and the place attitudes, and the word clouds presented here show that the textual contents of tweets partly reflect the character of the region from which they are sent. This indicates that Twitter-derived variables may provide a valuable input for models relating people to places at a different scale than was presented in this thesis.

One potentially attractive, but likely infeasible direction to take this research would be to reach out to heavy Twitter users and survey them about their attitudes towards areas they travel through most. By pairing this with a user-targeted tweet collection effort, this would allow for much more direct linkage between social media behavior, place attitudes, and travel behavior. Some questions to consider for this research effort would include:

- How many subjects would we be required to extract significant results?
- Is this type of study practical in Santa Barbara or would it require a larger area?
- Would we be able to ask people to turn on geotagging?
- Would explicit knowledge of surveillance change the ways our subjects tweeted?
- How do you reward someone for their participation in a study, when their participation is largely passive?

- How could we provide anonymity within our database?
- Could it return sufficiently interesting results to be worth the incredible effort and IRB headaches it would certainly entail?

Figure 13 Happy and Sad Word Clouds for Hexagons 5 (left) and 20 (right), copyright Tagxedo.com

Future Work – Place Attitudes

The GeoTrips survey has been an incredibly valuable resource for the GeoTrans lab, but its potential value has not been exhausted. This dataset and information harvested from a real estate website like Zillow could be used to develop a two-level housing choice model that accounted first for overall spatial preference in the region, and then for specific home within a hexagon. Alternatively, since the place attitudes variables may represent multiple attempts to measure the same thing, they could be recast into a latent variable model.

Deutsch's hexagon-based collection of place attitudes metrics in GeoTrips was innovative and valuable, but future work in this area would benefit from a finer scale of measurement. One solution would be to repeat the survey with a much larger number of much smaller hexagons, with each respondent given a random subset to rate. This would achieve three main goals: 1) Smaller hexagons may be more accurately summarized by measurable attributes; 2) hexagon-level relationships could be investigated with many more degrees of freedom, substantially improving the strength of any model we build; 3) random spatial sampling for each respondent would diminish the spatial autocorrelation of each person's responses; 4) though this may increase or decrease hexagon-level spatial autocorrelation, much of this is due to true similarities among nearby hexagons, which can be addressed by including other spatially correlated variables in a model, and these metrics may well be more useful at a finer spatial scale (since they will capture more variability).

If I were to rerun a survey like GeoTrips, in addition to using smaller hexagons, I would consider the following changes:

- Ask fewer questions about the abstract reasons people make decisions.
- Ask for actual travel data in terms of trips to or through hexagons.
- Ask about typical destinations for certain types of activities.

- Ask about atypical destinations. This could potentially include destinations for celebratory dinners, shopping trips (including long-distance trips) to meet special needs or special occasions (e.g. Korean families driving to Los Angeles to acquire hard-to-find staple ingredients or Latino families acquiring dough to make tamales on Christmas Eve).

References

About Twitter, Inc. (2015). Retrieved June 8, 2015, from https://about.twitter.com/company

Agnew, J. A. (2002). *Place and Politics in Modern Italy*. University of Chicago Press.

Agnew, J. A., & Duncan, J. S. (2014). *The Power of Place (RLE Social & Cultural Geography): Bringing Together Geographical and Sociological Imaginations*. Routledge.

Aitken, S. C., & Prosser, R. (1990). Residents' Spatial Knowledge of Neighborhood Continuity and Form. *Geographical Analysis*, *22*(4), 301–325. http://doi.org/10.1111/j.1538-4632.1990.tb00213.x

Arentze, T. A., & Timmermans, H. J. P. (2005). Information gain, novelty seeking and travel: a model of dynamic activity-travel behavior under conditions of uncertainty. *Transportation Research Part A: Policy and Practice*, *39*(2–3), 125–145. http://doi.org/10.1016/j.tra.2004.08.002

Asparouhov, T., & Muthen, B. (2014). *General Random Effect Latent Variable Modeling: Random Subjects, Items, Contexts, and Parameters*. Statmodel. Retrieved from https://www.statmodel.com/download/NCME_revision2.pdf

Ben-Akiva, M., Mcfadden, D., Train, K., Walker, J., Bhat, C., Bierlaire, M., … Munizaga, M. A. (2002). Hybrid Choice Models: Progress and Challenges. *Marketing Letters*, *13*(3), 163–175. http://doi.org/10.1023/A:1020254301302

Bertrand, K. Z., Bialik, M., Virdee, K., Gros, A., & Bar-Yam, Y. (2013). *Sentiment in New York City: A High Resolution Spatial and Temporal View*. Cambridge, MA: New England Complex Systems Institute. Retrieved from http://www.necsi.edu/research/social/newyork/

Bhat, C. R. (2000). A multi-level cross-classified model for discrete response variables. *Transportation Research Part B: Methodological*, *34*(7), 567–582. http://doi.org/10.1016/S0191-2615(99)00038-7

Bhat, C. R., & Dubey, S. K. (2014). A new estimation approach to integrate latent psychological constructs in choice modeling. *Transportation Research Part B: Methodological*, *67*, 68–85. http://doi.org/10.1016/j.trb.2014.04.011

Bhat, C. R., & Guo, J. (2004). A mixed spatially correlated logit model: formulation and application to residential choice modeling. *Transportation Research Part B: Methodological*, *38*(2), 147–168. http://doi.org/10.1016/S0191-2615(03)00005-5

Bolton, R. (1989). *AN ECONOMIC INTERPRETATION OF A SENSE OF PLACE* (Department of Economics Working Paper). Department of Economics, Williams College. Retrieved from http://econpapers.repec.org/paper/wilwileco/130.htm

Brechin, G. (2007). *Imperial San Francisco : Urban Power, Earthly Ruin*. Berkeley, CA: Univ. of California Press.

Brown, B. B., & Werner, C. M. (2008). Before and After a New Light Rail Stop: Resident Attitudes, Travel Behavior, and Obesity. *Journal of the American Planning Association*, *75*(1), 5–12. http://doi.org/10.1080/01944360802458013

Buckley, J. M. (1997). A Factory without a Roof: The Company Town in the Redwood Lumber Industry. *Perspectives in Vernacular Architecture*, *7*, 75–92. http://doi.org/10.2307/3514386

Canter, D. (1977). *The Psychology of Place*. St. Martin Press.

Canter, D. (1983). The Purposive Evaluation of Places A Facet Approach. *Environment and Behavior*, *15*(6), 659–698. http://doi.org/10.1177/0013916583156001

Cao, X., Mokhtarian, P. L., & Handy, S. L. (2007). Do changes in neighborhood characteristics lead to changes in travel behavior? A structural equations modeling approach. *Transportation*, *34*(5), 535–556. http://doi.org/10.1007/s11116-007-9132-x

Certoma, C. (2009). Environmental Politics and Place Authenticity Protection. *Environmental Values*, *18*(3), 313–341. http://doi.org/10.3197/096327109X12474739376497

Cervero, R., & Kockelman, K. (1997). Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research Part D: Transport and Environment*, *2*(3), 199–219. http://doi.org/10.1016/S1361-9209(97)00009-6

Chan, R., & Schofer, J. (2014). Role of Social Media in Communicating Transit Disruptions. *Transportation Research Record: Journal of the Transportation Research Board*, *2415*, 145–151. http://doi.org/10.3141/2415-16

Clarke, M. (2012). Achieving Memorable Places … "Urban Sense of Place" for Successful Urban Planning and Renewal? In I. Convery, G. Corsane, & P. Davis

(Eds.), *Making Sense of Place: Multidisciplinary Perspectives* (pp. 119–132). Suffolk, UK: Boydell Press.

Convery, I., & O'Brien, V. (2012). Gypsies, Travellers and Place: A Co-ethnography. In I. Convery, G. Corsane, & P. Davis (Eds.), *Making Sense of Place: Multidisciplinary Perspectives* (pp. 43–53). Suffolk, UK: Boydell Press.

County GIS Spatial Catalog. (2015). [Government]. Retrieved March 20, 2015, from http://cosb.countyofsb.org/gis/default.aspx?id=2802

Davenport, M. A., & Anderson, D. H. (2005). Getting From Sense of Place to Place-Based Management: An Interpretive Investigation of Place Meanings and Perceptions of Landscape Change. *Society & Natural Resources*, *18*(7), 625–641. http://doi.org/10.1080/08941920590959613

de Abreu e Silva, J., Golob, T., & Goulias, K. (2006). Effects of Land Use Characteristics on Residence and Employment Location and Travel Behavior of Urban Adult Workers. *Transportation Research Record: Journal of the Transportation Research Board*, *1977*(-1), 121–131. http://doi.org/10.3141/1977-17

DeJong, T. M. (1975). A Comparison of Three Diversity Indices Based on Their Components of Richness and Evenness. *Oikos*, *26*(2), 222–227. http://doi.org/10.2307/3543712

Deutsch, K. (2013, June). *An Investigation in Decision Making and Destination Choice Incorporating Place Meaning and Social Network Influences* (Doctoral Dissertation). University of California, Santa Barbara, Santa Barbara, CA.

Deutsch, K., & Goulias, K. (2010). Exploring Sense-of-Place Attitudes as Indicators of Travel Behavior. *Transportation Research Record: Journal of the Transportation Research Board*, *2157*, 95–102. http://doi.org/10.3141/2157-12

Deutsch, K., & Goulias, K. G. (2013). Decision makers and socializers, social networks and the role of individuals as participants. *Transportation*, *40*(4), 755–771. http://doi.org/10.1007/s11116-013-9465-6

Deutsch, K., Ravualaparthy, S., & Goulias, K. (2013). Place Happiness: Its Constituents and the Influence of Emotions and Subjective Importance on Activity Type and Destination Choice.

Deutsch, K., Yoon, S. Y., & Goulias, K. (2013). Modeling travel behavior and sense of place using a structural equation model. *Journal of Transport Geography*, *28*, 155–163. http://doi.org/10.1016/j.jtrangeo.2012.12.001

Dodds, P. S., Harris, K. D., Kloumann, I. M., Bliss, C. A., & Danforth, C. M. (2011). Temporal Patterns of Happiness and Information in a Global Social Network: Hedonometrics and Twitter. *PLoS ONE*, *6*(12), e26752. http://doi.org/10.1371/journal.pone.0026752

Douglas, G. C. C. (2012). The Edge of the Island: Cultural Ideology and Neighbourhood Identity at the Gentrification Frontier. *Urban Studies*, *49*(16), 3579–3594. http://doi.org/10.1177/0042098012448549

Duggan, M., Ellison, N. B., Lampe, C., Lenhart, A., Madden, M., Rainie, L., & Smith, A. (2015). *Social Media Update 2014: While Facebook remains the most popular site, other platforms see higher rates of growth*. Washington, D.C.: Pew Research

Center. Retrieved from http://www.pewinternet.org/2015/01/09/demographics-of-key-social-networking-platforms-2/

Ewing, R., & Cervero, R. (2001). Travel Behavior and the Built Environment. *Transportation Research Record: Journal of the Transportation Research Board*, *1780*, 87–114.

Farag, S., Schwanen, T., Dijst, M., & Faber, J. (2007). Shopping online and/or in-store? A structural equation model of the relationships between e-shopping and in-store shopping. *Transportation Research Part A: Policy and Practice*, *41*(2), 125–141. http://doi.org/10.1016/j.tra.2006.02.003

Ferguson, M. R., & Kanaroglou, P. S. (1995). Utility Variability within Aggregate Spatial Units and its Relevance to Discrete Models of Destination Choice. In P. D. L. Anselin & D. R. J. G. M. Florax (Eds.), *New Directions in Spatial Econometrics* (pp. 243–269). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-79877-1_11

Fielding, A. (1997). On scoring ordered classifications. *British Journal of Mathematical and Statistical Psychology*, *50*(2), 285–307. http://doi.org/10.1111/j.2044-8317.1997.tb01146.x

Fielding, A., & Goldstein, H. (2006). Cross-classified and multiple membership structures in multilevel models: an introduction and review. Retrieved from http://dera.ioe.ac.uk/6469/1/RR791.pdf

Francaviglia, R. V. (1971). The Cemetery as an Evolving Cultural Landscape. *Annals of the Association of American Geographers*, *61*(3), 501–509. http://doi.org/10.1111/j.1467-8306.1971.tb00802.x

Ghose, R. (2004). Big Sky or Big Sprawl? Rural Gentrification and the Changing Cultural Landscape of Missoula, Montana. *Urban Geography*, *25*(6), 528–549. http://doi.org/10.2747/0272-3638.25.6.528

Golledge, R. G., & Spector, A. N. (1978). Comprehending the Urban Environment: Theory and Practice. *Geographical Analysis*, *10*(4), 403–426. http://doi.org/10.1111/j.1538-4632.1978.tb00667.x

Golledge, R. G., & Stimson, R. J. (1997). *Spatial Behavior: A Geographic Perspective*. New York: The Guilford Press.

Goodchild, M. F. (2007). Citizens as sensors: the world of volunteered geography. *GeoJournal*, *69*(4), 211–221. http://doi.org/10.1007/s10708-007-9111-y

Grant, J. (2002). Mixed Use in Theory and Practice: Canadian Experience with Implementing a Planning Principle. *Journal of the American Planning Association*, *68*(1), 71–84. http://doi.org/10.1080/01944360208977192

Greene, W. H., & Hensher, D. A. (2010). *Modeling Ordered Choices*. Cambridge.

Grilli, L., & Rampichini, C. (2011). ch 19 Multilevel Models for Ordinal Data. In *Modern Analysis of Customer Surveys: With Applications Using R* (pp. 391 – 411).

Groth, P. E. (1994). *Living Downtown: The History of Residential Hotels in the United States*. University of California Press.

Harvey, D. (1993). From space to place and back again: Reflections on the condition of postmodernity. In J. Bird, B. Curtis, T. Putnam, & L. Tickner (Eds.), *Mapping the Futures: Local Cultures, Global Change*. London: Routledge.

Harvey, D. (2006). The Sociological and Geographical Imaginations. *International Journal of Politics, Culture, and Society*, *18*(3-4), 211–255. http://doi.org/10.1007/s10767-006-9009-6

Hiruta, S., Yonezawa, T., Jurmu, M., & Tokuda, H. (2012). Detection, Classification and Visualization of Place-triggered Geotagged Tweets. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing* (pp. 956–963). New York, NY, USA: ACM. http://doi.org/10.1145/2370216.2370427

Horlings, L. G. (2015). Values in place; A value-oriented approach toward sustainable place-shaping. *Regional Studies, Regional Science*, *2*(1), 256–273. http://doi.org/10.1080/21681376.2015.1014062

Hunt, L. M., Boots, B., & Kanaroglou, P. S. (2004). Spatial choice modelling: new opportunities to incorporate space into substitution patterns. *Progress in Human Geography*, *28*(6), 746–766. http://doi.org/10.1191/0309132504ph517oa

Jackson, J. B. (1958). The Abstract World of the Hot-Rodder. *Landscpae*, *7*(2), 140–151.

Jacobs, J. (1961). *The Death and Life of Great American Cities*.

Johnsen, S., May, J., & Cloke, P. (2008). Imag(in)ing "homeless places": using auto-photography to (re)examine the geographies of homelessness. *Area*, *40*(2), 194–207. http://doi.org/10.1111/j.1475-4762.2008.00801.x

Johnson, T. R. (2003). On the use of heterogeneous thresholds ordinal regression models to account for individual differences in response style. *Psychometrika*, *68*(4), 563–583. http://doi.org/10.1007/BF02295612

Jorgensen, B. S., & Stedman, R. C. (2001). Sense of Place as an Attitude: Lakeshore Owners Attitudes Toward their Properties. *Journal of Environmental Psychology*, *21*(3), 233–248. http://doi.org/10.1006/jevp.2001.0226

Jorgensen, B. S., & Stedman, R. C. (2006). A comparative analysis of predictors of sense of place dimensions: Attachment to, dependence on, and identification with lakeshore properties. *Journal of Environmental Management*, *79*(3), 316–327. http://doi.org/10.1016/j.jenvman.2005.08.003

Kuppam, A. R., & Pendyala, R. M. (2001). A structural equations analysis of commuters' activity and travel patterns. *Transportation*, *28*(1), 33–54. http://doi.org/10.1023/A:1005253813277

Lee, J.-H., Davis, A. W., & Goulias, K. G. (2015). Exploratory Analysis of Relationships Among Long-Distance Travel, Sense of Place, and Subjective Well-being of College Students. Presented at the 94th Annual Meeting of the Transportation Research Board, Washington, D.C. Retrieved from http://geog.ucsb.edu/geotrans/publications/LDTwithSOPandSWB%20_GEOTRANS2014-08-01.pdf

Lee, J.-H., Gao, S., & Goulias, K. G. (2015). Can Twitter data be used to validate travel demand models? Presented at the 14th International Conference on Travel Behavior Research, Windsor, UK. Retrieved from http://geog.ucsb.edu/geotrans/publications/Final%20Paper%20to%20Submit%20TweetOD%20Lee%20Gao%20Goulias_Geotrans%20Report_2015_05_03.pdf

Lewicka, M. (2008). Place attachment, place identity, and place memory: Restoring the forgotten city past. *Journal of Environmental Psychology*, *28*(3), 209–231. http://doi.org/10.1016/j.jenvp.2008.02.001

Ley, D. (2003). Artists, Aestheticisation and the Field of Gentrification. *Urban Studies*, *40*(12), 2527–2544. http://doi.org/10.1080/0042098032000136192

Low, S. M., & Altman, I. (1992). Place Attachment. In I. Altman & S. M. Low (Eds.), *Place Attachment* (pp. 1–12). Springer US. Retrieved from http://link.springer.com/chapter/10.1007/978-1-4684-8753-4_1

Massey, D. (1991, June). A Global Sense of Place. *Marxism Today*, 24–29.

Massey, D. (1995). Places and Their Pasts. *History Workshop Journal*, (39), 182–192.

McKenzie, G., & Janowicz, K. (2014). Coerced Geographic Information: The Not-so-voluntary Side of User-generated Geo-content. Presented at the Eighth International Conference on Geographic Information Science (GIScience '14), Vienna, Austria. Retrieved from http://grantmckenzie.com/academics/McKenzie_CoGI.pdf

Miller, H. J. (2005). A Measurement Theory for Time Geography. *Geographical Analysis*, *37*(1), 17–45. http://doi.org/10.1111/j.1538-4632.2005.00575.x

Mitchell, L., Frank, M. R., Harris, K. D., Dodds, P. S., & Danforth, C. M. (2013). The Geography of Happiness: Connecting Twitter Sentiment and Expression, Demographics, and Objective Characteristics of Place. *PLoS ONE*, *8*(5), e64417. http://doi.org/10.1371/journal.pone.0064417

Mokhtarian, P. L. (2005). Travel as a desired end, not just a means. *Transportation Research Part A: Policy and Practice*, *39*(2-3), 93–96.

Montello, D. R., Friedman, A., & Phillips, D. W. (2014). Vague cognitive regions in geography and geographic information science. *International Journal of Geographical Information Science*, *28*(9), 1802–1820. http://doi.org/10.1080/13658816.2014.900178

Mozingo, L. A. (2011). *Pastoral Capitalism: A History of Suburban Corporate Landscapes*. MIT Press.

N. Stokols, S. A. S. (1981). People in places: a transactional view of settings.

Paleti, R., Bhat, C., Pendyala, R., & Goulias, K. (2013). Modeling of Household Vehicle Type Choice Accommodating Spatial Dependence Effects. *Transportation Research Record: Journal of the Transportation Research Board*, *2343*, 86–94. http://doi.org/10.3141/2343-11

Pellegrini, P. A., & Fotheringham, A. S. (2002). Modelling spatial choice: a review and synthesis in a migration context. *Progress in Human Geography*, *26*(4), 487–510. http://doi.org/10.1191/0309132502ph382ra

Pender, B., Currie, G., Delbosc, A., & Shiwakoti, N. (2014). International Study of Current and Potential Social Media Applications in Unplanned Passenger Rail Disruptions. *Transportation Research Record: Journal of the Transportation Research Board*, *2419*, 118–127. http://doi.org/10.3141/2419-12

Proshansky, H. M. (1978). The City and Self-Identity. *Environment and Behavior*, *10*(2), 147–169. http://doi.org/10.1177/0013916578102002

Puleo, T. J. (2010). Baroque Disruptions in Val Di Noto, Sicily*. *Geographical Review*, *100*(4), 476–493. http://doi.org/10.1111/j.1931-0846.2010.00054.x

Puleo, T. J. (2013). Parasitizing landscape for UNESCO World Heritage. *Geoforum*, *45*, 337–345. http://doi.org/10.1016/j.geoforum.2012.11.024

Rapoport, A. (1977). *Human Aspects of Urban Form: Towards a Man—Environment Approach to Urban Form and Design*. Elsevier.

Ravualaparthy, S. K., & Goulias, K. G. (2014). Characterizing the Composition of Economic Activities in Central Locations: A Graph-Theoretic Approach to Urban Network Analysis. Presented at the 93rd Annual Meeting of the Transportation Research Board, Washington, D.C. Retrieved from http://geog.ucsb.edu/geotrans/publications/submitted%20version.pdf

Richard Walker. (2015). Retrieved June 13, 2015, from http://geography.berkeley.edu/people/emeriti/richard-walker/

Smaldone, D., Harris, C., & Sanyal, N. (2005). An exploration of place as a process: The case of Jackson Hole, WY. *Journal of Environmental Psychology*, *25*(4), 397–414. http://doi.org/10.1016/j.jenvp.2005.12.003

Solon, O. (2014, March 24). IBM can find you using non-geotagged tweets. Retrieved November 25, 2014, from http://www.wired.co.uk/news/archive/2014-03/24/twitter-location-algorithm

Sparks, T. (2010). Broke Not Broken: Rights, Privacy, and Homelessness in Seattle. *Urban Geography*, *31*(6), 842–862. http://doi.org/10.2747/0272-3638.31.6.842

Stedman, R. C. (2003). Is It Really Just a Social Construction?: The Contribution of the Physical Environment to Sense of Place. *Society & Natural Resources*, *16*(8), 671–685. http://doi.org/10.1080/08941920309189

Steinberg, D. Sustainable Communities and Climate Protection Act of 2008, California Health and Safety Code (2008). Retrieved from http://www.leginfo.ca.gov/pub/07-08/bill/sen/sb_0351-0400/sb_375_bill_20080930_chaptered.pdf

Tuan, Y.-F. (1974). *Topophilia: A Study of Environmental Perception, Attitudes, and Values*. New York: Columbia University Press.

Tuan, Y.-F. (1984). *Dominance and Affection: the Making of Pets*. Yale University Press. Retrieved from http://agris.fao.org/agris-search/search.do?recordID=US201300606563

Ukkusuri, S., Zhan, X., Sadri, A., & Ye, Q. (2014). Use of Social Media Data to Explore Crisis Informatics: Study of 2013 Oklahoma Tornado. *Transportation Research Record: Journal of the Transportation Research Board*, *2459*, 110–118. http://doi.org/10.3141/2459-13

Walker, R. (1989). A Requiem for Corporate Geography: New Directions in Industrial Organization, the Production of Place and the Uneven Development. *Geografiska Annaler. Series B, Human Geography*, *71*(1), 43–68. http://doi.org/10.2307/490506

Walker, R. (2009). *The country in the city: the greening of the San Francisco Bay Area*. University of Washington Press.

Wilson, C., & Groth, P. (Eds.). (2003). *Everyday America: Cultural Landscape Studies After J.B. Jackson*. University of California Press.

Zukin, S. (2011). Reconstructing the authenticity of place. *Theory and Society*, *40*(2), 161–165. http://doi.org/10.1007/s11186-010-9133-1