Understanding Geographic Bias in Crowd Systems

A Dissertation

SUBMITTED TO THE FACULTY OF THE

UNIVERSITY OF MINNESOTA

BY

Jacob Thebault-Spieker

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

OF

DOCTOR OF PHILOSOPHY

Brent Hecht and Loren Terveen

December, 2017

Acknowledgements

I would like to acknowledge the massive amount help and support I have received throughout my doctoral education from the people in my life, including:

- **My advisors** Loren Terveen and Brent Hecht, who have taught me so much and have been incredibly patient mentors. Thank you.
- **My partner** Johanna Desprez, who has stuck by me through the long nights, the times of incredibly stress, and at times across large distances. You are my foundation. Thank you.
- **My parents** John Spieker and Kim Thebault, and my **siblings** Micayla, Abbie, and Luke Thebault-Spieker. Your (at times snide) support has been critical. Thank you.

I would like to acknowledge my GroupLens labmates, without whom this process would have been a lot less fun. Y'all have provided an incredible number of laughs, and very valuable lunch table conversations. Thank you.

I would also like to remember John Riedl. His keynote talk at MICS 2010 was transformative for me. Had it not been for his presence at a regional conference and his willingness to engage with interested students, I would not have found GroupLens, nor my intellectual home – social computing and HCI. Thanks John.

I would like to acknowledge the other members of my thesis committee, Dan Keefe and Kate Starbird. Thank you for being game to participate, despite my irregular timeframe. Thank you.

Finally, I wish to acknowledge my other collaborators. Despite my use of the pronoun 'I' below, the research discussed here would not have been possible without my collaborators, their conversations, their work, and their support. Thank you.

This research was supported in part by NSF grants IIS-1707296, IIS-1218826, IIS-0808692, IIS-1707319, and a Microsoft FUSE Sharing Economy Research Award.

Abstract

Crowd platforms are increasingly geographic, from the sharing economy to peer production systems like OpenStreetMap. Unfortunately, this means that existing geographic advantages or disadvantages (e.g. by income, urbanness, or race) may also impact these crowd systems. This thesis focuses on two primary themes: (1) how these geographic advantages and disadvantages interact with crowd platform services, and (2) how people's geographic behavior within these platforms may lead to these biases being reflected. The first chapter in my thesis finds that sharing economy services fare less well in low-income, non-white, and more suburban areas. This chapter introduces the *spatial Durbin model* to the field of HCI, and shows that geographic factors like distance, socioeconomic status and demographics inform where sharing economy workers provide service. The second chapter in my thesis provides focuses on people in peer production communities contribute geographic content. By considering peer production as a *spatial interaction* process, this study finds that some kinds of content tend to be produced much more locally than others. Finally, my third contribution focuses on individual contributor behavior, and shows geographic "born, not made" trends. People tend to be consistent in the places, and *kinds of places* (urban, and non-high poverty counties) they contribute. The findings of this third study help identify mechanisms for how geographic biases may come about. Looking forward, my work helps inform an exciting agenda of future work, including building systems that provide individual crowd members sufficient geographic context to counteract these worrying geographic biases.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Many prominent online platforms depend heavily on the work of individual crowd members. Wikipedia is consistently in the top five most popular websites every month and OpenStreetMap data underpins some of the most prominent map technologies [181,182]. Neither platform could exist without crowd members' contributions. Uber is one of the largest sharing economy services, with forty million active monthly users, and a presence in over 663 countries. But without crowd members, no one would be available to provide on-demand rides. The citizen science platform eBird produces the largest biodiversity dataset of its kind, but this dataset could not exist at its current scale without citizen scientists crowd members.

Despite the apparent success of these crowd platforms, prior work has also shown that these platforms exhibit systemic biases in their service. One example of these systematic biases is the content 'gender gap'. Lam et al. [92] found disparities in content quality in Wikipedia articles about movies oriented toward women, versus those oriented toward men. Menking et al. [114] show that more generally, there is less content in Wikipedia about topics that are perceived to be of interest to women.

Just as the success of these platforms hinges on crowd members, these biases manifest based on who participates as crowd members and the behavior they exhibit. The simplest example of this was shown by Panciera et al. who found that a small group of people do an outsize proportion of the work [128] in these crowd platforms, and that this this behavior manifests as soon as they join. Others have shown that differences in the gender makeup of crowd workers leads to disparities in gender-oriented content [92]. With regard to crowd members' behavior, there is a misalignment between what Wikipedia readers seek out, and what crowd members work on (e.g. to increase the quality of content)[168]. Further, a recent study [63] identifies concrete ways in which individual autonomy and behavior are at odds with achieving globally standardized structured data.

Incorporating spatial context is another fundamental part of what makes these platforms successful. Sharing economy platforms like Uber and TaskRabbit promise on-demand rides between places or help with chores at a customer's home or office. In Wikipedia, geotagged articles are among the most popularly read articles on the platform. OpenStreetMap data could not exist without a concept of spatial context, and geotagged bird sightings in eBird have proven to be an invaluable resource for science. In some cases, the volunteered geographic information (VGI) [58] from these crowd platforms serves secondary purposes as well. For instance, algorithms are increasingly learning from VGI to understand the world [81], and scientists use VGI for scientific discovery in epidemiology [53] and earthquake prediction [145].

Platforms like Uber, OpenStreetMap, and eBird could not exist without this geospatial component, but this spatial context also paves the way for systemic *geographic* biases. For instance, Johnson et al. [80] found that most of the content in rural Wikipedia articles is generated automatically. Most of the content in Wikipedia about sub-Saharan Africa is written by westerners [150], and contributions by non-locals tend to be lower quality [38], or more superfical [80]. Haklay [61] found that low-socioeconomic status regions have lower quality content in OpenStreetMap.

While many forms of systemic biases exist, my thesis here focuses on one particular

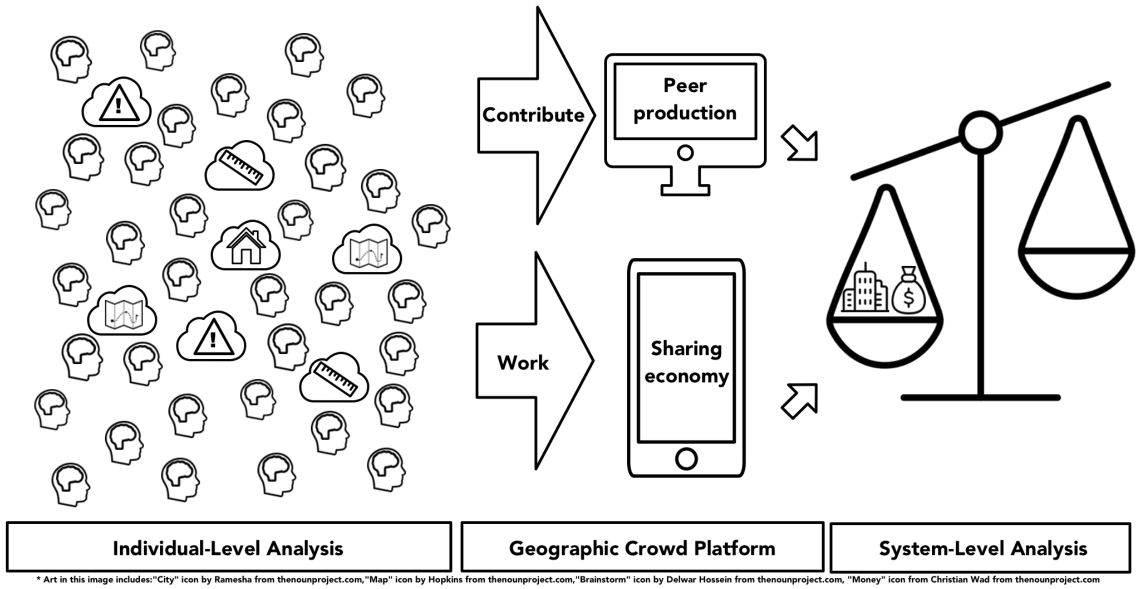

Figure 1.1: A diagram demonstrating the conceptual relationship between individual-level behavior and system-level geographic biases

dimension: understanding and mitigating systemic *geographic* bias. My thesis speaks to this topic by weaving together two distinct threads of prior work: identifying systemic geographic biases in crowd platforms, and understanding the impact of individual crowd member behavior. Holistically, my work shows that crowd members' geographic behavior impacts how well crowd platforms fare in low-income, non-white, and non-urban areas, and points to reasons for how, and perhaps why, these kinds of geographic biases come about.

Prior work in these previously distinct threads tends to follow two conceptual approaches, both of which I use here. First, as noted above, many studies take a 'system-level' view on identifying biases. Second, a different body of work focuses on 'individual-level' behavior that impact systemic results. Figure 1.1 illustrates the relationship between geographic behavior at the 'individual-level', and 'system-level' biases that manifest in these platforms. In Chapter 3 and Chapter 4, I take a 'system-level' view, and use robust geostatistical methods to measure biases, and identify some of the underlying geographic processes at play. In Chapter 5, I turn my focus to individual behavior, and describe how, and perhaps why, these biases occur over time. Throughout Chapters 3, 4 and 5 my work suggests that the principles from human geography (discussed in more detail in Chapter 2) manifest in the system-level biases of crowd platforms, and impact individual-level crowd member behavior. I now discuss each of these chapters in turn.

My first contribution, in Chapter 3, is a set of two studies on the availability of sharing economy service across different geographic and socioeconomic dimensions. I use robust, system-level geostatistical methods – and introduce the *spatial Durbin* model to HCI – to show that the the sharing economy disadvantages areas that tend to be low-income, non-white, and non-urban. Further, the evidence suggests that this disparity is due to geographic trends in who sharing economy crowd members are and the individual-level geographic decisions crowd members make. My findings suggest troubling implications for the sharing economy as it exists today, and identifies implications for the *geographic* design of these sociotechnical systems.

Chapter 4 represents the second contribution of my thesis. In this chapter, I study the effect of distance on contribution behavior, across hundreds of types of peer produced content (in eBird, Wikipedia, and OpenStreetMap). I formalize peer production as a *spatial interaction* process where contributions flow between regions. This study finds that (a) distance plays a different role depending on the type of content, and (b) that some types of content tend to be produced much more locally than others. Prior work has shown that locally-produced content tends to be richer and higher-quality than content produced by non-locals. My findings suggest implications for recruitment within these platforms.

Chapter 5 comprises the third contribution of my thesis, and focuses on *individual* geographic behavior of crowd members over time. Prior work [128] has found that contributors in Wikipedia tend to be 'born, not made' – that is, individuals begin participating in Wikipedia a certain way, and are consistent in that behavior over time. My research shows that individuals' geographic behavior tends to be 'born, not made' as well. Contributors in OpenStreetMap tend to be consistent in the places they contribute, as well as the *kinds* of places (non-rural and non-high poverty counties) they contribute. As noted above, prior work has shown that there tends to be more, and higher-quality OpenStreetMap content in urban and wealthier areas. My research points to how, and perhaps why, these geographic biases occur in crowd platforms like OpenStreetMap.

Taken together, these three chapters point to clear directions in the future. I discuss this research agenda in more detail later, but one particularly interesting direction is research towards mitigating these geographic disparities that occur. A common theme within my findings is that crowd members rely on their prior context when making decisions their geographic behavior. For instance, in Chapter 3, TaskRabbit crowd members defined entire sections of the city where they were unwilling to work, based on their generalized heuristics about safety. Further still, in Chapter 5, crowd members in OpenStreetMap were consistent in both the places, and kinds of places they contribute content. The concrete instantiations may be different for the sharing economy and peer production platforms, but one strong direction for future work is clear: developing tools to provide individual crowd members more, and better geographic context about underserved areas. I discuss this idea, among others, in more detail below.

Below, I first contextualize my work, and discuss how prior work informs my studies here. I then discuss each of the three contributions of this thesis in turn. Finally, I take stock of the findings of each of my there contributions, and discuss a set of holistic implications and directions of future research.

Chapter 2

Related Work

2.1 Principles from Human Geography

All of the work presented in this thesis is informed by geography, and specifically the large branch of geography known as human geography (c.f. [14]). My work here builds on four key principles from human geography: (1) residential clustering (i.e. the "Big Sort"), (2) structured geographic variations in population density, (3) distance decay, and (4) mental maps. It is likely that other principles from human geography also play a role in crowd platforms, but I focused on these four principles because (1) they have been found to figure into similar geographic processes (e.g. transportation geography or communication) and (2) they have been observed to play a role in other online social systems that have geographic footprints (see below). Below, I describe each of these four principles in more detail.

Residential clustering – in which people of similar characteristics reside close to one another – is a key property of the human geography of nearly all places around the world, and has been well-known in human geography and related fields for decades (e.g. [14,112]). For those familiar with the social networks literature, residential clustering can be understood as a type of "spatial homophily", and indeed this term has been used to describe similar phenomena (e.g. [113,174]). Within North America, residential clustering occurs along racial, ethnic, and socioeconomic lines, among other dimensions [18,112]. As I note below, clustering in some North American cities has become somewhat extreme, with tremendous socioeconomic (and other) gradients occurring across a metropolitan area. For instance, as of 2010, in the New York City metropolitan area, 78 percent of black residents would have to move to match the geographic distribution of white residents of the metropolitan area [45,46]. The same is true for 62 percent of those of Latino descent. Residential clustering in the United States (along with its concordant challenges) was the subject of the prominent book "The Big Sort" by Bill Bishop [178], which has led to the widespread use of the term "Big Sort" to refer to residential clustering. As such, I adopt this terminology in this thesis.

A longstanding subject of interest in the economic and urban geography communities has been understanding and modeling variations in population density across urban (and rural) areas (see [18] for an introduction and overview). These variations occur in structured – but diverse – patterns in cities and regions around the world. In North America, due to the character of local transportation networks, work/life behaviors and other factors, areas with very high population density tend to occur in city centers, which can have high socioeconomic and low socioeconomic status regions (as per the "Big Sort" phenomena). Outside city centers, density nearly always decreases, and in the suburban areas around cities (prior to entering rural areas), one usually finds low density, high-SES regions. It is important to note that outside of North America, population density patterns (and related SES patterns) can vary, resulting in different impacts on the sharing economy. While I briefly address implications for non-North American cities below, future work should seek to extend my research to the other metropolitan (and rural) structures that exist around the world. Brunn et al. [18] provides an overview of different metropolitan area structures that may be useful for this investigation.

The third major geographic principle in my work is distance decay, or the tendency for interaction between two places to decrease as the distance between the places increases [14,31,140,155]. Distance decay plays a role in a tremendous variety of human geographic process (and many processes from physical geography as well), with trade patterns [34] (trade declines with distance), transportation behavior [140] (destination choice is often largely defined by distance), and information dissemination [124,159,169] being some of the most well-known processes in which distance decay is a primary factor. Closely related to distance decay is the modeling of distance as a cost function in economic geography, leading to location-allocation problems [42] (e.g. what's the optimal place to put my Coca-Cola bottling plant given transportation costs of water, syrup, etc.). As I will show below, distance decay – when coupled with Big Sort processes along SES dimensions – suggest that distance may be an indirect agent of structural geographic bias in crowd platforms.

Finally, mental maps are, broadly speaking, the representations of the world that each individual has in their minds, both in terms of the geometry of the world and the attributes of those geometries. Work on mental maps dates back at least to Lynch's well-known 1960 book The Image of the City [104], and has continued for many decades, including prominent works by Gould and White [59] and Matei et al. [109]. Matei et al. focused on how communication infrastructures (e.g. mass media), coupled with "Big Sort" phenomena, has resulted in dramatically varying "comfort" levels across metropolitan areas. Namely, people from one type of area – e.g. high-SES areas that tend to be populated by people of certain races and ethnicities – feel unsafe and otherwise "uncomfortable" in other types of areas, and vice versa. Critically, the mental maps literature also points to a certain degree of ignorance associated with these comfort contours. That is, people tend to be less knowledgeable about places they feel less comfortable, both in terms of important attributes of these areas like crime rates [109], but also in terms of the geometries of these areas [91]. Mental maps and distance decay also have overlap, with knowledge about an area being in general inversely associated with the area's distance from one's home region [59]. my findings below point to mental maps – especially the associated knowledge and comfort factors – as playing a key role in crowd member behavior.

A key theme present in my human geographic principles is that socioeconomic status plays an important role: SES is one of the primary dimensions on which a "Big Sort" occurs in most cities around the world, SES and population density have important interplay (especially with respect to low-density suburbs), and people's mental maps and corresponding comfort and knowledge levels tend to vary across neighborhoods of differing SES [91,109]. As such, below, I adopt SES as a primary query mechanism with which to explore these geographic factors, using SES as an independent variable in both studies. As we will see, SES indeed sheds light on the impact of our human geographic properties in the sharing economy. I augment SES as an independent variable with other variables of interest, particularly targeting distance to capture a detailed picture of distance decay and population density to understand how its variation affects the sharing economy.

It is important to note that there are also other themes present in our geographic principles, most notably ethnicity and race, which are also important "Big Sort" dimensions and mental map determinants. Indeed, a number of factors in North America's history and present have led SES and race and ethnicity to be strongly correlated. In my studies below, in addition to using SES, I also discuss implications for race and ethnicity where I have sufficient data to support this analysis.

2.2 Geography in Social Computing Platforms

The broader body of literature studying geography in social computing platforms includes studies of geowikis, citizen science projects, and geotagged social media. This work tends to be focused on more traditional social computing topics like developing geographic platforms (e.g. [105,119,133]) and contributor behavior within these platforms (e.g.[129,162]).

By contrast, my work focuses on the *geographic* facets of these crowd platforms. That is, I focus on the geographic nature of these platforms, the geography of their participants, and the crowd members' geographic behavior. Below, I discuss contextualize my work within the body of research that focuses on the geographic aspects of these crowd platforms. Many of these geographic HCI [71] studies fall into three broad categories: (a) geographic variations by socioeconomic status and population density, (b) localness, and (c) geographic behavior of individual crowd members.

2.2.1 Variations by Socioeconomic Status, and Population Density

Prior research has shown that crowd platforms can have substantial geographic variations in the service they provide (e.g. [61,108,118,136,177]). One example of this is can be found in the research on volunteered geographic information (VGI) crowd platforms. A growing body of work shows that demographic factors are often associated with geographic variations in the quantity and quality of VGI contributions (e.g. [61,80,99]). Two demographic factors that are particularly linked to VGI content variations are socioeconomic status (SES) and the rural/urban divide. In short, low-SES and rural areas have been found to have fewer and lower-quality VGI contributions than wealthier and more urban areas [61,80,99]. For instance, Quattrone et al. show that more egalitarian (measured as a lower Power Distance) countries with higher incomes (GDP) have better geographic coverage in OpenStreetMap [136]. Similarly, Haklay finds that within Britain, the most disadvantaged areas (according to the Index of Deprivation, an aggregate metric of SES factors) tend to have worse coverage than those areas that are less disadvantaged [61]. Similarly, in Wikipedia, Johnson et al. [80] reported a similar trend concerning the rural/urban divide, observing that Wikipedia content about rural areas is often little more than bot-written template articles.

In a different domain, these geographic variations have been found in social media platforms as well. Li and Goodchild [99] found fewer tweets and photos submitted from low-SES regions of California. Similarly, researchers have found that found that people from rural areas produce less geotagged social media (e.g. posts to Twitter, Flickr, or Foursquare) per capita than their urban counterparts [70] and that this information is less likely to be produced by locals [82], and Johnson et al. identified that peer production crowdsourcing is less effective at describing urban areas than it is rural areas [80]. Even location-based games with social components (e.g. Pokémon GO) have been found to have similar coverage issues [24]. `As I discuss in more detail later`, it is likely that many of these findings can be attributed to the same geographic principles discussed above. Exploring this in more detail is an important direction of future work.

This research aligns with two geographic principles mentioned above: variations associated with residential clustering and population density. These trends suggest that these crowd platforms are *geographically biased*, and prior work has quantified these biases in many different social media and crowd platforms. However, little work has been done to understand the mechanisms behind these biases. My work below seeks to fill this gap in the literature by focusing on both system-level and individual-level geographic behavior within these crowd platforms.

2.2.2 Localness

'Localness' is another topic of particular interest in the "geographic HCI" literature, focused on content from both crowd platforms like OpenStreetMap and social media. This content has been considered a largely local phenomenon since the term "volunteered geographic information" was coined a decade ago [58]. As a result, the extensive and interdisciplinary literature in this space tends to presume that this VGI content is contributed by locals. Studies and systems that make a VGI "localness assumption" [82] range from studies of crowd members' "spatial footprints" (e.g. [101,122] to epidemiological analyses (e.g. [53]) to applications of sentiment analysis algorithms (e.g. [117]). Johnson et al. [82] provides a summary of applications of the localness assumption in VGI research and practice.

Implicit in this 'localness assumption' is a conception that the content producer resides in the region associated with the content (e.g. geotagged tweets). Put another way, assuming content is local implicitly assumes that the content is produced very near to where the producer lives. Under this 'localness assumption', the variations in content coverage discussed above would be directly attributable to the degree of local participation in these crowd platforms.

However, recent work has begun to call this localness assumption into question. This work has found that a substantial proportion of this VGI is in fact, non-local. For instance, Hecht and Gergle [69] observed that between 75 and 93% of edits to geotagged Wikipedia articles by anonymous (non-registered) editors were non-local, depending on the language edition (77% for English). Similar findings were observed by Hardy et al. [65], who modeled the relationship between IP-geolocated anonymous Wikipedia editors and the locations of their geotagged contributions (with an approach similar to my baselines, discussed in Chapter 4). Hecht and Gergle also observed that geotagged Flickr photos tended to be more local – although far from exclusively so – with around 50% of photos being taken by people outside their home region (100km). Sen et al. [150] found that geotagged Wikipedia articles about certain areas are significantly more local than others, with, for instance, articles about sub-Saharan Africa being written almost entirely by foreigners in most language editions. Finally, with respect to social media VGI (e.g. geotagged tweets), Johnson et al. [82] also observed a substantial degree of non-local contribution, averaging roughly around 25% depending on the definition of local.

The above work robustly establishes that the localness assumption in VGI is problematic. A substantial amount of VGI content is created outside of the producer's 'local' region, or 'from a distance'. The pervasiveness of non-local VGI content suggests that VGI content production is likely subject to another geographic principle discussed above: distance decay patterns. Put another way, geographic principles suggest that people are less likely to make contributions as the distance from a contributor increases. Indeed, some prior work has found this to be the case. Researchers have shown that Wikipedia contributions are subject to distance decay, with the likelihood of an editor contributing to an article about a place being a function of distance to the place [65,69]. On the social media side, Liben-Nowell established that roughly two-thirds of friends on Live Journal in 2004 could be attributed to a notion of geographic distance [100], and similar phenomena have been observed in other social networks [49,147,159].

Of particular note, one preliminary study suggests that the *rate* of this distance decay may differ by the type of content being produced. Hecht and Gergle [69] compared different 'spatial content production models' for generating volunteered geographic information [58] and found that Flickr contributions tended to be much closer to a contributors' 'home location' than was the case with Wikipedia. I present a much more detailed discussion of this topic in Chapter 4, where my work is first to systematically measure how these differences change between content types.

2.2.3 Geographic Behavior of Individual Crowd Members

Within the social computing literature, a number of different Finally, the geographic principle of mental maps, discussed in detail above, another body of literature in the social computing literature that The final geographic principle that manifests in social computing platforms, Much of the work in social computing spaces has focused on crowd member behavior, rather than formalizing this process as one stemming from mental maps.

Contributors' Geographic Patterns

The literature examining contributor geographic patterns falls broadly into two categories: the geographic ranges of contributors' work, and where contributors focus. The former largely overlaps with the body of research focused on 'localness', discussed above. With regard to the latter, I discuss this in more detail below.

Several different studies have sought to understand and characterize the geographic focus of contributors to peer production platforms. For instance, Panciera et al. [130] examined geographic trends in the Cyclopath platform, an early bicycling-centered community. In particular, they found that "Cyclopaths" (defined as the top 5% of contributors) had geographically constrained contribution regions, even within the relatively small area in which Cyclopath operated. Zielstra et al. [176] described the geographic extents of 13 OpenStreetMap contributors and show a method of characterizing which contributions are a part of a contributors' 'home location', and which are not. They found that the contribution ranges of these 13 people do not generally exceed approximately 50 square kilometers. Lieberman et al. [101] conducted a similar study, exploring the geographic extent of Wikipedia editors' contributions.

2.3 Temporal Evolutions of Individual Behavior

Whereas the work described above focused on geographic behavior, there has also been some research focused on the evolution of individual crowd member behavior over time. Much of this work is non-geographic in nature, despite occurring in VGI crowd platforms like Wikipedia, and OpenStreetMap.

In one of the seminal studies in this space, Priedhorsky et al. [134] took a temporal approach to understanding how value is created in Wikipedia and by whom. Panciera et al. [128] built on this paper with a study of 'Wikipedian' lifecycles and found that 'Wikipedians' (the term they use to describe those who contribute most of the Wikipedia content) begin contributing at a high level and maintain this trend over time, resulting in distinctive differences in contribution behavior between different classes of users. In other words, "Wikipedians are born, not made" [128].

Other work uses temporal evolution as a way to characterize the status of a geographic region (versus focusing on contributors and their behavior). One example of such a study is work by Gröchenig et al [60], who computationally estimated the 'completeness' of twelve urban areas, based on identifying three temporal stages ('start', 'growth', and 'saturation'), and modeling the development of a region through these stages.

More recently, others have begun to explore what roles contributors play in peer production communities, and how that changes over time. Arazy et al. [6] described 'career paths' of Wikipedia editors. Rehrl et al. [139] took a similar approach, and considered the different roles that people have in OpenStreetMap. Dittus et al. [35] explored the activation of newcomers and reactivation of previously dormant contributors during disaster events on Humanitarian OpenStreetMap (HOT).

My study in [Chapter #sec:chptbnm] is deeply informed by the work of Panciera et al. [128], and the studies mentioned above. Whereas prior work has focused on understanding geographic behavior or the temporal evolution of behavior, my study sits at the intersection. A spatiotemporal lens helps inform our understanding how contributors' geographic behavior evolves at the individual-level, and how these individual geographic behaviors may impact the geographic variations seen in OpenStreetMap.

Chapter 3

A Geographic Understanding of the Sharing Economy: System-Level Biases in TaskRabbit and UberX

In this chapter, I discuss two studies on different crowd platforms within the sharing economy, TaskRabbit and UberX. These studies make one clear conclusion: the sharing economy biases towards wealthier, more population dense regions. The first of these two (labeled Study 1 below) studies the decision making process of TaskRabbit crowd members – where are they available to provide service, and how much do they want to earn? To make sense of this individual-level decision process, I use geostatistical models to understand the system-level trends in this decision making process, to understand the geographic availability of TaskRabbit service. These findings are further informed by qualitative findings as to *why* individual TaskRabbit crowd members made the decisions they did. These responses point reasons rooted in the geographic principles discussed in Chapter 2.

The second of these two studies (Study 2, below) extends my TaskRabbit findings to UberX, a different sharing economy platform. For reasons discussed in more detail below, this study takes an exclusively system-level perspective on where UberX crowd members make their service available. To do so, this study introduces the HCI community to a robust geostatistical approach known as the *spatial Durbin model.* This modeling method provides important interpretive value, providing insight into how a locations' *surrounding geographic region* influences the spatial patterns described by the model.

3.1 Why the Sharing Economy?

While much attention has been paid to the economics and labor conditions of UberX, Airbnb, TaskRabbit and similar services (e.g. [33,77,120,138,144]), there has been much less focus on a topic that is critical in nearly all sharing economy platforms: geography. Regardless of whether we consider ride-hailing services (e.g. UberX, Lyft), peer-to-peer rental services (e.g. Airbnb), mobile crowdsourcing services (e.g. TaskRabbit), or even non-commercial sharing economy platforms (e.g. CouchSurfing), geography plays a key role. For instance, in the case of ride-hailing, a driver must travel from his or her current location to the location of the ride requester and drive the requester to a desired destination. For Airbnb, customers must decide where to stay, and prices are in part defined by the geographic context of each option. In TaskRabbit – a well-known platform that allows people to "outsource household errands and skilled tasks" [160] – a crowd member ("tasker") commutes to the task requester's location and/or to the locations involved with the specific errand.

The goal of this chapter is to better understand the role of geography in the sharing economy. I work towards this goal through two studies that provide evidence that the four principles from the field of human geography (established in Chapter 2) are key factors in the relative success of the sharing economy.

Through my consideration of these four principles, I demonstrate the critical importance of a geographic lens when examining the sharing economy, showing that this lens can reveal structural inequalities that would be otherwise invisible. Focusing on TaskRabbit and UberX in the Chicago region, I find that the four geographic principles lead to structural geographic biases in which the sharing economy is more effective in some types of areas than other types of areas. Namely, sharing economy platforms appear to succeed in areas with high socioeconomic status (SES) and population density and struggle in areas with low-SES and low population density. My evidence, for instance, shows that people in poor neighborhoods and outer-ring suburbs in the Chicago region wait longer for UberX cars and will have a harder time finding a TaskRabbit worker ("tasker") to complete a given errand.

Additionally, as in many parts of the world, population density and SES in my study region (the Chicago area) are correlated strongly with membership in certain protected classes, in particular those defined by race and ethnicity. This relationship results in an unfortunate corollary to my findings: in some cases, the sharing economy appears to be most effective in areas with fewer minorities and much less effective in black and Latino neighborhoods.

The work in this chapter triangulates my high-level findings across multiple, system-level, methodological approaches. These approaches include controlled experiments, the analysis of qualitative survey responses, and (to my knowledge) the first use of an advanced geostatistical technique known as spatial Durbin modeling in the human-computer interaction literature. Spatial Durbin models are emerging as a best practice in the social and natural sciences, and are an approach that I believe can be broadly useful for studying the sharing economy and in the growing "geographic human-computer interaction literature" [71] more generally.

While this chapter focuses on the descriptive analysis of the geography of sharing economy platforms, my studies also provide evidence for potential solutions to the challenges I identify. For instance, I observe that very few TaskRabbit crowd members live in low-SES neighborhoods and that large crowd member-to-job distances contribute to the higher prices and decreased willingness to accept jobs in these neighborhoods (a manifestation of distance decay). Since in TaskRabbit, UberX, and most other sharing economy services, low-SES individuals have a harder time satisfying crowd member enrollment requirements (e.g. crowd members must have a bank account in many cases), this likely reduces crowd member participation in low-SES neighborhoods and thereby diminishes the effectiveness of the overall platforms in these neighborhoods. Below, I discuss how my results – in combination with a geographic perspective on the sharing economy – suggest that removing or relaxing some of these crowd member requirements may address some of the problems raised by my findings.

In summary, this chapter makes the following contributions to the literature on the sharing economy:

1. I present a broad examination of the role geography plays in the sharing economy and solidify the importance of a geographic perspective in the sharing economy literature. In particular, my results point to the influence of four key principles from human geography: "Big Sort" residential clustering, geographic variation in population density, distance decay, and mental maps.
2. I present evidence that the interaction between common sharing economy platform design decisions and these four geographic principles lead to structural geographic biases in the sharing economy, biases that reinforce existing advantages. Specifically, my results suggest that high-population density, high-income neighborhoods receive the largest benefits from the sharing economy and poor urban neighborhoods and outer-ring suburbs receive fewer benefits.

3. I find evidence that, due to the pervasive correlation between poverty and race/ethnicity in the United States and many other parts of the world, in many cases, black and Latino neighborhoods tend to be less well-served by the sharing economy.
4. I discuss the design implications of my research, including evidence from my studies that points to means by which the benefits of the sharing economy may be more widely distributed.
5. Finally, this work makes a lower-level, methodological contribution: this chapter introduces spatial Durbin models to the human-computer interaction literature and discusses why spatial Durbin modeling is important for robustly understanding many sharing economy geospatial processes (and many geographic HCI processes more generally).

Below, I describe in detail prior work that is focused on the sharing economy. I then present my methods and results for my TaskRabbit study. Next, I describe my geographic analysis of UberX, introduce spatial Durbin modeling, and show how Durbin modeling can robustly identify structural geographic biases in UberX wait times. I conclude with a summative discussion that cuts across both studies and provide an overview of design implications for the sharing economy.

3.2 Sharing Economy Related Work

Sharing economy platforms (e.g. TaskRabbit, Uber, and Airbnb) have become a subject of intense public discussion, which has led to increased attention from researchers (e.g. [33,39,40,51,76,77,97,105,163]). Initial work has focused on addressing non-spatial issues, usually involving the adaptation of research questions from the virtual crowdwork literature (e.g. [10,11,78,87]) to the sharing economy context. For example, Teodoro et al. [162] conducted a qualitative study to investigate the motivations of workers in TaskRabbit and Gigwalk (a platform broadly similar to TaskRabbit but with different primary use cases). They found that monetary compensation and control of working conditions (time of day, rate of pay, the tasks they do) were primary factors in workers' motivation to participate in these platforms as crowd members. Alt et al. [2] independently developed an experimental system similar to TaskRabbit. They asked people to complete tasks using a smartphone and observed their behavior. They found that workers were more willing to do tasks that were, for example, relatively straightforward (e.g., taking photos) and that could be done before and after business hours. Ikkala and Lampinen [77] explored the social role that payment plays in an Airbnb study, and discuss how payment modifies the social relationship between hosts and guests.

One thread of sharing economy work has focused on non-commercial "peer-to-peer exchange" platforms (a term used instead of "sharing economy", which some have problematized [148,149]). One prominent example in the HCI community is the body of work on timebanking, or time-based currency made possible through technological support (e.g. [8,151]). A thread of work related to these non-commercial systems focuses on the social dimension of the commercial sharing economy. This thread suggests that even in commercial systems, there is a social 'economy' between the worker and the person receiving service, and that this dimension is a critical attribute of what people like about and expect from these systems [73]. While my work here focuses on large, commercial sharing economy platforms, examining the role of geography in non-commercial peer-to-peer-exchange platforms is an important direction of future work.

Another major focus of sharing economy research has involved examining the challenges associated with being a sharing economy crowd member. Lee et al. [97] found that tensions arise between supervisory task assignment algorithms and crowd members in ride-hailing systems like UberX and Lyft. Rosenblat and Stark [144] consider the power structures that arise from the reputation system in UberX, and what effect his has on drivers. In a similar vein, Raval and Dourish [138] find that part of what sets "good" drivers apart in UberX is the emotional labor they carry out, and argue for the importance of recognizing this labor. Glöss et al. [57] compare the differences in work and perspectives between taxi and Uber drivers. Ahmed et al [@ SyedIshtiaqueAhmedPeertopeerworkplaceview2016] consider a very similar juxtaposition to Glöss et al., but focus on a different, international context: the Ola ride-hailing platform in India. Ola connects passengers to rickshaw rides, similar to UberX. Ahmed et al. explore the differences between auto-rickshaw drivers who do not use a sharing economy platform, and those who do.

Recent work has examined the relationship between demographics and worker participation in sharing economy platforms, and this research played a major role in informing the research present in this chapter. For instance, Lee et al. [97] found that UberX drivers often turned off 'driver mode' in the Uber app when they are near areas where they feel unsafe or avoided unsafe areas entirely, a finding that provides key context to some of my results below. Dillahunt and Malone [33] identified that there are barriers to participation in the sharing economy for people who live in low-SES areas. Dillahunt and Malone's work, in particular, provides important scaffolding for my design implications as I discuss below. Along the same lines, Edelman and Luca [39] found that black Airbnb hosts systemically earn less than non-black hosts, and that users with stereotypically African American names Airbnb are less likely to be accepted as guests compared to identical profiles with stereotypically white names [40]. A similar preliminary set of results was recently identified in UberX in Seattle and Boston by Ge et al. [52] with respect to wait times and cancellations.

My research is most directly motivated by recent work on the sharing economy that has begun to identify geographic phenomena as potential factors of interest. For instance, Teodoro et al. and Alt et al. (as well as others) observed that how far people would need to travel to a task appears to influence their attitude toward the task. This is a finding that I both replicate and formalize in a controlled experiment on TaskRabbit, identifying this phenomenon as a manifestation of distance decay. Through modeling and qualitative analysis, I are also able to provide the first evidence that distance decay in the sharing economy – coupled with "Big Sort" residential self-selection – has substantial effects on the availability of sharing economy services and the price of these services (and, subsequently on the bias in the geographic effectiveness of these services).

Similarly, Quattrone et al. [137] explore the geographic and demographic factors that contribute to Airbnb growth and penetration in London. Through this study, they make policy recommendations based on their findings that would allow regulators to be more responsive to the changing attributes of Airbnb. Expanding on Diakopoulos's past work on UberX surge pricing [32] (variable prices based on demand) in a recent blog post for the The Washington Post, Stark and Diakopoulos [153] describe initial explorations of UberX availability (with a particular focus on surge pricing) with regard to socioeconomic and demographic attributes of the Washington, D.C. area. Hughes and MacKenzie [76] performed a similar analysis in Seattle. Among other extensions of this work (see below), I am also able to replicate and formalize the findings from both of these studies using a robust statistical framework (spatial Durbin modeling) that can provide new insight given the spatial properties of relevant data.

Critically, I also show how these principles interact with design decisions in sharing economy platforms to create important structural geographic biases that disadvantage people living in low-density and poor areas. The robustness of these contributions is supported by this work being the first to examine multiple sharing economy platforms (and multiple types of platforms) with a geographic lens and by my adoption of a new statistical framework from the domain of spatial statistics (spatial Durbin modeling). I believe this framework will prove useful for other researchers in future examinations of the sharing economy. Overall, this chapter, supported by the previous work in this space, paints a clear two-part picture: (1) when it comes to the sharing economy, geography matters and (2) one way it matters is that key human geography principles interact with sharing economy design decisions to create structural geographic biases.

3.3 Study 1: TaskRabbit (Mobile Crowdsourcing Sharing Economy Platform)

I begin the discussion of my empirical work with my analysis of TaskRabbit. TaskRabbit is a canonical example of the mobile crowdsourcing branch of the sharing economy. As noted above, it is used by task requesters for the completion of physically-situated tasks such as delivering flowers, building IKEA furniture, and helping task posters move large items [157].

The goal of my TaskRabbit study [163] was to understand the effectiveness of the TaskRabbit platform in different regions with respect to our four geographic principles. To address this goal, I first had to define effectiveness within a TaskRabbit context. I did so by decomposing the notion of TaskRabbit effectiveness into two basic dimensions: (a) the ability to find a crowd member to complete a task (i.e. the willingness of a worker to do a task) and, if a crowd member is willing to do the task, (b) the price at which the crowd member will complete the task. These dimensions led directly to my two research questions for this study, each of which I explicate in turn immediately below.

Research Question #1: RQ-Willingness: Where will participants in TaskRabbit be willing to go to complete tasks?

As noted above, SES plays a key role in three of our human geographic principles: the "Big Sort" (i.e. people of similar SES cluster together), population density variation (i.e. some parts of metropolitan areas like the suburbs tend to be wealthier than others), and mental maps (e.g. people who live in higher SES areas tend to know less about low-SES areas and have low comfort levels in these areas). As such, in this study and in the UberX study below, I used SES as a straightforward probe into the function of these principles in the sharing economy. This was a decision that turned out to be supported in my results (see below). In particular, the human geography literature on our principles suggests that workers would be less willing to complete tasks in low-SES areas, which amounted to my first hypothesis: *H-Willingness-SES*.

The one principle that is not directly addressed through an investigation of SES in this context is distance decay. As such, I also included distance to a task (from a worker's frequently visited areas) as an independent variable. Distance decay suggests that as this distance increases, willingness to complete a task should go down: *H-Willingness-Distance*.

Research Question #2: RQ-Price: How does geography affect how much participants in TaskRabbit request in payment?

At the time of data collection, TaskRabbit had a straightforward auction system for tasks in which workers would bid on tasks posted by users. As such, I believed that the amount workers would charge for a task would be subject to similar processes as their willingness to do the task. Specifically, I hypothesized that distance and task price would be positively correlated (indeed, cost can be a primary mechanism for distance decay, e.g. [23]) (*H-Price-Distance*) and SES and price would be inversely correlated (*H-Price-SES*).

It is important to note that TaskRabbit's pricing model is subject to frequent iteration, as is the case with many aspects of most sharing economy platforms. As of this writing, TaskRabbit's model has changed to a more complex approach that involves several options for workers and requesters. However, the model still incorporates relatively significant user input in some cases, making it more liable to principles from human geography, something that is in theory not the case with UberX (although driver behavior with respect to surge pricing problematizes this notion [32]). I highlight the role of pricing model design, the differences between UberX and TaskRabbit with respect to pricing, the relationship between willingness and pricing, and innovation in this area in my Discussion section below.

3.3.1 TaskRabbit study design

To address the above research questions and evaluate the corresponding hypotheses, I developed an experiment and recruited TaskRabbit workers as participants. This recruitment was done in an organic fashion by posting tasks to TaskRabbit's Chicago metropolitan area site just as a typical task requester would post a task. Only TaskRabbit workers local to the Chicago area could participate in my experiment, for which I paid participants $5 in 15-minute intervals, capped at an hour (e.g. a

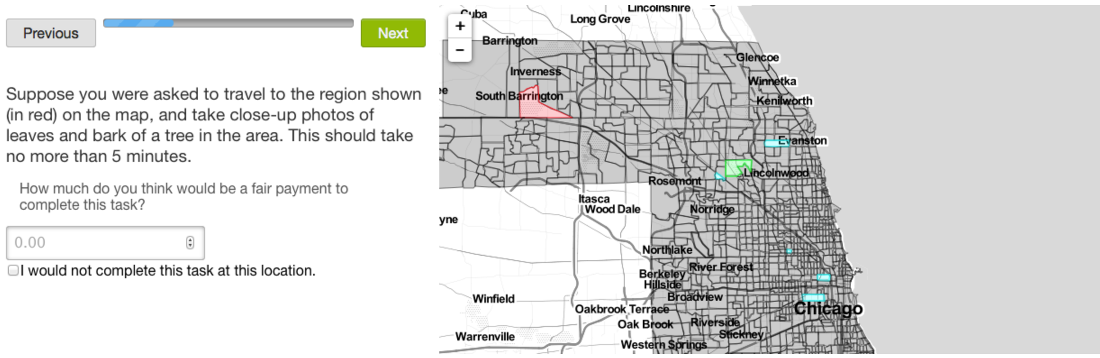

Figure 3.1: An example of what participants saw in my TaskRabbit experiment. The green census tract is the worker's self-reported home tract and the blue tracts are those that the worker reported visiting at least once a month. The red tract is the tract about which the worker is currently being questioned (Note: the image is cropped for space and, for privacy reasons, the figure does not depict an actual worker's responses).

person who took more than 15 minutes but less than 30 would receive $10).

To add context to my results, I first asked participants a number of questions about themselves, covering topics such as gender, preferred mode of transportation, and activity level on TaskRabbit. I also asked participants to select their home census tract on a map I provided, and to do the same for census tracts that they visited at least once a month.

After participants answered questions about themselves, they began the main portion of the experiment. This portion of the experiment involved prompting participants with census tracts in Cook County, Illinois, which contains Chicago and many of its suburbs . For each census tract, the participant had the option to either check a box labeled "I would not do this task at this location" (RQ-Willingness) or to name what they felt would be a fair price to complete the task (RQ-Price). I did not ask TaskRabbit workers to complete the tasks, only to say if they would complete them and at what price. Figure 3.1 shows an example of the experiment interface.

Table 3.1: Experiment Tasks and Their Hypothesized Engagement Level

Task	Engagement Level
Task 0: Suppose you were asked to travel to an intersection in the region shown (in red) on the map, and photograph all of the signs at that intersection. This should take no more than 5 minutes.	Low
Task 1: Suppose you were asked to travel to the region shown (in red) on the map, and take close-up photos of leaves and bark of a tree in the area. This should take no more than 5 minutes	Medium
Task 2: Suppose you were asked to travel to the region shown (in red) on the map, visit someone's home, and ask the owners to respond to a single question about local politics. This should take no more than 5 minutes	High

Each tract was randomly assigned one of three hypothetical tasks designed to vary the level of engagement with the local area (Table 3.1), an important variable considering the mental maps literature and its findings relating to geographically-variable comfort levels (e.g. [109]). These tasks ranged along an engagement spectrum from a task that could be done without leaving a vehicle to a task that required interaction with a person in the area. Tasks were designed so that they would not take more than five minutes.

Each participant received 20 census tracts. Fourteen of the tracts were randomly selected (without replacement), and these tracts were the tracts that were considered in my quantitative analysis below. In order to enrich our qualitative understanding of the geography of mobile crowdsourcing markets, I also considered four special-case tracts: the highest-income and the lowest-income tracts in my study area and, in according with Matei et al.'s work on mental maps [109], the highest-crime and the lowest-crime tracts. The remaining two tracts were repeated from the randomly chosen set of 14 tracts to verify intra-rater reliability. Each repeated tract was presented no fewer than 5 tracts after the original.

Upon seeing and responding to all 20 of the tracts with either a price or by stating that they would not complete the task, I asked participants several open-ended questions whose answers were entered into text boxes. Specifically, I asked participants about how they made their pricing decisions and why they would not complete certain tasks (if they checked that box at least once).

3.3.2 TaskRabbit Results

Forty participants completed the experiment (20% of active TaskRabbit workers in Chicago at the time of the study), which I ran during Spring 2014. 57.5% of participants identified as women (42.5% men), which aligns well with gender distribution in the platform overall [160]. The median participant performed a task on TaskRabbit between once a week and once every two weeks. 30% of participants indicated that they complete multiple tasks per week, while only 20% of participants indicated that they complete a task once a month or less.

RQ-Willingness.

Because price is irrelevant if a worker will not complete a task, I first sought to understand the geography of worker willingness. To do so, I built a logistic mixed effects model with three fixed effects:

- Distance to task from the closest census tract visited by the participant at least once a month (as indicated in the experiment) [This helped us understand the role of distance decay].
- Median household income of the task tract, as an indicator of socioeconomic status (e.g. [99,154]). As noted above, many other socioeconomic variables are well known to be correlated with income (e.g. educational attainment, occupation). To reduce the effect of the long-tailed distribution of wealth, I log-transformed this variable. Median household income data was gathered from the United States Census' American Community Survey 2006-2010 dataset. [This helped us see the effect of the "Big Sort", population density effects, and mental maps]
- Task ID, to make sure I understand the effect of distance and median income in the context of a given task.

The model's random effects were intercepts for participant and by-participant slopes for the effects of income and distance. The model's dependent variable was whether the participant had checked the "I would not do this task at this location" box for a given tract. It is important to note that this model used standard mixed effects techniques rather than my more advanced spatial Durbin modeling approach, which is employed in my analysis of UberX data. I discuss the relationship between these two models and their appropriateness for each setting in Section 3.4.

To operationalize distance, I used travel time rather than Euclidean distance to better match actual mobility in Chicago. I used the Google Distance API to calculate the off-peak travel time between the centroid of the task tract and the centroid of the nearest tract (to the task tract) that the participant indicated visiting frequently (more than once a month). The API supports multiple transportation modes, and I calculated travel time with participants' self-reported preferred mode of transportation.

Overall, participants indicated that they would not do 34% of the tasks. The few census tracts that had a reported median income of zero (e.g. the tract that consists of O'Hare International Airport and a few hotels) were excluded from further analysis.

Table 3.2: The Results of my *Willingness* Model

Fixed Effect	**Estimate**	**p-value**
Travel time (in hours)	-3.15 (0.99)	0.001
log_2[Task tract income in $10]	0.87 (0.36)	0.014
Task ID (baseline = Task 0)	1: 0.37 (0.40), 2: -0.92 (0.40)	0.003
Constant	1.81 (0.82)	0.028

Table 3.2 shows the results of my model. All fixed effects are significant and I find that both *H-Willingness-Distance* and *H-Willingness-SES* are supported. Socioeconomic status of the task location and distance to the tract both have an effect on whether a worker is willing to complete a task, with SES having a significant positive relationship and distance having a significant negative one.

The effect sizes are relatively large. According to the model, for every doubling of task area median income, there is a 2.38x increase in likelihood that a worker

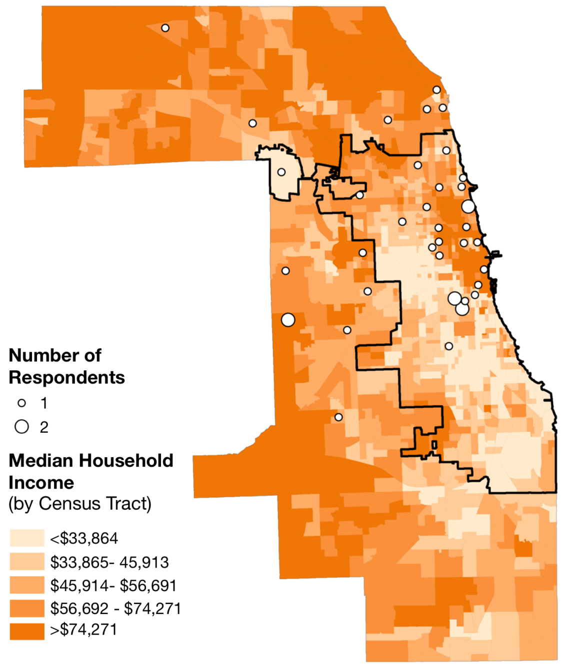

Figure 3.2: Experiment participants' self-reported home census tracts and median income in Cook County, Illinois. Very few participants live in low-income tracts. Note that the low-SES "South Side" of Chicago (Chicago is outlined in black) has only one participant, and no participants live in the poorest parts of the South Side. Median income color classes are determined via the quantile method, meaning each class represents a quintile of the household income dataset. Participant are displayed at the centroid of their home census tract.

will accept a task. In other words, holding the other variables constant, my model suggests that the likelihood of a worker accepting a task will more than double if the task is in a tract with a median income of, for instance, \$60K rather than a tract with a median income of \$30K. As shown in Figure 3.2, \$60K is a relatively standard median household income in northern Chicago and the Chicago suburbs, with \$30K median household incomes common on the "South Side" (as the southern part of Chicago is commonly known).

With respect to travel time, my model indicates that for every hour of travel time there is a substantial decrease in willingness to complete a task. In this case, the geographic interpretation is clear: this result directly validates a rather large presence of distance decay. Specifically, TaskRabbit workers are about 4.3% as likely to complete a task an hour away than they are tasks in their immediate vicinity.

Examining my willingness results in more detail, I found an interesting result with regard to gender. While 78% of women said they would not complete at least one task, the equivalent number for men was 53%. In addition, the grand mean willingness (mean of the means for each participant) for women was 57.1% but for men it was 77.7%. my qualitative results below suggest that both distance and crime factors play a role in willingness decisions by women (in part mediated by mental maps), but these are the same factors also indicated by men. Although further research is needed, it is likely that women have a lower threshold for one or both of these factors.

RQ-Price.

I now turn my attention to my analysis of the price participants indicated that they would charge for a task (assuming they were willing to complete the task). I began this analysis by ensuring that it had sufficiently high intra-rater reliability. I did so by calculating the Pearson's correlation coefficient between the first and second price judgments for the repeated tracts. The coefficient was r =0.96 across all participants, indicating that participants' pricing decisions were very consistent. To understand the effect geography has on task prices in TaskRabbit, we built a linear mixed effects model with identical independent variables as my willingness model but with reported task price as the dependent variable.

Table 3.3: The Results of my *Price* Model

Fixed Effect	**Estimate**	**p-value**
Travel time (in hours)	10.10 (2.27)	< 0.001
log_2[Task tract income in $10]	0.40 (0.52)	*n.s.*
Task ID (baseline = Task 0)	1: -1.73 (0.85), 2: 0.28 (0.87)	0.024
Constant	16.92 (2.90)	< 0.001

The results of this price model can be seen in Table 3.3. This table reveals that travel time was positively associated with price, supporting *H-Price-Distance* and the distance-as-cost-function view of distance decay. Indeed, the model suggests that for every hour of travel time, the price goes up at a rate of $9.97/hour. Task tract income, on the other hand, was not significant; the median household income of the tract does not have a significant effect on price. In other words, *H-Price-SES* was not supported.

This, however, is where the important role of "Big Sort" phenomena becomes clear: due to these phenomena, even though SES is not a significant predictor of price, I found that people who live in large low-income areas are indeed likely to be charged more for the same task. To understand how this works, consider Figure 3.2, which shows the self-reported home tracts of all 40 participants on top of a map of income by census tract in Cook County. Immediately visible in Figure 3.2 is that very few participants live in the heart of low-income areas. Indeed, most participants seem to live in middle-income areas next to the very high-income portions of northern Chicago (the "North Side"). Only a single participant lives well within the lower-income South Side of Chicago. As a result, low-income residents on the South Side are almost always a large distance away from any given TaskRabbit worker, making distance an agent of higher prices for these low-income neighborhoods. In other words, a low-income resident of the South Side would have to pay more to receive a given TaskRabbit service, for instance someone to take care of errands to make time for longer-term goals [15,167]. Moreover, as per my findings above, South Side residents also likely have a harder time finding a TaskRabbit worker to accept a request for services in the first place.

This result suggests a specific character for the effect of the Big Sort on the sharing economy. I found that most sharing economy workers live near (but not in) high-SES regions. If this result generalizes, people who live in large, low-income districts like the South Side will need workers to travel greater distances to get to their task locations, resulting in longer travel times, and, ultimately, higher prices. Where low-SES pockets are much smaller (e.g. the lower income pockets in the suburbs just north of Chicago), the effect on travel time, and therefore price, will be more minimal. However, these smaller pockets may get rarer and rarer in a "Big Sort" world.

It is also important to point out that these "Big Sort" effects also play a role in willingness decisions. Since distance and willingness were found to be inversely associated, the fact that TaskRabbit workers live far away from large low-SES areas means that this inverse association will disproportionately affect people who live in these areas. Moreover, since I identified a separate effect in which willingness and SES are positively associated (as SES goes up, willingness goes up), the distance effects and SES effects likely compound each other to make the task willingness in large low-SES areas particularly low.

Lastly, although my consideration of population density largely lies in my UberX study, we do see an important effect for population density here. Figure 3.2 shows that workers are concentrated in the high-density city of Chicago rather than the low-density suburbs. As such, distance not only reduces availability of TaskRabbit and increases the price of TaskRabbit in lower SES areas, but does the same in suburbs. However, the situation in suburbs is generally quite different: as can be seen in Figure 3.2 (and is discussed in related work), suburban people have higher incomes than people on the South Side of Chicago, and thus they can potentially afford the increased costs. In addition, in some cases, even people in somewhat remote suburbs are closer to one of my participants than a person in southern Chicago. That said, while United States suburbs tend to be relatively wealthy, the opposite is true in many cities around the world (e.g. France and Latin America [14,18]). Where this is the case, services like TaskRabbit will likely be drastically more expensive and less available in these areas. As I note above, investigating these phenomena in cities with different socioeconomic segregation patterns is an important direction of future work.

3.3.3 TaskRabbit Qualitative Results

Thus far, my quantitative models have revealed evidence for the importance of distance decay, "Big Sort" phenomenon and, to a lesser extent, structured variations in population density, when considering the effectiveness of TaskRabbit. I have also seen these principles manifest in structural geographic biases in TaskRabbit, biases that lead TaskRabbit to be both more expensive and less available in low-SES regions in my study area. I now turn to my qualitative results to attempt to help understand why these dynamics exist. To do so, a single investigator looked for themes in the textual survey responses, focusing on ideas related to my four geographic principles. Below, I outline the results of this analysis, which identified qualitative data relevant to mental maps (and their interactions with "Big Sort" phenomenon) and distance decay.

Mental Maps.

A theme that was very clear in my participants' answers to why they ticked the "will not do [a task]" box is the importance of mental maps, and specifically a large region of low comfort levels in their mental maps corresponding to Chicago's South Side (and to a certain degree the "West Side"). Participants reported that these low comfort levels were driven mostly by perceptions of high crime. Indeed, some participants' responses read as if they came directly out of Matei et al.'s study that examined the role of crime in neighborhood-level comfort assessments in mental maps. For example, consider this response from P27:

> "I think the high incidence of gang-related crime makes many Chicagoans too nervous to visit some parts of the city. We always refer to Chicago as being a"city of neighborhoods" but the truth is that many Chicagoans feel uncomfortable visiting a huge portion of our city. The nature of the crimes that occur on the South and West Sides (gang-related) makes me particularly nervous because there's nothing you can do to prepare/protect yourself. I realize that I might have some biases but it's less about location for me and more about crime rate. I do wish Chicagoans (and visitors) could feel more comfortable exploring and enjoying more neighborhoods without worrying about crime." (P27)

P9 is a member of the TaskRabbit Elite. This is a designation one can earn within TaskRabbit after earning an average rating of 4.9 stars (ratings are given by task requesters upon completion), completing a large number of tasks, and not violating any of TaskRabbit's policies. P9 offered similar feedback to P27:

> "I am an Elite member of TaskRabbit and I do a lot of tasks. I do not do tasks anything below the loop of Chicago [i.e. the South Side] so it has to be on the north side for me to work. It is purely for safety concerns." (P9)

P4, a relatively new resident of Chicago, wrote that the comfort level overlay in her mental map also led to similar decisions about whether or not to accept a task. In this case both poverty and crime are mentioned:

> "I only moved to Chicago last May, so I don't know much about the city except that there are large pockets of poverty, inequality and high crime. In terms of general areas of the city I understand that large swaths of the south side and west side include these pockets of poverty and high crime. Without specifics about which neighborhoods/blocks/streets are safe I essentially ruled out anything on the south or west side of the city. For the most part, I think the western suburbs are safe but I know nothing about the southern suburbs so I erred on the side of safety and avoided those areas as well." (P4)

P39 specifically addressed her gender as part of the reason she did not consider certain tasks, saying:

> "I wouldn't feel safe in some areas as a female by herself." (P39)

P16 was very explicit about how he makes decisions between the contradictory signals from his comfort layer and the desire to increase his income:

> "Whether or not my assumptions of lack of safety were correct, I wouldn't put myself in danger for a few dollars" (P16)

The quotes above make it clear that a key sharing economy decision-making process – whether or not a crowd member agrees to accept a task – is subject to classic mental map effects. These are effects that have been observed in geography and related fields for decades (e.g. [59,109]). As has been observed by Gould and White [59], Matei et al.[109], and others, humans tend to ascribe large regions of their mental maps with positive and negative emotions, with Matei et al. [109] specifically focusing on comfort levels assigned to neighborhoods in mental maps as a function of perceived crime in those areas. My qualitative results suggest that this is a primary driver behind the results of my willingness model.

These findings also dovetail with recent findings by Lee et al., who, as noted above, did qualitative work with UberX drivers. Lee et al. identified that UberX drivers often manually disable their availability to the UberX platform when they are traveling through what they perceive to be unsafe neighborhoods, a finding that can be easily understood through a mental maps lens. This is also the direct UberX analogy to a TaskRabbit worker not accepting tasks in specific neighborhoods, and is a point to which I return in my UberX study.

Another theme in the above responses that is also present in Gould and White's and Matei et al.'s results is a lack of geographic nuance in mental maps. While the South and West sides do indeed experience much higher levels of crime than other parts of Chicago, there are pockets of these areas that are quite safe [165]. However, “Big Sort” processes have driven most people of higher SES out of the South Side and the West Side (as well as many people of White, non-Latino descent; see below), likely leading to large regions that are unknown to people who do not live there, both in terms of geometry and personal comfort levels. This is roughly analogous to a finding observed in mental maps of another major urban city, Boston [91]. The mental maps of my participants clearly are not sufficiently nuanced to support knowledge of the lower crime pockets on the South and West side.

Before moving on to my analysis of the presence of distance decay themes in my qualitative results, it is important to point out that prior work (e.g. [109]) suggests that, even though SES and crime are the only two attributes directly cited by my participants in their willingness decisions, race and ethnicity may also be involved. This is a point I address in my discussion below.

Distance Decay.

Participants’ qualitative feedback supports the finding from my quantitative modeling exercise that proximity of the task location is a very important factor in task willingness and pricing decisions. Here, P4 explicitly discusses the role of distance decay in her pricing decisions:

> “Mostly how much of a pain it was going to be to get there. If it was a place I could stop by on my way to or from work or the gym= cheap. If it required getting in my car=more. If it required an extensive drive to a far flung suburb=more.” (P4)

My qualitative data also shed some of light on the specific form of distance decay in TaskRabbit pricing and willingness decisions. Distance decay can take many and multiple forms (e.g. gravity models (e.g. [44,140,155]), thresholds (e.g. [1]), and my qualitative results suggest that distance decay in this context contains a threshold component. Four participants explicitly or implicitly mentioned thresholds in explaining why they said they would not complete a specific task:

> "The distance was too far to justify any fair price for completing task. The price would have to be higher/greater than 25 dollars to justify it." (P31)

> "Getting there would take me longer than actually completing the task." (P39)

> "Other areas were too far from the Metra [the commuter rail system in Chicago] to make it worth my while. Others were still close to the Metra but far enough away where the ticket round trip would be a bit pricy." (P16)

> "I didn't think any price would be worth the commute and risk while still offering even a marginally fair price." (P23)

More specifically, these participants suggested that when the cost of commute time (either in raw time or money) rises above a certain level (in two cases the financial or temporal cost of the task), they would no longer be willing to accept the task. This feedback should help guide future work involving modeling distance decay's role in willingness decisions in the sharing economy.

3.3.4 Summary of TaskRabbit study

Above, we have seen that three of our human geographic principles – the "Big Sort", distance decay, and mental maps – play a key role in the effectiveness of the sharing economy. I have also shown that these principles manifest in structural geographic biases in TaskRabbit, in which people who live in high-SES regions in the urban core gain most of the benefits of TaskRabbit's rendition of the sharing economy (at least in the Chicago area). These biases also mean that TaskRabbit is both more expensive and less accessible to people in low-SES areas. I have also observed a smaller role for our fourth geographic principle, structured variation of population density, observing that prices are also higher (and to a lesser extent, service is less available) in high-SES, low-density suburbs as well. Below, I explore whether these same trends persist in an entirely different rendition of the sharing economy: the well-known ride-hailing platform, UberX.

3.4 Study 2: UberX (Ride-Hailing Sharing Economy Platform)

As noted above, the goal of my UberX study is to identify whether the key findings from my TaskRabbit study – the importance of the four geographic principles and their manifestation through specific geographic biases – generalize to UberX. Most (if not all) studies of the sharing economy thus far have focused on a single sharing economy platform. By taking this multi-platform approach, I aimed to gain a more general understanding of the role of geography in the sharing economy (rather than a platform-specific understanding).

I focus my attention in this section on an analysis of UberX wait times, a dimension of effectiveness related to my willingness variable in TaskRabbit. In UberX, drivers can show their willingness to pick up a passenger by accepting or rejecting a fare, by avoiding (or spending time in) certain areas or by selectively turning on and off their availability as they approach certain areas (as was mentioned above). All of these expressions of willingness manifest in the amount of time a potential customer has to wait before an UberX driver arrives at her/his location. Moreover, these wait times can be automatically obtained, affording a quantitative understanding of geographic effectiveness just as was the case with TaskRabbit. It is important to note that I do not analyze price in UberX, as this is determined automatically by a basic formula in all cases (a point to which I return below).

More specifically, I structure this investigation of UberX wait times around the following research question:

Research Question #3: RQ-Wait Times: How does human geography affect UberX wait times?

Motivated by the geographic principles considered in this paper and in analogy to my TaskRabbit study, I made two hypotheses with respect to this question. First, I hypothesized that, all other factors being equal, wait times would be higher in low-SES areas than in high-SES areas (*H-Wait Times-SES*). Secondly, I also hypothesized that structured variations in population density would be a significant factor in wait times, with denser areas having more convenient access to UberX (*H-Wait Times-Population Density*).

In addition to highlighting the similarities and differences in the geography of UberX versus that of TaskRabbit – as we will see, there are far more similarities than differences – this section also makes a methodological contribution to the sharing economy literature. Specifically, to test my hypotheses, I adapt spatial Durbin modeling, an advanced spatial statistical technique from the natural and social sciences, and show how it is often critical for conducting robust geographic sharing economy analyses. To my knowledge, this study represents the first use of spatial Durbin modeling in the human-computer interaction community , and I expect that this work can provide statistical assistance for other sharing economy researchers and researchers in other domains who encounter similar types of spatial data. As such, I dedicate a significant portion of the methods section below to explaining the character, intuition, and proper execution of spatial Durbin modeling.

The remainder of this section will proceed as follows: I first introduce the datasets I utilize in the analysis of UberX wait times. Next, I describe spatial Durbin modeling and explain why it is essential for understanding the geography of the sharing economy in many cases. Following my discussion of methods, I then present the results of my models, highlighting similarities and differences with my TaskRabbit results and discussing connections to my four geographic principles.

3.4.1 Datasets

In every metropolitan area where UberX provides service, there is a defined region where the service is available. At the time this analysis was performed (late 2014), this region did not encompass all of Cook County. Specifically, there are 275 census tracts (shown in Figure 3.3) on the southern end of Cook County that were outside of UberX's operating area. Thus, I excluded these tracts from my study. It is important to note that UberX's service area has since expanded. I return to this issue in the Discussion section, in which I highlight potential "early access" benefits provided to certain types of areas over others in geographic social computing systems.

The tracts in which UberX did not provide service at the time of analysis are systematically poorer than the tracts in which it did offer service. This provides my

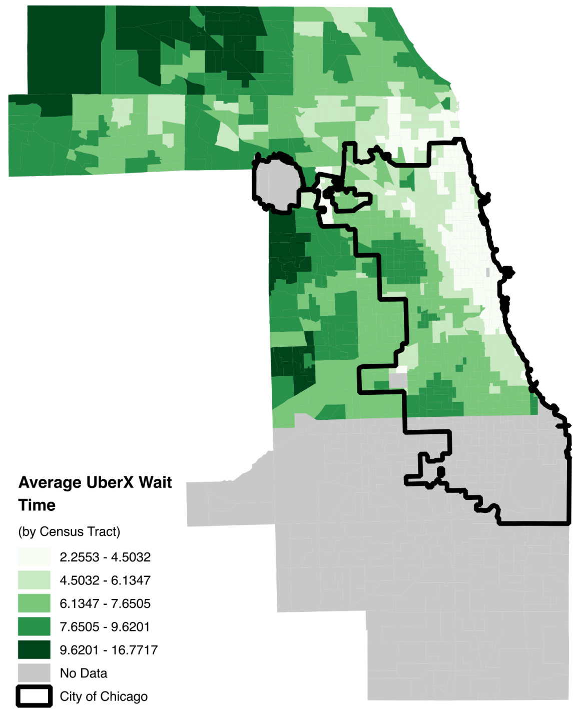

Figure 3.3: Average UberX wait times in Cook County, Illinois. In grey are tracts excluded from my analysis because (a) they were not within UberX's operating area (the large block to the south) as of time of data collection or (b) the Uber API did not provide a single ETA value for these tracts.

first evidence that some of the factors associated with effectiveness in TaskRabbit – particularly the "Big Sort" along the SES dimension – play the same role in UberX. More specifically, tracts excluded from UberX's service area had an average median household income of \$53,122 (sd = \$19,247), whereas the tracts served by UberX have an average median income of \$56,369 (sd = \$29,761).

I used Uber's Time Estimate API (available through their developer website) to measure wait times in the 1,041 Cook County census tracts within the UberX operating area. I sampled the centroid of each census tract every hour for a period of 7 days, leaving us with 168 samples per tract in the ideal case. Uber's API never provided wait times for the three census tracts containing the Chicago area airports, which is likely due to Chicago city ordinances that prohibited Uber from providing service at airports at the time [74]. Thus, I excluded these tracts from my data, leaving us with 1,038 census tracts for my analysis.

In some cases, Uber's API did not provide a wait time every time a tract was sampled, the reasons for which are unclear. However, ninety-eight percent of tracts in my sample received a wait time more than 80% of the time, and the census tract with the fewest samples received valid responses approximately 60% of the time (100 of 168). I compute the mean of all wait times for each census tract, and use this average wait time as the core dependent variable in my analysis of UberX.

To examine *H-Wait Times-SES*, I used the same SES data as I did for the TaskRabbit study: median household income (MHI) (in $10K increments) from the United States Census' American Community Survey 2006-2010 dataset. Again, I log-scaled this variable to reduce the effect of a long-tailed wealth distribution. To examine *H-Wait Times-Population Density*, I utilized United States census data on the number of people per square kilometer in each tract. I also log-scaled this variable, again to reduce the skew of a long-tailed population density distribution.

Finally, I note that while the methods I used to study TaskRabbit enabled us to compute a "distance from home region" variable, this was not the case for my UberX work: UberX's API does not give a regular starting point for a given UberX driver. As such, in this study, my ability to speak to distance decay directly is limited (although, as is discussed below, my results below do suggest several indirect findings).

3.4.2 Spatial Modeling

If the datasets I considered in this study had not been not geospatial, my modeling task would have been straightforward. Specifically, using standard regression modeling techniques like ordinary least squares (OLS), I could have assessed whether there were significant relationships between MHI and population density (independent variables) and UberX wait times (dependent variable), as well as determined the effect sizes of these relationships.

However, the geospatial nature of my datasets demand that my methods be considerably more sophisticated. For instance, let us consider a wealthy census tract that is surrounded by less wealthy census tracts (e.g. as is the case near the University of Chicago on the "South Side"). One might imagine that this tract's UberX wait times may be affected by the fact that its neighbors are less wealthy. After all, UberX drivers might not want to drive through poorer areas to get to this tract, which they may perceive to be less safe (as we saw in the case of TaskRabbit), or they may prefer to stay in an area that has consistent and widespread high incomes. Conversely, a poorer tract near richer tracts may see opposite effects. However, traditional regression modeling assumes that all samples (tracts) are independent and cannot incorporate potentially critical information about a tract's neighbors' income in its estimates of wait times for the tract. In other words, one can think of traditional regression modeling as failing to account for Big Sort effects when applied in many types of sharing economy analyses (and other types of analyses in geographic HCI).

More generally, when a dataset consists of specific geographic locations associated with attributes – which is the case for UberX wait times – it is common for individual data points to be affected by other nearby data points, or to be spatially autocorrelated [4]. The Big Sort is an instance of spatial autocorrelation in which the variables of interest are demographic in nature.

The presence of spatial autocorrelation, including in the Big Sort case, means that the data from one location often is not independent of data from neighboring locations, violating a core assumption of traditional regression modeling techniques like OLS. The output of these techniques – significance, effect sizes, etc. – are all conditioned on the assumption of independence of observations in the independent and dependent variables, something that often will not be true with geographic sharing economy data. As such, specialized techniques are needed, not just to acquire more statistical power and understanding, but also simply to gain reliable insight on the associations in question and the role of spatial relationships in these associations.

When engaging in spatial statistics, a frequent first step is to model the spatial structure of the study area. A common approach to generating a representation of this structure – and the approach I use in this work – is called a Queen's weights matrix (there are also other distance-based schemes, e.g. k-nearest-neighbors). With a Queen's matrix, the neighbors of a given census tract (or polygon more generally) are assumed to be all the tracts (or polygons) that are directly adjacent through either an edge or a vertex (similar to the moves available to a queen in chess).

Once the spatial structure has been encoded, we can turn my attention to modeling spatial effects between neighbors. These effects can be of three distinct types [107]:

1. A correlated spatial relationship: unknown factors lead to similar outcomes (e.g. wait times) between two neighboring locations.
2. An endogenous spatial relationship: an outcome (e.g. wait time for a given census tract) for one location is dependent on the outcomes (e.g. wait times) of neighboring locations.
3. An exogenous spatial relationship: an outcome (e.g. wait time for a given census tract) for one location is associated with the predictors of neighboring locations

(e.g. the median income of its neighboring tracts).

In the context of my modeling exercise, this means that a given census tract's UberX wait time may be (1) similar to its neighbors because of some unmeasured factors, and/or (2) dependent on the wait times of its neighbors, and/or (3) dependent on the population density and median income of its neighbors (my predictors / independent variables).

Until recently, the majority of the focus in geostatistical modeling has fallen into two camps: modeling the correlated relationships (#1 above) by accounting for any spatial relationships using the error term (known as a spatial error model), and modeling the endogenous spatial relationship (#2 above) using the weights matrix to lag (or spatially weight) the dependent variable (known as a spatial lag model). The third type of spatial relationship – exogenous spatial relationships – had until recently largely been ignored in the natural and social sciences, let alone in the human-computer interaction literature. While most treatments of spatial data in the human-computer interaction literature do not consider spatial autocorrelation at all and instead utilize standard regression techniques for spatial data, to the extent that spatial models have been used, they have been the more traditional spatial lag and spatial error models (e.g. [80,82,106]).

Spatial error models may be sufficient when model interpretation is unimportant and addressing independence assumptions while maximizing predictive power is the only consideration [41]. However, recent work in the spatial statistics literature has argued that both endogenous and exogenous relationships need to be examined when using spatial modeling to shed light on the underlying spatial processes [41,173], as I am doing here (and as is common in the HCI community's consideration of spatial data more broadly). In this vein, a more generalized modeling approach – the spatial Durbin model – that accounts for both endogenous and exogenous relationships has begun to be recommended as best practice [41,98].

A spatial Durbin model can be understood as having multiple versions of each variable corresponding to the endogenous and exogenous spatial relationships discussed, supporting better interpretation of the spatial relationships in the data. In my case, a spatial Durbin approach models the average UberX wait time for a given tract as a linear combination of (1) the values of the independent variables in that tract (as is typical in OLS regression), (2) the average of the values of each independent variable (MHI and population density) in neighboring tracts according to the Queen's matrix (spatially exogenous relationships), (3) the average of the wait times in neighboring tracts calculated in the same fashion (spatially endogenous relationships, i.e. spatial lag term), and (4) an error term, which functions similarly to the error term in a traditional OLS. So, in other words, whereas a traditional OLS regression for my experiment would involve two independent variables and an error term, a spatial Durbin approach applied to my problem would have 2 + 2 (spatially exogenous) + 1 (spatially endogenous) = 5 independent variables and an error term.

In light of the tremendous interpretative advantages of spatial Durbin modeling – none of the key distinctions between the "direct" and "indirect" effects below would be possible without Durbin modeling – I employed spatial Durbin modeling as my primary analytical tool for understanding my UberX data. I was not able to apply spatial Durbin modeling to my TaskRabbit analysis for one critical reason: my TaskRabbit experiment required the employment of mixed effects models and, to my knowledge, mixed effects models have not yet been integrated with spatial Durbin approaches. Indeed, spatial Durbin approaches have only recently become feasible;

up until several years ago, they were too computationally demanding for common practice [41]. In the TaskRabbit case, mixed effects models were required to control for the lack of independence between observations gathered from the same individual. In concert with my university's statistical consulting center, I determined that the violations of independence that are due to observations coming from the same participant (which are handled by mixed effects modeling) were more serious than those due to spatial autocorrelation (thanks in large part to the fact that I was not considering many immediate neighbors in my TaskRabbit observations). Because no existing modeling approach (to my knowledge) allows us to account for both types of independence assumption violations, I used the mixed effect models discussed above. As we will see below, I found nearly identical high-level results in my TaskRabbit and UberX analyses, adding credence to both the high-level results and the modeling choices in each.

The majority of future geographic sharing economy research will likely not face the challenges I did with TaskRabbit and will be able to gain the advantages of spatial Durbin modeling as I do here with UberX. Indeed, the two studies that most directly resemble my research – Quattrone et al. [137] and Stark and Diakopoulos [153] – did not face the statistical challenges associated with my TaskRabbit study.

Interpreting Spatial Durbin Models.

In general, when examining the results of Durbin models, traditional outcome metrics like the value and significance of coefficients (betas) are far less useful for interpretation than a series of specialized metrics, in particular the Rho term, indirect effect values, and direct effect values.

The Rho term captures the effect of spatial diffusion in the dependent variable (wait times) as one neighbor's value affects another neighbor's value, which then affects another neighbor's value, and so on. More specifically, the Rho term encapsulates endogenous spatial relationships and in the context of my work, describes how a given tract's wait time should be affected by the wait times of its neighbors (endogenous spatial relationships). This term is not interpreted like a traditional model coefficient, but instead is multiplied by the spatially-weighted average of neighbors' measured wait times.

The need for direct and indirect effect values arises out of the fact that spatial dependence invalidates the interpretive benefits of model coefficients in traditional regression (e.g. OLS). Traditional approaches compute regression coefficients through partial derivatives of the regression formula with respect to each independent variable. Because of spatial dependence, however, when these partial derivatives are computed in a spatial model that incorporates data from the neighboring area, any given partial derivative will in turn be dependent on the values of the neighboring tracts' partial derivatives. These feedback loops, caused by the spatial structure, get built into the models. This means that the variable coefficients in a spatial model cannot be interpreted directly because the partial derivatives are not orthogonal [41].

On the surface, these feedback loops seem like a challenge to the interpretative power of spatial Durbin modeling. However, LeSage and Pace [98] introduced direct effects and indirect effects to explicitly address this issue. Indirect and direct effects are calculated by averaging across all of the relevant partial derivative values at every location (calculated based on the lagged value of the variable in question). I follow Yang et al. [173] and use a Markov Chain Monte Carlo (MCMC) approach to randomly permute input data in order to estimate the average effects and generate an average over the permuted output .

Direct effects describe average relationships between the dependent and independent variables that are analogous to the relationships modeled in traditional regression approaches. Specifically, a direct effect for a given independent variable describes the average impact the value of that variable at a specific location has on the value of the dependent variable at that location. In the context of my work, this means, for instance, the effect the median household income of a given tract has on the wait times of that specific tract. Each independent variable has its own direct effect.

Conversely, indirect effects model the relationship between the value of a dependent variable at a given location and the values of independent variables at neighboring locations (spatially exogenous relationships). In my analysis, indirect effects capture, for instance, the effect of the average median household income of neighboring census tracts on a given census tract's wait time. Like is the case with direct effects, each independent variable has its own indirect effect value (so each independent variable in my model has both a direct effect value and an indirect effect value).

When describing the results of a spatial Durbin model, it is considered best practice [98] to present the Rho term (endogenous effect), the modeled coefficients, and the direct and indirect effects (exogenous effects), but interpret only the endogenous and exogenous spatial relationships. This is due to the unclear meaning of the standard modeled coefficients. I follow this best practice below.

3.4.3 UberX SES Results

Table 3.4: The Results of my Cook County UberX *Spatial Durbin* Model.

Fixed Effect	Estimate	p-value
Rho (Weighted effect of neighbors' wait times)	0.92	< 0.001
log_2[people/km^2] (population density)	-0.37	*n.s.*
lagged log_2[people/km^2] (population density)	-5.63	0.002
log_2[Tract income in \$10,000] (median income)	6.09	0.09
lagged log_2[Tract income in \$10,000] (lagged median income)	-10.21	0.02
Intercept	114.88	< 0.001

Table 3.4 shows the model coefficients for my UberX wait times model. In my model, the endogenous relationship between a tract and its neighboring wait times is quite strong: 92% of the average wait time of immediate neighbors is contributed to the wait time of a given tract. This is intuitive: a tract should not have drastically different UberX wait times than its neighboring tracts due to the nature of wait times. For instance, if a tract has five neighbors and the sum of their average wait times is 1,500 seconds (mean = 300 seconds), the neighboring tracts would contribute 259 seconds (92% of the 300 second mean) to the predicted wait time of the tract in question. Coefficients in grey are commonly reported, but not interpreted.

Table 3.5: The *Direct* Effects of my Cook County Model

Direct Effects	**Estimate**	**p-value**
log_2[people/km^2] (population density)	-3.01	0.03
log_2[Task tract income in \$10,000] (median income)	4.02	*n.s.*

Table 3.6: The *Indirect* Effects of my Cook County Model

Indirect Effects	**Estimate**	**p-value**
log_2[people/km^2] (population density)	-69.09	< 0.001
log_2[Task tract income in \$10,000] (median income)	-53.68	0.08

Tables 3.5 and 3.6 show the direct and indirect effects of my independent variables. The tables reveal that, when examining the entirety of UberX's service area across Cook County, population density is significant in both its direct and indirect effects (supporting *H-Wait Times-Population Density*). The indirect effects for population density suggest a strong and inverse relationship between population density and wait times (note that the indirect effect for population density in Tables 3.5 and 3.6 is both large and negative). Specifically, if the average population density across all of a tract's neighbors were to double, we would expect a decrease in average wait time of approximately 70 seconds for that tract. This is not an extreme scenario: the mean population density in my study region is 6,183 people/km^2 and the standard deviation is 7,616 people/km^2. This means that tracts in a very dense area should expect much lower wait times than areas where there are fewer people per km^2.

While significant, the direct effects of population density – i.e. the role played by the population density of the tract in consideration itself – had a much smaller effect size. A doubling of a specific tract's density would only lead to an average decrease in wait time of approximately 3 seconds, for that tract. This is a trend that we will see repeated below: the characteristics of a tract's region appears to matter more than the characteristics of the tract itself.

More generally, the results in Tables 3.5 and 3.6 substantiate the importance of the principle of structured variation of population density that I observed in the TaskRabbit study. Specifically, the sharing economy seems to be significantly less effective in the suburbs relative to the central city. For UberX, this finding is quite visible in Figure 3.3, where we see wait times of over 10 minutes in the very-low-density distant suburbs and under 3 minutes in dense urban cores.

While Tables 3.5 and 3.6 strongly suggest that structured variation in population density is a prominent factor in UberX wait times, they display less clarity about income's role. Tables 3.5 and 3.6 show no significant direct effects for income and only marginally significant indirect effects (providing little support for *H-Wait Times-SES* at this stage). With regard to these indirect effects, it appears that if a tract's neighboring region became more wealthy, UberX wait times would decrease in that tract. However, at least at this point, I only have marginal confidence in this relationship.

Table 3.7: The Results of my Cook County UberX *Spatial Durbin* Model.

Fixed Effect	Estimate	p-value
Rho (Weighted effect of neighbors' wait times)	0.95	< 0.001
log_2[people/km^2] (population density)	-0.07	*n.s.*
lagged log_2[people/km^2] (population density)	0.52	*n.s.*
log_2[Tract income in \$10,000] (median income)	5.29	0.04

Fixed Effect	Estimate	p-value
lagged log_2[Tract income in $10,000] (lagged median income)	-14.55	< 0.001
Intercept	33.29	0.05

Table 3.8: The *Direct* Effects of my Chicago Model

Direct Effects	**Estimate**	**p-value**
log_2[people/km^2] (population density)	0.19	*n.s.*
log_2[Task tract income in $10,000] (median income)	-0.41	*n.s.*

Table 3.9: The *Indirect* Effects of my Chicago Model

Indirect Effects	**Estimate**	**p-value**
log_2[people/km^2] (population density)	8.95	*n.s.*
log_2[Task tract income in $10,000] (median income)	-190.61	<0.001

Tables 3.5 and 3.6 show results for a model that considered the entirety of Cook County. However, many sharing economy decisions and debates occur at the municipal (city) level (e.g. [22,37,48]), where there tends to be less variation in population density. To understand the relationships between income, population density, and wait times within a central city itself – rather than an entire metropolitan area that includes suburbs and exurbs – I re-ran my model focusing only on census tracts that fall within the borders of the city of Chicago. I present the coefficients of this model in

Table 3.7 and the average direct and indirect effects in Tables 3.8 and 3.9. Table 3.7 shows that, as expected, in my Chicago model, the endogenous interaction between a tract's wait time and its neighboring wait times is quite high, just as it was for my Cook County model. A tract's immediate neighbors contribute 95% (Rho = 0.95) of the average of the neighboring wait times to the predicted wait time (as opposed to 92% in the Cook County model).

Tables 3.8 and 3.9 show that I identified no statistically significant direct effects in my Chicago-only model. That is, the income and density values of a specific tract in Chicago do not seem to play a role in that specific tract's wait time. However, we do see a significant and substantial indirect effect for median income: if the income of an area goes up, wait times go down by a large margin (supporting *H-Wait Times-SES*). To be more specific, my model suggests that if a Chicago census tract's neighbors experienced a doubling in their average median household incomes, I would expect to see that tract's UberX wait time decrease by over 3 minutes and 10 seconds (190.6 seconds) on average. This suggests that while individual poor census tracts surrounded by higher-income tracts should benefit from the wealth of their neighbors, UberX is less effective in regions of Chicago where there is more widespread poverty. This can be seen in Figure 3.3, in which the low-income neighborhoods in the southern and western areas of Chicago have much higher wait times. I discuss how this may be related to mental maps in Section 3.5 below.

The results in Tables 3.8 and 3.9 can be read as strong support for the Big Sort's influence on the sharing economy within regions of similar population density (e.g. within the city limits of Chicago). While the SES of a specific tract does not seem to have an impact on sharing economy effectiveness in that tract (as instantiated by UberX wait times), the SES of a tract's neighborhood has a substantial impact the tract's wait times. This leads to high wait times in large low-SES areas (e.g. the South Side) and low wait times in large high-SES regions (e.g. northern Chicago). This also means that even if a neighborhood in a low-SES area begins to improve its SES, there will likely continue to be a damper on sharing economy effectiveness in this neighborhood. If those who argue that the sharing economy will become a dominant economic paradigm are correct, this is a troubling implication.

In the section that follows immediately below, I provide further discussion of these results in the context of my TaskRabbit results and four geographic principles.

3.5 Discussion

In my investigation of the geography of the sharing economy, I examined two different sharing economy platforms using diverse methodologies that ranged from quantitative and qualitative analyses of survey results to the application of spatial Durbin autoregressive models on data gleaned from APIs. In both cases, however, I found very similar high-level findings regarding the geography of the sharing economy: the sharing economy is more effective in dense, wealthy neighborhoods and significantly less effective in suburbs and low-income urban neighborhoods. Moreover, my results pointed to the underlying geographic principles responsible for these structural geographic biases: the Big Sort, structured variations in population density, distance decay, and mental maps.

In this section, I discuss the implications of these findings along several key dimensions: the role of race/ethnicity, suggested improvements to the design of sharing economy platforms, and directions for future work.

3.5.1 Examining my findings with a lens informed by race and ethnicity

As noted above, due to the tremendous economic inequalities that occur across racial and ethnic lines in the United States (and elsewhere), SES and race and ethnicity tend to be closely linked. Indeed, I observed that the percent of the population that self-identifies as white (non-Latino) has a strong correlation with income in my study area ($r = 0.67$). White (non-Latino) is a demographic variable provided by the U.S. census whose inverse (the percent of the population that does not identify as white or Latino) is often interpreted as the percent of the population that identifies as a racial or ethnic minority [19].

Because of this correlation, I hypothesized that many of the patterns I saw with SES and the Big Sort would also occur with race and ethnicity and the Big Sort. To test this hypothesis, I replaced SES as an independent variable in both the TaskRabbit and UberX models. For TaskRabbit, I found that the percentage of the population that is white (non-Latino) was a marginally significant predictor of willingness (p = 0.06). As was the case with SES, I did not observe an explicit effect for white (non-Latino) with respect to price. However, due to the correlation above, poor neighborhoods tend to be minority neighborhoods in Chicago (and in many other places in the world), so I also observed the same distance decay effects with respect to white (non-Latino) as I did with SES.

I identified a similar finding for UberX wait times as I did for TaskRabbit willingness: in a version of my Chicago-only spatial Durbin model with SES replaced by the percent of the population that is white (non-Latino), I saw a significant indirect effect for percent white (non-Latino). If the area around a given tract changed from 0% white (non-Latino) to 100% white (non-Latino), we would expect to see that tract's

UberX wait time decrease, on average, by 317 seconds (5 minutes and 17 seconds). In other words, if an entirely non-white area in the city of Chicago were to see a complete demographic shift towards being entirely white, my results suggest that tracts in that area may see UberX wait times decrease drastically, making the UberX service much more effective. This suggests that while individual non-white census tracts surrounded by white neighbors should benefit from better UberX service, UberX is less effective in regions of Chicago where there are large minority populations. Interestingly, these results dovetail with very recent findings by Ge et al. [52] that suggest that people with African American-sounding names wait longer for UberX service in Seattle.

It is important to note that in the above results, the core geographic principles at work are no different than is the case with SES, they are simply manifest in race and ethnicity rather than SES. For instance, just as was the case for low-SES neighborhoods, many minority neighborhoods also see reduced TaskRabbit effectiveness due to the Big Sort and due to distance decay interacting with the Big Sort (with respect to the location of TaskRabbit workers' residences, which tend to be farther away from large minority districts than from large white (non-Latino) districts). Similarly, the Big Sort has the same effect on large minority neighborhoods in UberX as it does on large low-SES neighborhoods. Indeed, as per the correlation mentioned above, many minority neighborhoods and low-SES neighborhoods are one and the same and thereby suffer from identical lower sharing economy effectiveness. More generally, a sharing economy platform that does not serve low-income people will, in general, also fail to serve non-white populations, at least in North America (a key concept in the sociological theory of intersectionality [29]). Similarly, it is likely that investigating sharing economy effectiveness across other demographic properties affected by the

Big Sort (e.g. educational attainment, religion, age, political affiliations), this would show similar findings.

However, regardless of the geographic mechanisms at work, just as was the case for SES, the structural racial and ethnic biases in the sharing economy identified in this section are quite important in their own right. Groups defined by race and/or ethnicity are protected classes in the United States [27]. If these results are found to generalize across other cities – with sharing economy systems working better in white (non-Latino) neighborhoods than minority neighborhoods – this could become an important data point in the ongoing debate about the sharing economy occurring across the U.S. and around the world.

3.5.2 Additional Relevant Geographic Principles

In this chapter, I have discussed how four geographic principles (introduced in Chapter 2) play a key role in the sharing economy and result in structured geographic biases along SES lines (and those defined by race and ethnicity). However, these four principles are almost certainly not the only aspects of human geography that are important to consider when examining the sharing economy. For instance, two human geography principles worthy of exploration are border effects and edge cities.

Border effects are a well-known human geography principle that describe what occurs when two neighboring places that are on the opposite sides of an administrative boundary have tremendously different circumstances with respect to a variable of interest. My results suggest that as some municipalities like Austin, Texas begin to place restrictions on sharing economy services (especially Uber and Lyft) [21,125], border effects will become increasingly important in the sharing economy. Specifically, I (unsurprisingly) found that a census tract's wait times were highly dependent on neighboring tracts' wait times. This means that adjoined municipalities will likely suffer reduced sharing economy effectiveness when one municipality places restrictions on sharing economy services. For instance, my results suggest that Austin's suburbs are going to be severely affected by Austin's sharing economy-related decisions, even if they have no say in these decisions. Examining the effect of differing sharing economy regulations within the framework of border effects is an important direction of future work.

My results also suggest that if and when the sharing economy becomes prominent in suburbs, "edge cities" [50] will lead the way and become secondary sharing economy hubs. An "edge city" is a concentration of work and leisure resources in a suburb that has good vehicle accessibility with respect to the rest of the metropolitan area, usually due to a nearby intersection of multiple freeways (e.g. a "ring road" and an intersecting highway) [50]. Common examples include Tyson's Corner, VA, Bloomington, MN, and the Rosslyn-Ballston Corridor in the Washington, D.C. metro area. Edge cities emerged in the second half of the twentieth century, becoming competitors with central business districts for shopping, employment, and entertainment services. The vehicle accessibility advantages that led to the agglomeration of traditional services in edge cities should also apply to sharing economy services.

Namely, relative to other suburbs, edge cities will be significantly less affected by the negatives associated with distant decay due to these accessibility advantages. Moreover, in the ride-hailing space, there is reason to believe that the limited availability of mass transit options in edge cities makes them even more suited to the sharing economy. Another more uncomfortable potential advantage for edge cities in the sharing economy is that they and their surrounding residential areas tend to be relatively high -SES and populated by non-minority racial and ethnic groups.

Overall, given the numerous properties of edge cities that interact with properties of the sharing economy, many branches of sharing economy research – e.g. work on improving sharing economy effectiveness in new regions, work seeking to understand its long-term impact on urban areas – should likely consider edge cities as in important near-term direction of inquiry.

3.5.3 Implications for "Geosociotechnical" Design

As noted above, in almost every case, the structural geographic biases that are the result of the four geographic principles are not destiny for the sharing economy: they are the outcome of interactions between these principles and the "geosociotechnical" design of existing sharing economy platforms. In other words, given the design choices in these platforms and the inherently geographic nature of the sharing economy, it is not a surprise that these principles manifest in the biases I observed. Indeed, awareness of these principles led us towards my SES-related hypotheses.

Fortunately, the important role of geosociotechnical design in the structural biases identified in this chapter means that there is an opportunity to address these biases with design changes. In the remainder of this sub-section, I outline a number of geosociotechnical improvements that could lead to the reduction of the SES, racial, and ethnic biases I identified in the sharing economy.

Using Design to Improve the Geographic Distribution of crowd members

My results suggest that the Big Sort residential geography of sharing economy workers – coupled with distance decay – is a primary causal factor behind the geographic variation in the effectiveness of sharing economy systems that I observed. For instance, in my TaskRabbit study, I found that workers were not willing to travel long distances for a task. This would not result in geographic disparities in the effectiveness of TaskRabbit if TaskRabbit workers were evenly distributed across the Chicago area. However, as noted above, TaskRabbit workers are heavily clustered in relatively wealthy and dense parts of the region (as per Big Sort processes), leading to higher prices and fewer jobs accepted in other types of areas. One might call these '(commercial) sharing economy deserts' as an analogy to 'food deserts' [17], which tend to occur in similar types of areas. Along the same lines, with respect to UberX, Dillahunt and Malone [33] found that few participants in a workshop on the sharing economy for job-seekers were even aware of sharing economy services.

These results suggest that recruiting new sharing economy workers in areas that suffer from the wrong end of the structural biases I identified would go a long way to eliminating these biases. For instance, it is likely that a relatively small number of TaskRabbit crowd members who live on the South Side of Chicago could have substantially diminished any price and willingness disadvantages in these areas. With respect to Dillahunt and Malone's findings, the same may be true for UberX wait times. Moreover, service quality improvements in disadvantaged areas could lead to a larger (and more diverse) customer base, increasing the incentive to recruit more workers from disadvantaged areas.

The question then becomes: how can sharing economy platforms do such recruitment? Making different design choices can likely contribute to the answer. For instance, TaskRabbit requires that workers have a bank account to participate on their platform, with low-income and minority neighborhoods having a much greater percentage of the population that is "unbanked" [156], inherently reducing the potential working population in these places. UberX additionally requires workers to own a car and have active insurance, among other requirements , which likely has a similar effect (Dillahunt and Malone [33] found that two-thirds of people in a sharing economy workshop for disadvantaged communities did not have a car that met Lyft's requirements). Other worker restrictions may also play a role: both TaskRabbit and UberX require that workers pass a background check, and it is unclear if minor prior offenses would result in rejection (e.g. a minor drug possession arrest). Low-income neighborhoods have a higher rate of these minor offenses [67].

Some sharing economy platforms have begun to make design changes in this direction. For instance, recent efforts by Uber to provide banking services to its drivers [94] and to make obtaining car leases easier for potential drivers [171] could potentially address some of the problems that I identified in this chapter. Of course, these initiatives could also lead to new problems, including the serious risk of exacerbating debt-related challenges in disadvantaged areas, and these leases have been accused of being predatory [123]. My results suggest that research into these and other mechanisms for increasing worker participation rates in 'sharing economy deserts' should be a top priority for future work.

Addressing Workers' Mental Maps

My TaskRabbit results showed that comfort levels in workers' mental maps played a key role in their willingness to accept tasks and specifically made them less likely to accept tasks in wide swaths of southern and western Chicago. A similar finding was identified by Lee et al. [97], who found that UberX drivers turned off their availability when they were in neighborhoods they perceived as undesirable. In both cases, workers cited perceptions of crime as the reasons for their discomfort in certain areas. The mental maps literature, however, suggests that in many cases perception does not match reality. For instance, Matei et al. [109] showed that comfort levels associated with regions in the Los Angeles metropolitan area were effectively uncorrelated with crime levels. Additionally, as noted above, our mental maps generally struggle to incorporate sufficient detail to be able distinguish pockets of low-crime areas in unfamiliar high-crime districts.

The gap between perception and reality in mental maps presents a potentially powerful opportunity for geosociotechnical design. One straightforward design improvement would be to provide workers with geographically-linked crime statistics in an easily-digestible format that would allow for design-making on-the-fly. In most areas, crime statistics are public information and could be surfaced via a map in an app for crowd members quite easily. A more interesting and likely more useful approach would be to provide this information in the context of a given task, e.g. when a TaskRabbit worker is deciding whether to accept a task or when an UberX driver is driving through a specific neighborhood. This information could take the form of basic crime statistics or, matching the norms of the sharing economy, could be reported as a "geographic reputation score". This score could take into account both public crime information as well as geographically-linked incident reports privately held by a sharing economy platform. Based on the work of Matei et al., it is likely that many areas that currently are associated with high discomfort would have high geographic reputation scores (and perhaps vice versa). If this information were made available to workers, it could address some of the TaskRabbit willingness and UberX wait time bias that I observed in this study .

A Role for "Sociotechnical Auditing"

My results add to evidence that auditing has an important role to play in protecting the sharing economy from bias, just as has been argued in more explicitly algorithmic domains (e.g. [85]) and for other sociotechnical platforms (e.g. [3,146]). Fortunately, the geographic techniques I developed and adopted here – especially my spatial Durbin modeling approach – can provide a useful lens in this process. The relative geographic effectiveness of sharing economy platforms in a given administrative district likely would be a valuable data point for the many sharing economy debates that are occurring around the world. My experiments outlined above should be replicable in most (if not all) areas in which TaskRabbit and UberX are active, and my techniques should relatively easily generalize to similar platforms (e.g. Lyft).

To make repeating my work in other areas as straightforward as possible, I am releasing my UberX data collection and spatial Durbin statistical framework under an MIT license . This package should allow someone with technical training to quickly repeat my UberX experiment in their area with relatively little effort. Moreover, it should be relatively straightforward to adapt my code to other outcome metrics besides wait time and to other similar sharing economy platforms.

Additionally, my approach here has been to examine system-wide effects, but the auditing of the sharing economy should also likely occur at the worker-specific level. By examining the geographic history of a given worker, it should be possible to determine if that person is exerting implicit or explicit bias in their pricing and willingness decisions. If the worker has a long enough history with a platform, techniques similar to those I described above can be employed (e.g. adapting wait times to job history-specific attributes). In more traditional workplaces, "substantive oversight of decision making" is one facet of minimizing workplace gender and racial bias [13], and it is intuitive that the sharing economy could learn from this body of literature. Correcting for implicit bias may sometimes be as simple as making decisions more legible, e.g. "98% of your completed jobs have been in areas that are at least 95%

white (non-Latino)".

Lastly, while the term "algorithmic auditing" has been adopted for doing this type of auditing work in a technological context, this term is not ideal for the sharing economy. Algorithms play a role – especially in the case of UberX – but sharing economy platforms are sociotechnical, not just technical. As we have seen, it is human biases – in the form of the Big Sort, mental maps, etc. – that are the drivers of many of the structural geographic disadvantages that we observed in this study. As such, auditing in the sharing economy is "sociotechnical" auditing (and even perhaps "geosociotechnical auditing") and needs to adopt approaches from both the algorithmic auditing literature and the large literature on detecting bias in human decision-making, e.g. the implicit bias reduction technique mentioned immediately above.

Task-specific vs. Global Pricing

My results above suggest that UberX's decision to fix prices globally as a function of distance may have reduced pricing-related bias in its platform relative to TaskRabbit, which at the time of my analysis allowed for per-task pricing. Namely, whereas in TaskRabbit both price and willingness were entirely dependent on human decision-making processes that are subject to bias, in UberX, this is only true of willingness (manifest in wait times). TaskRabbit's pricing model has changed since my study and now more closely aligns with that of UberX. Specifically, in most cases, workers now define hourly wages for categories of tasks, and are algorithmically presented to task requesters. As such, it is likely that the price-related biases I identified above are either reduced or manifest differently in the design of the TaskRabbit platform that is current as of this writing (like many sharing economy platforms, TaskRabbit is frequently changing its pricing structure).

However, basic economics tells us that price and willingness are not independent, and the relationship between the two was specifically addressed by a number of my TaskRabbit participants above. Specifically, when price controls are employed, shortages can emerge [161]. As such, if a sharing economy platform uses fixed pricing, and these prices are set too low for tasks in a specific area for whatever reason (e.g. distance, mental maps), willingness will likely drop in this area. This could, indeed, be a factor behind some of the relatively large wait time effect sizes observed in my UberX models. Better understanding the relationship between price and willingness in the context of the geosociotechnical design of pricing models is an important area of future work. Reimagining my work under a variety of different pricing models would be a good place to start.

3.5.4 Other Areas of Future Work

Gender and the Sharing Economy

One critical area of future work highlighted by this research is further examination of the relationship between gender and the sharing economy. I found in the TaskRabbit research that women were significantly less likely to be willing to do a task than men (willingness rates were about 20% lower). While I hypothesized that the effects associated with discomfort and mental maps may be exacerbated for women, resulting in the 20% difference, future studies will be necessary to (1) confirm this difference in other sharing economy contexts and (2) isolate its cause. The import of investigating these two points cannot be understated: if willingness is lower for women, it could have important effects on women's ability to earn comparable amounts as men in the sharing economy: with less competition in areas perceived to be unsafe, men could charge higher rates. There have also recently been high-profile developments associated with the relationship between gender and the sharing economy that are worth studying and that may provide key sources of qualitative and quantitative data with regard to these issues, e.g. a ride-sharing service designed explicitly to serve the safety needs for women passengers by specifically hiring women drivers [43].

Temporal Bias in Access to Sharing Economy Services

In my UberX analyses, I observed that UberX was launched in a higher SES portion of the Chicago region before it became available to the metropolitan area more widely. This is a pattern we see at a more global geographic scale as well, with many sharing economy platforms launching first in relatively high-SES, high-density metropolitan areas (e.g. San Francisco, New York) before their developers open them up to other metropolitan areas.

Relative to some of the other challenges associated with the sharing economy identified in this chapter, this 'temporal bias' is likely less significant, assuming wide launches eventually occur. However, one concern I have is that this temporal pattern may lead to higher SES individuals gaining first mover advantages both as consumers and crowd members, e.g. with regard to reputation scores. Better understanding the launch patterns of sharing economy platforms (and other geographic technologies, more generally) and their possible follow-on effects could be a valuable direction of future work.

Putting Sharing Economy Bias into Context

This chapter is interested in understanding the relative effectiveness of the sharing economy in different areas and the geographic mechanisms behind this variation, not in comparing the effectiveness of sharing economy platforms with their traditional economy equivalents. It may be, for instance, that UberX has significantly lower wait times than traditional cab companies in all areas of Chicago, regardless of the demographic makeup of a neighborhood. Similarly, TaskRabbit may open up new opportunities for acquiring low-cost paid help in small low-SES areas near high-SES regions.

Given that it is widely believed that sharing economy services will substantially displace their traditional economy equivalents in the near- and mid-term future [12,115], understanding the geographic variation in the effectiveness of these services is critical. However, many sharing economy-related debates have been framed as a comparison with traditional economy competitors. As such, it is important that future work provide much-needed robust data points with respect to this comparison. The methodological frameworks I developed above can be used for studies of this type. For instance, one could adopt my TaskRabbit experiment to collect data from taxi drivers. Similarly, my spatial Durbin modeling approach could be used with large-scale trip data collected by taxi companies and obtained by municipal governments [102]. Indeed, I have completed early work comparing UberX to New York City's green and yellow cabs. This research suggests that UberX provides better service to areas with large minority populations compared yellow cabs. However, green cabs, which serve outer boroughs, significantly outperform UberX in this respect.

Of the comparisons between sharing and traditional economy services that have been made [7,103,143], one important factor has tended to be excluded: the informal economy services that often arise to address limitations in traditional economy services (e.g. [95,142,158]). For instance, 'vernacular cabs' [158] – ride-hailing services that are informally organized and have "fares based on negotiations or 'gentlemen's agreements' " – have existed in many low-income areas in the United States for years [158]. In many ways, vernacular cabs (and related systems like sluglines [110]) can be considered "peer-to-peer UberX" and their relationship to (digital) sharing economy services and traditional economy services should be a consideration in any comparative analysis of the sharing and traditional economies.

Vernacular cabs also present several intriguing possibilities for sharing economy researchers and practitioners. Can we develop technologies to support these networks in addition to (or instead of) attempting to adapt centrally-run commercial sharing economy platforms to be more effective in low-SES areas? What would be the effect of having separate platforms for low-SES and high-SES areas? Would a peer-to-peer model work well in high-SES areas? Given the limited amount of information about vernacular cabs, likely the first step in this research direction is formative qualitative work on vernacular cab networks with an eye towards implications for design.

3.6 Conclusion

In this chapter, I have demonstrated how four geographic principles – the "Big Sort", variation in population density, distance decay, and mental maps – result in structural geographic biases in the effectiveness of the sharing economy. These biases lead sharing economy services to be both more expensive and less available in low-SES areas and suburban areas than in high-SES and high-density urban areas. Moreover, SES and race/ethnicity are often strongly correlated in many parts of the world, and I observed that, at least in the city of Chicago itself, areas in which the population is more white (non-Latino) have better access to sharing economy services.

Overall, this chapter provides evidence that (1) in the sharing economy, geography matters, and geographic principles should be strongly considered in examinations of the sharing economy and (2) one way in which the importance of geography manifests is that key geographic principles interact with common design decisions in sharing economy platforms to create important biases in the effectiveness of the sharing economy. As discussed above, engaging with both of these takeaways can lead to 'geosociotechnical' design improvements in sharing economy platforms that reduce these biases, among other benefits.

Chapter 4

Geographic Behavior as Spatial Interaction: A System-Level View

One of the key findings from the previous chapter found was that distance plays an important role in where sharing economy crowd members are willing to provide service (e.g. perform TaskRabbit tasks, or pick up rides). Put another way, crowd members in the sharing economy incorporate the distance to the task in their decision-making. Based on what is known about distance decay, this finding is intuitive. After all, a TaskRabbit crowd worker cannot help build IKEA furniture without being 'local' to where the furniture is.

Similarly, VGI content production has also been broadly assumed to be local. That is, despite not needing to travel to e.g. edit geotagged Wikipedia articles, the assumption is that people contribute information nearby. This idea was so fundamental to Goodchild's [58] initial conception of VGI that he suggested that "the most important value of VGI may lie in what it can tell about local activities in various geographic locations". Goodchild's intuition is logical. It is probably easier to contribute nearby, both because of proximity and because someone is more likely to be knowledgeable about their home area.

However, as I discussed in detail in [Chapter #sec:chptrw], recent work suggests that this 'localness assumption' [82] is largely untrue for peer-produced VGI. While the localness assumption postulates that people make all of their VGI contributions nearby, these findings in from prior work suggest that there is a lot more variability in people's local contribution behavior. This tension between common localness assumptions and the reality of peer-produced VGI – as well as how this variability influences our understanding of VGI content and VGI production behavior – calls for an alternative model of how VGI is produced.

In this study, I propose and evaluate such an alternative model. My approach is based on 'spatial interaction models', long used as a means of understanding geographic interaction patterns in the social sciences. I compare these models against two baselines that evaluate opposing perspectives on VGI production. The first (my local production baseline) directly models the localness assumption, i.e. that where people contribute is decided by distance alone. The second (my distance is dead baseline) represents the complete inverse of the localness assumption, i.e. that people merely contribute based on the attractiveness of the contribution location, and distance has absolutely no effect. Conceptually, gravity models merge the ideas of both these two baselines – distance impacts where contributions occur, but attractiveness of the location also informs where contributions occur, and may counteract the effect of distance. As a fundamental part of this evaluation, I follow recent calls to increase the ecological validity of social computing research by moving beyond single-community analyses to studies that simultaneously consider multiple communities (e.g. [5,93]). Specifically, I look at three different VGI platforms: Wikipedia, OpenStreetMap, and eBird.

My gravity models yield an important theoretical insight: I find that spatial interaction models perform meaningfully better than either baseline and describe more than 90% of the 'VGI flows' from one region to another in some cases. I see strong performance of these models across hundreds of different types of content from my three peer-produced VGI platforms. My results show a large degree of variation in how local content is, and suggest that distance does play a role – people's contribution rates decrease as the places they contribute about get further away. However, the effect of distance clearly varies between different content types, and some are much more local than others.

My results also have implications for multiple types of VGI stakeholders and suggest important areas of future work. In particular, I discuss how my findings problematize some approaches to VGI "editathons", might suggest mechanisms for understanding demographically-linked coverage biases in VGI, and help to define where local perspectives may be present in VGI and where they may be absent.

4.1 Related Work

Beyond two of the topics discussed in [Chapter #sec:chptrw] – localness in VGI, and geographic variations by socioeconomic status and population density – my research here is also informed by applications of gravity models in the social sciences and the various ways VGI content is used. Below, I describe these two latter areas and how they informed my work.

4.1.1 Spatial Interaction Models in the Social Sciences

Modeling spatial interactions between regions has a long history in the social sciences, particularly in the field known as economic geography. Gravity models, which date back to 1948 [141], are the most common approach. Gravity models aim to capture the interaction between two regions based on the 'gravitational' pull of each region and the friction of distance between the two regions. In the almost 70 years of their existence, gravity models have been used to effectively model a wide variety of spatial phenomena, primarily in two domains: (1) transportation of goods and people (e.g. international wheat transactions [88], inter-state gun trades [84], international meat trades [89]) and (2) communication patterns (e.g. inter-city [90] and international phone calls [36]).

My research is directly motivated by the effectiveness of gravity models in explaining these phenomena. As I describe below, I hypothesized that a contribution to one region by a VGI contributor based in another region could be modeled similarly to a product (e.g. meat, wheat) being exported from one region and imported to another. In other words, even with the complex dynamics associated knowledge production in online communities, I believed the spatial dynamics of VGI contribution can be thought of as transfers of units of information from one region to another. My results indicate that this hypothesis was supported.

Beyond the critical implicit value of better understanding VGI production processes, my work also highlights that gravity models may be useful in computing domains further afield from their typical applications in transportation and communication. In this case, I identify that gravity models are surprisingly effective at capturing a knowledge production relationship between a person and a place as mediated by complex online community dynamics. I return to this point in the Discussion section.

4.1.2 Applications of VGI Stakeholders

There are three main types of applications of VGI: (1) direct consumption by readers/users, (2) scientific studies, and (3) intelligent technologies and other systems.

The success of these applications tends to be closely tied to the coverage and quality of their underlying VGI datasets. This raises the stakes of the applied implications of my work, in which I help to explain geographic variation in VGI coverage and quality.

With regard to direct consumption, geotagged articles are some of the most persistently popular articles on Wikipedia [72] and OpenStreetMap powers many prominent mobile maps applications like Apple Maps [80]. In this case, VGI coverage and quality have a direct impact and one that is highly visible to the public. Scientific applications of VGI that rely on the coverage and quality of VGI include the effects of tourism on water quality [86], detecting the epicenter of earthquakes [145], and others discussed in more detail by Venerandi et al. [166] and Wood et al. [172]. VGI has also become a key input to many intelligent technologies, like geolocation inference techniques (e.g. [26,83]), among many others [25,47]. Indeed, geolocation inference (e.g. of users and documents), is a domain in which coverage and quality has verified importance [81]. Further still, there is some evidence [75] that VGI can impact economic growth.

4.2 Methods

4.2.1 Datasets

One of the key findings in previous work is that the spatial interaction dynamics in VGI may differ based on the community (e.g. OSM vs. Wikipedia, [69,122]). Therefore, to more robustly evaluate spatial interaction dynamics in VGI production, I examined three VGI platforms: Wikipedia, OpenStreetMap, and eBird.

Further, contribution in each of these communities is a heterogeneous process; that is, some types of content in a given community may support different types of spatial interaction behavior than other types of content. For example, editing Wikipedia articles about national parks (which are globally known) may have a different spatial interaction profile than editing Wikipedia articles about elementary schools (for which information is more locally concentrated). A similar dynamic may exist in OSM with respect to, for example, encoding state borders versus tracing and labeling ("tagging") specific buildings.

Therefore, within each of my three platforms, I systematically examine contributions at the level of the content type. I analyze the effect of spatial interaction for each content type individually, as well as at the overall platform level. Example content types include articles about schools for Wikipedia (as defined by WikiProjects), residential buildings for OSM (as defined by tags), and bald eagles (*Haliaeetus leucocephalus*) for eBird (as defined by species). In total, we have 561 different content types, with 101 content types in Wikipedia, 192 content types in OpenStreetMap, and 268 content types in eBird.

As is common in VGI research (e.g. [81,99,116,137]), I focus on data from a single study area: the continental United States. I explore how my research can be expanded to other study areas in my discussion of future work below.

I next describe in more detail the datasets I developed for each of my three VGI platforms.

Wikipedia

My Wikipedia dataset focused on contributions to geotagged Wikipedia articles. A contribution can be anything from creating new article text to fixing a typo. I queried the English Wikipedia public database for all contributions by registered users to geotagged articles that were saved in the year between Oct 2015 and Oct 2016 (resulting in 3.5 million edits). I then limited the data to articles located within the continental United States, leaving 644,480 total contributions.

For each edit, I used its associated WikiProject as the content type (approximately 4% of the contributions had no WikiProject assigned and were excluded). A WikiProject is a self-organized group of people working to improve Wikipedia content on a certain topic. For instance, WikiProject Schools is a group of contributors who work to curate school-related content in Wikipedia. Each content type served as an independent dataset in my analyses. I excluded the smallest WikiProjects (with fewer than 1,000 contributions) in order to ensure sufficient data to fit a model, resulting in 101 Wikipedia content types.

OpenStreetMap

My OpenStreetMap dataset focused on node (point) contributions. An OpenStreetMap node may be a tree, a traffic circle, or a label point for a building. Entities like buildings or roads are normally represented by 'ways', logical groups of nodes. However, attributes of the way (e.g. height of the building) are not associated with individual nodes, and therefore they would not be included in my dataset. I used the full history of OpenStreetMap nodes in the continental USA through February 2014. I excluded nodes that did not have one of the 1,000 most-popular tags (e.g. to eschew typos). From this set of nodes, I than randomly sampled 2,000,000 nodes for analysis.

I used the tags of a node to define its content type(s). A tag consists of a key-value pair, with only one value allowed per key. For example, a 'natural=tree' tag on a node denotes that this node represents a tree and 'junction=roundabout' denotes a traffic circle node. As such, all tree contributions were defined as one content type, all traffic circles as another, and so on. As noted above, I excluded the smallest content types (with fewer than 1,000 contributions) in order to successfully fit my models, resulting in 192 total OSM content types.

eBird

eBird is an observational citizen science project in which a contribution is a bird sighting. As opposed to Wikipedia and OSM, in which one does not need to be physically present in order to contribute, eBird contributors need to be at or near the location of their contributions. This geographic proximity requirement makes eBird an interesting comparison point to Wikipedia and OpenStreetMap. As is shown in the Results section, this comparison point will prove to be a valuable reference for understanding contributions in OSM and Wikipedia.

To gather an eBird dataset, I began with the full history of eBird observations through April 2015. I then randomly sampled 2,000,000 observations from this data set, and again limited this data to the continental United States, resulting in 1,573,798 total observations. To understand spatial interaction by content type, I defined content type by sightings of a particular bird species. Again, I excluded the smallest species (with fewer than 1,000 observations) to ensure successful model fitting, resulting in 268 eBird content types.

4.2.2 Defining The Geographic Origin of Contributions

Prior to modeling spatial interaction processes in peer-produced VGI, I first had to verify my three datasets actually are largely non-local. To do so, I needed to define two properties for each contribution: (1) the local (home) region of its contributor (i) and (2) the region in which the contribution was made (j). I also had to determine the spatial scale at which a region would be defined. For this, I used the scale of U.S. counties, a common choice in VGI analyses [80].

Determining the county where a contribution is made (j) is straightforward: I use the geotag attached to each contribution and perform a reverse geocoding operation. Determining the home region of a contributor (i), on the other hand, is significantly more complex. Unlike social media user profiles, contributors to my VGI repositories have no widely-used means by which they state their home location. Although some contributors do so voluntarily in venues like Wikipedia user pages, participation is low and available only in certain repositories. Similarly, prior work has used IP address geolocation [66,69] when studying Wikipedia, but contributor IP is not available in all of my repositories (and would likely suffer accuracy problems at the county scale [131]). Moreover, even within Wikipedia, IP addresses are only available for anonymous editors [150].

As such, it was necessary to do home location inference to determine the county i of each contribution. Fortunately, this is common, and numerous solutions exist [82]. I adopted the home location inference technique known as plurality, which defines a contributor's home region (county) as the region (county) in which s/he has made the plurality of their contributions; this technique has been used in a number of VGI and VGI-related studies (e.g. [70,82,120]). I excluded contributors with fewer than 5 contributions, in order to be confident in the inferred county. Following recent calls for researchers to validate home location results across multiple inference techniques [82], I also calculated the home location of each contributor using the geographic median approach [26,82,83]. I found that well over 90% of identified home counties were identical across the two approaches, giving us high confidence that both approaches would lead to very similar results in a spatial interaction model. Therefore, I used the plurality approach in my analysis.

To verify that my datasets violate the assumption of being local, I examined the percentage of contributions in which the contributor's home county i is not equal to the contribution county j. The results of this simple analysis made clear that the large degree of non-local contributions identified in prior work is replicated in my datasets: only 26% of Wikipedia contributions, 23% of OSM contributions, and 57% of eBird contributions occurred in the plurality-defined home county of their contributor.

These findings justified my further exploration of spatial interaction as an alternative model of VGI production. Below, I describe how I performed these analyses using gravity models.

4.3 Gravity Modeling

4.3.1 Intuition

Spatial interaction models seek to explain the relationship between two locations (i and j) using the distance between them and their individual attributes. More formally, they ask the following: how does location i interact with location j, based on the attributes of i, the attributes of j, and the distance between i and j? Gravity models specifically assume that these relationships can be modeled through an analogy to the basic formula for gravity in the physical world [155]:

$$F_{ij} = \frac{M_i \ M_j}{D_{ij}^2}$$

When considering the physical gravitational pull two objects have on one another, the mass of each object describes their attraction to one another, which is moderated by the distance between them. The gravity model takes this intuition, and applies it to interaction between two regions (i, and j) across space. The amount of interaction – the dependent variable – is commonly represented as F_{ij} (or 'flow between regions'). The 'mass' variables (M_i for region i, and M_j for region j) are typically the population of the area, GDP of the area, or other 'attraction' attributes (e.g. [16,84,88,89]). Because gravity models are intended to help understand interaction, it is critical that the mass variables incorporate both potential outflow (leaving i) and potential inflow (entering j). For instance, using GDP for both mass variables (M_i and M_j) is common for physical processes like international meat trading [89], because it accounts for both exports (potential outflow from i) and imports (potential inflow to j). The final variable in a gravity model, distance (D_{ij}), is often operationalized as geodesic (straight-line) distance between two regions.

Airline travel is a common intuitive example for understanding how these variables relate to one another. Consider the case of three cities: New York City, Los Angeles, and Bangor, Maine (a city of about 33,000 residents), with the mass variables set to the population of each city. In this case, population operationalizes both the potential outflow from a city and the potential inflow to a city (more people usually means more business and personal travel, etc.). New York City and Los Angeles are on opposite coasts of the United States, and thus have a large D_{ij}. However, many people fly back and forth between New York City and Los Angeles due to the large 'attraction' (i.e. large product of masses) between the two cities, which overcomes the large distance (large D_{ij}). On the other hand, despite the much smaller D_{ij} between Bangor and New York City, the tiny mass of Bangor counteracts the shorter distance, and many fewer people fly between Bangor and New York City.

4.3.2 Applying Gravity Models to My Datasets

In the traditional formulation of the gravity model (above) the friction of distance is defined as -2, and the weights of M_i and M_j are held constant, predefining the degree to which they affected F_{ij}. Because of this, the traditional form was generalized and transformed to a log-linear OLS model (below) [44]. The friction of distance was no longer held constant (at -2) , and M_i, M_j, and D_{ij} all became independent variables, predicting the dependent variable F_{ij}.

$$F_{ij} = \frac{M_i^{\beta_1} \; M_j^{\beta_2}}{D_{ij}^{\beta_3}}$$

$$ln(F_{ij}) = \beta_0 + \beta_1 \ln M_i + \beta_2 \ln M_j - \beta_3 \ln D_{ij}$$

This straightforward approach, however, causes problems when the variables contain zeroes. After all, the natural log of zero is undefined. Further, the common practice of adding a small constant (such that there are no zeroes) produces biased estimates [44]. To address this problem, I employ one of the most common solutions (recommended by [44]): fitting a Poisson linear regression, which does not risk biased estimates in the scenario mentioned above. It is common to take the natural log of all independent variables, so I use this strategy in my models [44].

The first step in operationalizing gravity models is defining i and j. I use the same definitions as before: i is the 'home' region of a contributor, and j is the region in which a contribution is made. Every inter-county interaction is thus modeled as someone based in county i contributing information about county j, aggregated over all contributions in a content type. In other words, if i = Wayne County, Michigan and j = Baltimore County, Maryland, the goal of the models is to accurately predict the number of contributions about places in Baltimore County made by people whose home county is Wayne County (F_{ij}). To make these predictions, I define M_i to be the number of contributors from county i (e.g., Wayne County) that make contributions elsewhere (potential outflow), and M_j to be the number of contributors from anywhere that make contributions (potential inflow) into county j, (e.g. Baltimore County. I follow common practice, and consider D_{ij} to be the geodesic distance between i and j. I make these predictions separately for each content type in each repository. In other words, I run a separate gravity model for each of my 561 (101 Wikipedia + 192 OSM + 268 eBird) content types.

Traditionally, predictions for intra-regional flow are excluded when constructing gravity models, for two primary reasons. First, intra-regional flows are not intuitive for many physical processes, e.g., we don't speak of a country trading wheat with itself. Second, it is not intuitive what the distance from a region to itself ought to be, and using zero can be problematic for reasons mentioned above.

However, for my purposes, these reasons do not hold. First, in my data, non-trivial quantities of VGI content is locally produced (intra-regional) – as much as 43% in the case of eBird. Second, because I follow more recent common practice and implement my gravity models as Poisson regressions, defining a small intra-regional distance will not cause biased estimates. Therefore, I adopt two approaches for intra-regional distances defined in the literature [55,64]:

- constant 1 km for every region, and
- $\frac{1}{2}\sqrt{regionalArea}$.

Below, I compare results from these approaches, which I term constant-distance and regional-distance, respectively. I evaluate both options in any models that include distance as a variable – my baselines, and my gravity models.

Because I build so many models, statistical intuition suggests that a small portion of any significant results would be due to chance. However, as I will see below, my overall results are sufficiently widespread that they are quite robust against the occasional Type I error.

4.3.3 Contextualizing Gravity Models

To provide context for my evaluation of how well gravity models describe VGI production, I also construct two baselines against which to compare my gravity models. To make sure my baselines and my gravity models are comparable, I construct both baseline approaches with a Poisson regression. As I do in my gravity models, I ensure that all variables I include in these models are log-scaled.

My first baseline is a set of distance is dead models. This represents an alternative interpretation of VGI production, and postulates that the distance between i and j is irrelevant. This model seeks to predict F_{ij} with only M_i and M_j as independent variables, for each content type. If indeed these models perform better than my gravity models, it will indicate that distance is unimportant in VGI production.

Second is my local production baseline, which is composed of two different sets of models, one for each approach to intra-regional distances I discuss above. This baseline represents a second alternative interpretation of VGI production, that Goodchild is correct and contributions only happen nearby. More formally, this baseline postulates that distance between i and j is the only meaningful factor in why contributions flow between i and j. This model seeks to predict F_{ij} using only D_{ij} as an independent variable, for each content type. If these models perform better than my gravity models, it will indicate that attributes of i and j have no bearing on VGI production.

To properly evaluate if gravity models are even effective at characterizing VGI production, and because the literature suggests two alternatives for incorporating intra-zonal predictions into gravity models, I construct five separate models, across hundreds of different content types. Specifically, I construct one distance is dead baseline, two instances of my local production baselines, and two instances of gravity models. I then compare all five, and evaluate which are most successful at describing peer-produced VGI.

4.4 Summary of Methods

To summarize: - M_i is the number of contributors from county i, M_j is the number of contributors who contribute in county j, and D_{ij} is geodesic distance between i and j. - I construct all models as Poisson regressions, following the recommendations of 9. - Because some VGI contributions occur in the same county where their contributor lives, I evaluate two approaches for defining D_{ij} when i and j are the same county: 1 km, and $\frac{1}{2}\sqrt{regional\_area}$. This is true for both my local production baseline, and my gravity models. - I construct five different models for each of the 561 different types of content.

4.5 Results

I now turn to results, first evaluating if my gravity models are even a reasonable approach to understanding VGI production, and then engaging in a deeper exploration into the effect of distance in spatial interaction models.

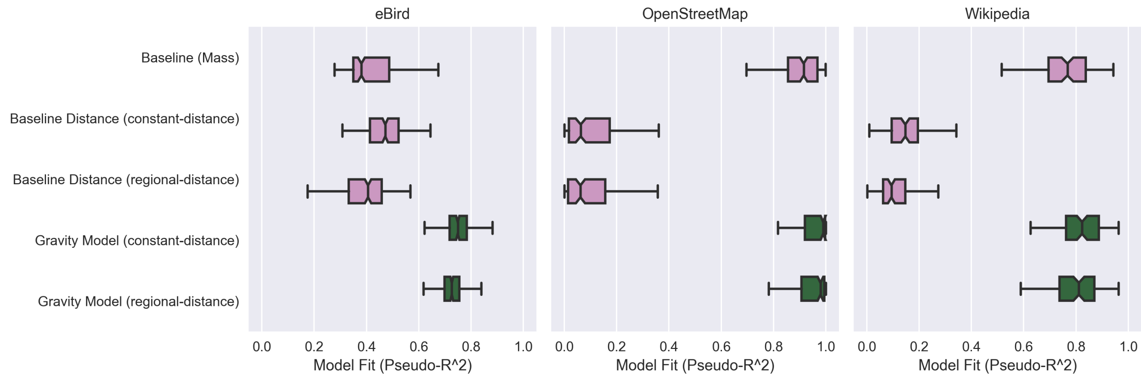

Figure 4.1: On the x-axis, each plot shows the model fit (pseudo-R^2). The y-axis shows each of the five models I are evaluating. Each distribution excludes outliers. The top three are baseline models, and the bottom two are gravity models.

4.5.1 Evaluating Model Fit

Informed by their long history and theoretical underpinnings, I believed that gravity models would likely be effective descriptors of VGI production, though this was by no means guaranteed. Therefore, my first task was to evaluate this conjecture. I did so by comparing the pseudo-R^2 values from each content type, across my two local production baselines, my distance is dead baseline, and both instances of my gravity models.

Figure 4.1 shows the distributions of pseudo-R^2 values along the x-axis (as measured by the pseudo-R^2 metric suggested in [28]). The y-axis lists each of my five models, and each chart indicates a different platform. The top three models are baselines, and the bottom two show my two different instantiations of gravity models .

To evaluate differences between distributions, I use a notched boxplot. When the notches do not overlap, the differences can be considered statistically significant [111]. However, because I are testing the significance of differences between groups of model output, the distributional assumptions are unclear and significance should be interpreted with some caution. Effect sizes, on the other hand, do not have this issue.

Figure 4.1 shows that all eight gravity models perform better than the distance is dead or local production baselines (seven of them significantly so). Examining the medians of each distribution, the general trend is clear: spatial interaction models are very successful at describing VGI contributions with median pseudo-R^2s as high as 0.99 in some cases. eBird has the lowest median pseudo-R^2s, at 0.74 and 0.72 for the constant-distance gravity models and regional-distance gravity models respectively. Wikipedia content types show better fits than eBird content types, with median pseudo-R^2s of 0.82 and 0.1 for the constant-distance and regional-distance gravity models. OpenStreetMap content types tend to show the highest median pseudo-R^2s, at 0.99 and 0.98 (for constant-distance and regional-distance, respectively).

Focusing on the baseline model distributions in more detail, I noticed some striking differences between platforms. In eBird, a platform where contributors must travel to make contributions, the distance is dead and local production baselines tend to be much more similar in terms of model fit. This is in contrast to OpenStreetMap and Wikipedia, where the distance is dead models fit substantially better than the local production models.

Comparing the pseudo-R^2s of the baselines to those of the gravity models provides additional insight into how the mass and distance affect the model fit of my gravity models. My gravity models fit quite well, and in OpenStreetMap and Wikipedia, the distance is dead baseline models also fit quite well. This suggests that in the 'wiki' platforms (Wikipedia and OpenStreetMap) where "armchair editing" is possible, the mass variables drive a substantial portion of the gravity model fit. Put another way: in OpenStreetMap and Wikipedia, the content type and its attraction dynamics between regions matters much more than geographic distance for how contributions flow between regions. However, in eBird both distance and content type matter for where people contribute. Someone may contribute to Wikipedia about golf courses in Florida, regardless of where they live. Conversely, but while some bird species may be more interesting, any bird-sighting still requires travel in order to contribute.

The high-level conclusions of my results at this stage are clear. First, distance is a substantial factor in how well gravity models perform in systems like eBird, where contributions are inherently a physical process. Second, even in OpenStreetMap and Wikipedia – where contribution is not an obviously physical process – most of my gravity models still have more explanatory power than my distance is dead baselines. Further still, while the distance is dead baseline indicates that the mass variables play a substantial role in model fit for OpenStreetMap and Wikipedia, the addition of a distance variable does improve model performance. This means in spite of fact that a contributor does not need to move at all to edit Wikipedia or OpenStreetMap, there is a degree to which contributions in these platforms are impacted by how far away they are from a contributor. Holistically, these findings suggest that VGI production in these platforms indeed can be understood as a gravity model spatial interaction process.

4.5.2 Interpreting My Models

My results in the previous subsection indicate that my gravity models perform better than my baselines , and effectively explain a large portion of the spatial contribution decisions of VGI contributors. Therefore, I now limit my discussion of results to my gravity models. While I present both my constant-distance and regional-distance

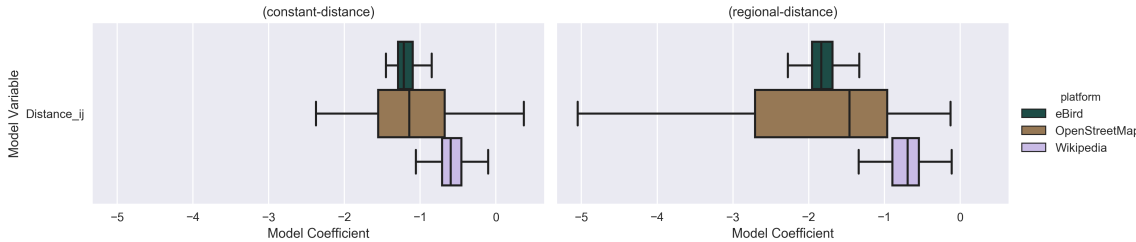

Figure 4.2: These plots show my friction coefficients for each platform. On the left are my constant-distance models, and the right shows my regional-distance models. Both exclude outliers.

models, I will focus this discussion of results around constant-distance gravity models, because all three perform significantly better than the distance is dead baseline (whereas only two regional-distance models outperform the distance is dead baseline). I exclude 22 types of content with variables that are not significant (predominantly from OSM), to ensure all distributions are comparable.

I focus specifically on the D_{ij} (friction) coefficient , in order to shed light on the degree to which contributions are likely to be local. Recall that the more negative a friction coefficient is (further left in Figure 4.2), the stronger friction effect exists.

Figure 4.2 shows the distributions of my D_{ij} coefficients. On the left are the coefficients from my constant-distance instance of a gravity model, and on the right are the coefficients from my regional-distance models. Each boxplot represents a different platform. eBird is on top, OpenStreetMap is in the middle, and Wikipedia on the bottom.

Immediately visible in Figure 4.2 is that there are clear differences in the friction coefficients between eBird and Wikipedia. Content types from both Wikipedia and eBird are quite clustered together, and the platforms themselves center around different points on the friction of distance spectrum. In some cases, eBird has species that have similar friction coefficients to some Wikipedia content types, but the overlap between these distributions is small. Surprisingly, content types from OpenStreetMap have a much wider overall distribution – some are much closer to eBird content types, and others are much closer to Wikipedia content types. Put simply, distant contributions are much more expensive in eBird than Wikipedia, and OpenStreetMap contains some types of content that have similar friction coefficients to fundamentally physical processes like eBird.

To explore these friction coefficients in more detail, I now discuss some examples from each platform, moving from left to right, and from top to bottom (highest friction coefficient to lowest, eBird to Wikipedia).

eBird

Near the high-friction end of the spectrum is the purple finch (*Haemorhous purpureus*) sighting content type, with a friction coefficient of -1.45. Succinctly, many contributions of this bird would be highly 'local', or nearby the contributors' home. One hypothesis for why this might be the case is that the purple finch has a very large range, spanning most of the eastern United States. This means that while eBird contributors might upload reports of purple finch sightings on an everyday basis, while traveling it perhaps might be somewhat boring to continue to upload sightings of the same species when there are novel species available.

On the low end of the friction spectrum for eBird we see the ladder-backed woodpecker (*Picoides scalaris*) with a friction coefficient of -0.85. The ladder-backed woodpecker has a range that contains very popular tourist areas in the United States (e.g. Las Vegas, the Grand Canyon). As such, one hypothesis is that this bird is often reported while eBird users are on vacation in these areas, thereby making these reports distinctly non-local (i.e. having a small friction of distance).

OpenStreetMap

The OpenStreetMap content types are much more heterogeneous, and have a much wider distribution than either of the other two platforms. Near the left-hand side of the OSM distribution is 'addr:city=San Diego', which has a friction coefficient of -2.58. Because San Diego is a city of 1.5 million people (has a large mass), it is likely that this friction coefficient reflects a highly-local bulk import done by a resident of San Diego. This would cause a large number of highly local contributions, and thus a high friction of distance. Another interesting example is 'building=residential' with a friction coefficient of -1.63. This tag might have a strong friction of distance for a simple reason: it is quite difficult to determine whether a building is a residential building or a commercial building from far away in many cases; one must know the area (especially in large cities, where there might be mixed-use development).

Wikipedia

What is initially clear is that the Wikipedia friction coefficients tend to be quite similar to one another. Starting from the left-hand side of the distribution is WikiProject Politics, with a friction coefficient of -1.05. Contributions to WikiProject politics decrease nearly linearly as the places they contribute about get further away. Intuitively, it seems likely that people are less interested or less aware of the details of politics that are further away from them – as the common saying goes: "all politics is local". On the other end of the spectrum is WikiProject Golf, with a friction coefficient of -0.1. This friction coefficient is quite low. One reason this may be the case is the topic itself – to participate in WikiProject Golf, a contributor would likely be highly motivated by Golf as a topic, and may treat golf courses as vacation destinations as well. The combination of being highly motivated and traveling to play Golf would lead to a quite low friction coefficient.

Summary and Generalizable Conclusions

To summarize my results, I found that in all of my content types, gravity models are very effective at describing VGI production. Additionally, I found that contributions in Wikipedia and OpenStreetMap are largely driven by attraction between regions, whereas distance is much more important when describing eBird contribution trends. Further, in two of my platforms (eBird and Wikipedia), the friction coefficients are quite consistent, indicating that some platforms facilitate a specific 'style' of spatial interaction. In contrast, in OpenStreetMap the content types span a large range of friction coefficients.

4.6 Discussion

My results have implications for a number of constituencies and research areas. Below, I outline these implications in more detail.

4.6.1 Implications for VGI Contributors and Platform Managers

My model fits align well with an idea implicit in the editing ethos of some large VGI communities – the belief that distance has minimal impact on VGI contribution. For instance, Wikipedia states "anyone can edit almost every page" [170], and OpenStreetMap's introductory documentation says "You can map from your armchair" [126]. From a purely technical perspective, it is just as easy for a person who lives in e.g. Montreal to log into Wikipedia or OSM and contribute information about McGill University as it is for that person to contribute information about, for instance, Nazarbayev University in Kazakhstan.

This raises a key question: what limits some content types from being advantaged by the affordances to map anywhere or write articles about anywhere from "armchairs"? A number of factors likely are responsible. For example, physical world processes are still highly correlated with knowledge about a region, and knowledge about a region can help one more easily write a Wikipedia article, do OSM mapping, or see and recognize a specific bird. Regional boosterism may also be at play, causing people to build up information about known locations. However, future work should seek to examine the presence and strength of these and other factors. One approach might be a qualitative study to understand where people choose to contribute, and why. This would help shed light on some of the mechanisms that underpin the large attraction processes we see in the results.

My work has several implications for the design of VGI communities and platforms. My results present challenges for a particularly common means by which VGI communities attempt to address coverage issues: "editathons". Editathons are usually in-person events and are typically held in urban areas where many potential new contributors can attend. My results show that for high-friction content types, these types of in-person contribution drives will not affect the variations in coverage. To do so requires localized contributors, and it is unlikely that editathons occur in places where contributors are needed most. This is especially troubling as editathons are often funded by the cash-strapped organizations that operate VGI platforms (e.g. the Wikimedia Foundation). My results suggest that organizations like the OpenStreetMap Foundation may want to redirect some of their resources towards efforts that work towards these goals.

4.6.2 Implications for Coverage Biases

More generally, my work may help to reveal mechanisms for the coverage biases linked to socioeconomic status, the rural/urban spectrum, and other demographics. One hypothesis as to the mechanisms for this coverage variation is that "self-focus bias" is playing a role [68]. That is, people are contributing about places where they have lived, and, given the demographics of VGI contributors (e.g. [56]), it is likely that they will have lived in higher-SES areas and urban areas. My results provide a direct means of testing this hypothesis: If this is true, then content types for which the friction of distance is high should exhibit more coverage bias then content types for which armchair mapping is more common. Evaluating this hypothesis is an immediate opportunity for future work.

My results also highlight a hypothesis for a potential second cause of these biases: preferential attachment. It may be that high-SES areas and urban areas were some of the first areas to be covered in these datasets, thereby making them more "attractive". Because of this attraction – and the importance of attraction shown in my baseline models more generally – these areas' early leads in coverage became effectively permanent. More generally, my baseline models suggest that, at least for OpenStreetMap and Wikipedia, preferential attachment may be a particularly potent force. Testing this "geographic preferential attachment" hypothesis is also an excellent direction of future work.

4.6.3 Implications for Algorithms

Hecht and Gergle showed that AI systems that use VGI for world knowledge can adopt the perspectives of their underlying VGI datasets [54]. Since my results suggest that certain VGI content types will innately contain more local perspectives than others, this suggests that VGI-based AI systems that rely on certain types of data may innately be biased towards local or non-local perspectives.

This is particularly an issue for Wikipedia-based systems given the number of AI systems that use Wikipedia data. My results suggest that these systems will adopt "local perspectives" for high-friction topics like politics and "armchair mapper perspectives" for low-friction topics like golf. Investigating this hypothesis in well-known Wikipedia-based AI systems would be a fruitful direction of future work.

4.6.4 Implications for Human Consumers

The exact same biases that may affect algorithms with respect to non-local and local perspectives will also affect human consumers of VGI. For instance, my results suggest that Wikipedia content about golf courses will be less local than its content about politics. This highlights a number directions of future work. Two of the most interesting might be (1) building tools that can surface the fact that local perspectives may not be present for certain content types and (2) using this surfacing to perhaps incentivize more contributions from the local area (e.g. using a prompt like "This article about your local golf course was written entirely by non-locals. Do you have any local expertise to add?")

4.6.5 Implications for Gravity Models and Social Computing

As discussed above, the predominant use of gravity models have in HCI and social computing contexts have tended to be in the traditional gravity model domains of transportation and communication (using datasets of interest to the HCI and social computing communities). My results suggest that gravity models are also quite useful for understanding processes further afield from transportation and communication.

At the very least, this work suggests that researchers who are examining the role of distance in a geographic HCI [71] process consider utilizing gravity model techniques. The primary challenge of moving beyond simple distance involves operationalizing the mass variables, and my discussion of my implementation of mass can provide a reference point along these lines.

4.7 Conclusion

This work established the value, both theoretical and practical, of understanding non-local contributions in VGI repositories. I showed that VGI contributions can be modeled effectively using spatial interaction techniques, and gravity models in particular. I also explored the implications of these findings for my understanding of VGI, for stakeholders currently managing large VGI communities, and for the development of future VGI platforms.

Chapter 5

Geographic Biases are "Born, Not Made": Individual Spatiotemporal Behavior

In the previous two chapters of this thesis, I have presented two studies that predominantly focus on geographic behavior from the system-level, using robust geostatistical analysis approaches. However, establishing crowd members' geographic behavior trends is distinct from understanding geographic crowd behavior at an individual-level.

Here, I do turn my focus to the individual. In particular, I take a spatiotemporal view on how contributors choose to focus their contribution efforts. Contributor choice is a fundamental characteristic of peer production: it differentiates peer production from other forms of crowdwork [9] and may even be necessary for the success of the peer production content generation model [9]. Indeed, the ethos of contributor autonomy is so foundational in peer production that, for instance, the introductory documentation of OpenStreetMap, states that "anybody can enter anything she wishes" [127].

Because of the importance of peer produced content and the role of contributor autonomy in producing that content, researchers have long sought to understand and model contributor focus in various peer production contexts (e.g. [35,62,128,134]). One common thread in this research involves studying how contributor focus evolves over the lifespan of a contributor (e.g. [6,128]). In other words, this research examines contributor focus through a temporal lens.

While a temporal lens is sufficient to understand contributor evolution in many peer production contexts, in geographic peer production – e.g. contributing to OpenStreetMap and editing geotagged Wikipedia articles – a purely temporal lens cannot detect another critical type of potential focus evolution: that which unfolds spatially. For instance, while it is useful to know that an OpenStreetMap contributor is increasing her/his contribution rate, it is also important to understand where and in which types of places the user is contributing, and how this changes over time. Among other applications, such knowledge can provide critical insight into the troubling coverage biases that have been observed in peer produced geographic datasets (e.g. on socioeconomic and urban/rural lines [61,80,150]).

In this research, I extend the literature on temporal focus evolution to geographic peer production with an exploratory analysis that examines contributor focus with a spatiotemporal lens. My work uses OpenStreetMap – the world's largest peer produced geographic dataset – as a case study and centers around two basic research questions adapted from the temporal literature [128]. First, I ask:

(RQ1) How does contributors' geographic focus change over time?

To address this question, I operationalize four geographic contribution metrics and explore if and how they change over time. Overall, my results suggest that contributors are broadly consistent in their geographic editing behavior over the course of their contribution lifespan, although there are some deviations from this trend. Further, the consistency is of a particular nature: people tend to consistently edit in relatively specific geographic areas.

These results recall the findings of one well-known GROUP paper that examined contributor focus with a temporal lens, finding that Wikipedia power editors have different editing behavior than other users from day one of their editing career, i.e. that power editors are "born, not made" [128]. In my study, I observed this "born, not made" dynamic in a very different peer production context: the geographic editing behavior of OpenStreetMap editors (although I observe a somewhat softer version of the dynamic).

My spatiotemporal approach also advances understanding of mechanisms behind a second (and concerning) trend that has been observed in the literature: geographic biases in peer production. In the face of peer production's immense success – which is predicated on the idea that "anyone can enter anything she wishes" – recent research shows that urban and wealthy areas receive better geographic coverage than rural and less wealthy areas [80,99]. While prior work characterizes these biases, few have studied their root causes. Thus, my second research question asks:

(RQ2) Can the spatiotemporal evolution of contributors' focus help to explain systemic coverage biases?

My exploratory results suggest that most contributors are "born" urban-focused and wealthier-focused and stay that way. In other words, for most editors, the proportions of edits in rural and poor areas are consistent and consistently low across contribution lifespans. I also find that the few editors who do consistently focus in rural and poorer regions tend to have lower survival rates, exiting OpenStreetMap sooner than their urban- and wealthier-focused counterparts

My study makes four primary contributions:

- I explore the geographic contribution behavior of OpenStreetMap editors over time and observe that most editors exhibit similar behavior across their entire contribution lifespans. Thus, for many people, I find evidence that geographic editing behavior is “born, not made”.

- I show how this consistent contribution behavior applies also to the types of regions people edit. In other words, I find some evidence that geographic biases also are “born, not made”.
- These focus biases are amplified by a survival bias – people who focus in rural and high-poverty areas tend to contribute for shorter periods of time.
- While I did observe a small group of people who focus primarily in rural or high-poverty areas, they produce only a small portion of OpenStreetMap content.

5.1 Related Work

My work here builds primarily on prior work in three areas: (1) peer production contributors’ geographic contribution behavior, (2) temporal evolution of contributor behavior, and (3) geographic biases in peer production. Below I situate my work relative to each of these areas.

5.1.1 Contributors’ Geographic Patterns

The literature examining contributor geographic patterns falls broadly into two categories: where contributors focus and the geographic ranges of contributors’ work. My research extends these two categories of prior work by considering the evolution of these types of geographic trends over time. Below, I describe each category in more detail and put each in the context of my work.

5.1.2 Where Contributors Focus

Several different studies have sought to understand and characterize the geographic focus of contributors to peer production platforms. For instance, Panciera et al. [130] examined geographic trends in the Cyclopath platform, an early bicycling-centered community. In particular, they found that "Cyclopaths" (defined as the top 5% of contributors) had geographically constrained contribution regions, even within the relatively small area in which Cyclopath operated. Zielstra et al. [176] described the geographic extents of 13 OpenStreetMap contributors and show a method of characterizing which contributions are a part of a contributors' 'home location', and which are not. They found that the contribution ranges of these 13 people do not generally exceed approximately 50 square kilometers. Lieberman et al. [101] conducted a similar study, exploring the geographic extent of Wikipedia editors' contributions.

5.1.3 Geographic Ranges of Contribution

Hecht and Gergle [69] compared different 'spatial content production models' for generating volunteered geographic information [58] and found that Flickr contributions tended to be much closer to a contributors' 'home location' than was the case with Wikipedia. Hardy et al. [66] considered geographic contribution as a spatial interaction process, using an exponential distance decay model for each language edition. They found that anonymous edits to geotagged Wikipedia articles decay exponentially as the contribution location gets further form a contributor's 'home'. I return to this idea of spatial interaction in the Discussion section.

5.1.4 Temporal Evolution of Contributor Behavior

Whereas the work described above focused on geographic behavior, others have focused on the evolution of non-geographic peer production contributor behavior over time. In one of the seminal studies in this space, Priedhorsky et al. [134] took a temporal approach to understanding how value is created in Wikipedia and by whom. Panciera et al. [128] built on this paper with a study of 'Wikipedian' lifecycles and found that 'Wikipedians' (the term they use to describe those who contribute most of the Wikipedia content) begin contributing at a high level and maintain this trend over time, resulting in distinctive differences in contribution behavior between different classes of users. In other words, "Wikipedians are born, not made" [128]. As noted above, this work strongly informs my study. One of the key takeaways of my work is that this finding, which describes temporal contribution levels in Wikipedia, also applies to spatiotemporal contribution behavior in OpenStreetMap. Panciera's work also inspired the methodologies in this paper: as described below, the spatiotemporal contributor class-specific analyses are a direct analogue to the temporal analyses done in Panciera et al.

Other work uses temporal evolution as a way to characterize the status of a geographic region (versus focusing on contributors and their behavior). One example of such a study is work by Gröchenig et al. [60], who computationally estimated the 'completeness' of twelve urban areas, based on identifying three temporal stages ('start', 'growth', and 'saturation'), and modeling the development of a region through these stages.

More recently, others have begun to explore what roles contributors play in peer production communities, and how that changes over time. Arazy et al. [6] described ‘career paths’ of Wikipedia editors. Rehrl et al. [139] took a similar approach, and considered the different roles that people have in OpenStreetMap. Dittus et al. [35] explored the activation of newcomers and reactivation of previously dormant contributors during disaster events on Humanitarian OpenStreetMap (HOT).

My study here is deeply informed by the work of Panciera et al. [128], and the studies mentioned in the subsection above. Whereas prior work has focused on understanding geographic behavior or the temporal evolution of behavior, my study sits at the intersection. A spatiotemporal lens helps inform my understanding how contributors’ geographic behavior evolves, and how this may impact the geographic variations seen in OpenStreetMap.

5.1.5 Geographic Biases in Peer Production

Geographic coverage biases in peer produced datasets have become a subject of relatively substantial research interest in recent years. For instance, Sen et al. [150] found that most content in some parts of the world (e.g. sub-Saharan Africa) is not produced by people from those parts of the world, but instead by Westerners. Other work shows that these biases manifest along two important human geography variables: the urban/rural divide, and socioeconomic status variation. As one example, Johnson et al. [80] found that the quality of Wikipedia and OpenStreetMap content is much greater in urban areas than in rural areas, a result that informs key analyses below. Haklay [61] found a similar result when considering socioeconomic status as well – the quality of OpenStreetMap data is much better in wealthier regions. Informed by these (and other geographic HCI [70,81,99]) studies, I focus one of my research questions on these two specific dimensions (I discuss this in more detail below).

Prior work in this area has quantified and shown the existence of these geographic biases in peer produced datasets, but little work has been done to understand the mechanisms behind these biases. As mentioned above, my work takes a spatiotemporal approach, at the intersection between studies of temporal contributor behavior and those characterizing the geographic behavior of contributors. For this reason, my work is well-situated to shed light on how the temporal evolution of geographic behavior may (or may not) facilitate the geographic biases that others have found.

5.2 Methods

To study the spatiotemporal evolution of contributors in OpenStreetMap, I needed to (1) develop my OpenStreetMap dataset, (2) define geographic variables of interest (i.e. the 'spatio' in spatiotemporal), and (3) characterize these variables of interest over time (i.e. the 'temporal'). I first provide a brief introduction to how contributions occur in OpenStreetMap and then discuss each of these three steps.

5.2.1 Introduction to Contribution in OpenStreetMap

Where Wikipedia editors help create articles, OpenStreetMap contributors help create a worldwide map (or, more formally, a worldwide spatial database). OSM contributions either add or annotate geographic entities, e.g. bus stops, roads, buildings or even logical collections of buildings like a university. Nodes (points) are the simplest geometric unit in OSM, and they may stand alone (e.g. a bus stop), or they may comprise other types of geometries, namely 'ways' (e.g. roads or buildings) and 'relations' (e.g. a university). Early in the life of OpenStreetMap, contributions depended heavily on "GPS traces" recorded as contributors moved about the world. However, it is now much more common to trace new entities from satellite imagery using a web-based tool [182].

Similar to Wikipedia, OpenStreetMap records a "version history" for each map entity. For instance, when the node for a bus stop is first created, it will be version 1. If the location is adjusted later, the version will be incremented to 2. If the bus stop is then annotated with the available bus lines, the version would be incremented again.

5.2.2 Dataset

My dataset focuses on OpenStreetMap nodes (points) and consists of the full, versioned history of OpenStreetMap, through February 2014. Because ways and relations are made up of nodes, nodes define the underlying geometry of contributions. For this reason, I limit my analysis to OSM nodes (I discuss implications for ways and relations later).

I limit my study site to the continental United States so that I can take advantage of readily-available government census data published by the U.S. Census – a common practice in geographic human-computer interaction studies (e.g. [70,71,80,81,99,164]). Because a key contribution of this work is developing an understanding of urban-rural and socioeconomic biases, it was necessary to ensure that there would be "urbanness" and socioeconomic census variables for my study site. I discuss how this work may extend to other geographic contexts in my Discussion section below.

From the broad OSM dataset, I first extracted all nodes in the continental United States, including every version of every node. I then excluded nodes created in an automated manner (e.g. large imported road datasets and bot-created geometries) using the technique in Johnson et al. [80]. Since I was interested in spatiotemporal trends, I excluded nodes created by people with fewer than five contributions out of sparsity concerns that I discuss in more detail below. Finally, I used a standard reverse geocoding approach to associate each node with the United States county that contains it. In total, I considered more than 28 million (28,021,802) contributions by 23,329 contributors.

Because contribution rates are so skewed in peer produced datasets (i.e. power-law dynamics [128,134]) and informed by Panciera et al. [128], I organize my analysis around three classes of contributors, defined by the number of edits they made:

- 1%ers: The 1% of contributors that produce the most content. In total, "1%ers" contribute 68% of all OpenStreetMap nodes.
- 9%ers: The "middle" 9% of contributors, i.e. those between the 1%ers and the 90%ers. "9%ers" produce 27% of OpenStreetMap content;
- 90%ers: The bottom 90% of contributors. They produce only 11% of OpenStreetMap content.

Note that the percentages above refer to statistics once contributors with fewer than five edits have been removed (these contributors made only 0.07% of edits in total).

5.2.3 Geographic Variables of Interest

I operationalize four geographic variables using my historical dataset of human-generated nodes in the United States. These variables were selected because they had one of two properties: (1) they (or close variants) had been employed in non-temporal characterizations of geographic contributor focus, or (2) they are metrics related to observed geographic biases in peer produced geographic data. My first two variables meet the first property and describe the geometric characteristics of contributors: (1) their geographic ranges [117] and (2) where they focus [101]. My second two variables meet the second property and capture the (1) urbanness [24,70,80,81,135] and (2) socioeconomic status [24,61,137,164] of where people contribute. Below, I detail each of my four variables in turn.

Geometric Variables

***std_dist*:** Standard distance is a common point-pattern analysis metric of geographic dispersion. *std_dist* is analogous to a standard deviation; it represents the geometric spread of a set of points relative to the geometric center of the set. Specifically, a *std_dist* describes the radius of a circle around the mean center point. Like a standard deviation, 68% of the points fall within this circle.

For my analysis, I computed the *std_dist* for each contributor simply by finding the mean center point of their contributions and then computing their dispersion. Prior to making this calculation, I projected all data points into a 2D reference system using the Albers' Equal Area Conic projection.

***plurality_focus*:** While my *std_dist* variable describes the spatial distributions of people's contributions, my *plurality_focus* variable describes the actual locations where people focus. Each contributor's *plurality_focus* county is simply the county in which a plurality of their contributions were made (i.e., the mode). Prior work in geographic HCI [82] often uses this approach to attribute the "home region" of a contributor, but here I interpret "plurality county" more conservatively: I just take it as the region where a contributor has focused their contributions.

Human Geography Variables

My next two variables focus on human geography and describe the kinds of places people contribute. In other words, while my first two variables describe the locations and geographic spread of contributions, the next two describe characteristics of the people who live in the contribution locations. Specifically, I define variables that describe the biases shown in prior literature: ruralness and poverty. Based on the county associated with each node, I label each contribution with: (1) a county urbanness class (from the National Center for Health Statistics' Urban-Rural Classification Scheme [121]), and (2) the percent of the county's population that is in poverty (from the US Census' American Community Survey [20]).

With these labels in place, I compute two variables for each contributor:

***pct_rural*:** This variable describes the percent of a person's contributions that occurred in counties with urbanness classes 5 and 6 (the two nonmetropolitan classes in the classification scheme mentioned above). In Florida, for example, Miami-Dade County (where the city of Miami is located) is a 1 on this urbanness scale, whereas Monroe and Hamilton Counties (near the border with the state of Georgia, approximately halfway between the cities of Jacksonville and Tallahassee) are urbanness classes 5 and 6.

***pct_high_poverty*:** This variable describes the percent of a person's contributions that occurred in 'high-poverty' counties, where at least 20% of the population is in poverty. I base this variable on the definition of 'high-poverty' provided by the United States Census American Community Survey [20]. For example, Webb County in Texas is a high-poverty county. Webb County is home to Laredo, Texas –one of the largest cities on the United States-Mexico border – and has an average per-capita

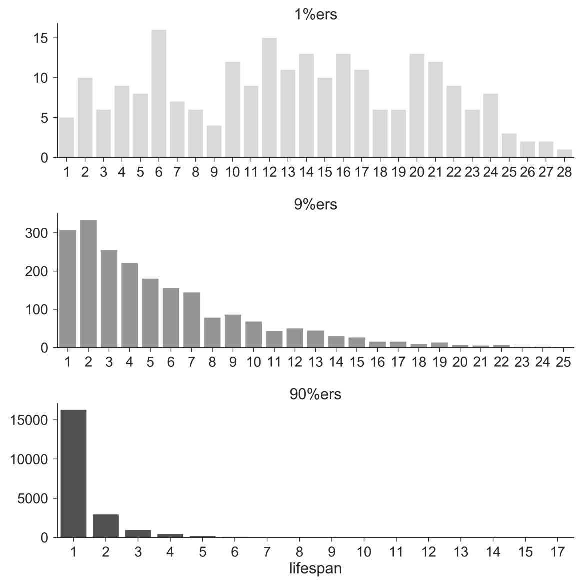

Figure 5.1: A histogram of contributors who participate for each number of quarters.

income of approximately $10,000 (approximately $2,000 below the US poverty line in 2015).

5.2.4 Temporal Units of Analysis

Each of my four variables are a descriptive summary of the geography of contributors' focus, but they are not temporal. To understand how these geographic summaries change, I temporally group each person's contributions into quarters (Jan. 1 - Mar. 31st, April 1 - June 30, July 1 - Sept. 30, and Oct. 1 - Dec. 31). I selected three-month periods to ensure that (a) there would be sufficient data in each period, and (b) the temporal periods were granular enough to analyze the evolution of contributors' behaviors over time. For each contributor-quarter, I computed my four geographic variables. As I noted above, I excluded contributors with fewer than five contributions to avoid drawing conclusions from excessively small samples.

Figure 5.1 shows a histogram of the number of quarters that people participate in OpenStreetMap. Most people (71%) participate in only one quarter. These contributors are (a) predominantly 90%ers, and (b) account for only about 4% of the total edits in my dataset. 1%ers participate for a median of thirteen quarters, 9%ers for a median of five quarters, and 90%ers for a median of two quarters. I discuss the implications of these medians below.

5.3 Results

I use my two main research questions to frame the presentation of my results. As I previewed, I generally find that most people are quite consistent throughout their contribution lifespans – contributors' geographic behavior tends to be 'born, not made'. Since this is exploratory work, I approach both research questions by identifying and characterizing the general trends in the data. I also highlight important deviations from those trends. I now discuss the results for each of my research questions in turn.

RQ1: How does contributors' geographic focus change over time?

The spatiotemporal trends in my *std_dist* and *plurality_focus* variables tell a relatively clear story: most contributors and contribution groups tend to have consistent geographic ranges and focus areas. In other words, most (though not all) contributors' geographic focus behavior is 'born, not made'. I now unpack these findings in more detail.

***std_dist*:** Figure 5.2, which visualizes contributors' quarterly geographic ranges over time as defined by *std_dist*, shows a relatively clear trend: contributor groups have meaningfully distinct standard distances, and these distinctions are mostly consistent over time. Along the y-axis in Figure 5.2 – following the method used by Panciera et al. [128] – I plot the mean and 95% confidence interval in each quarter. I

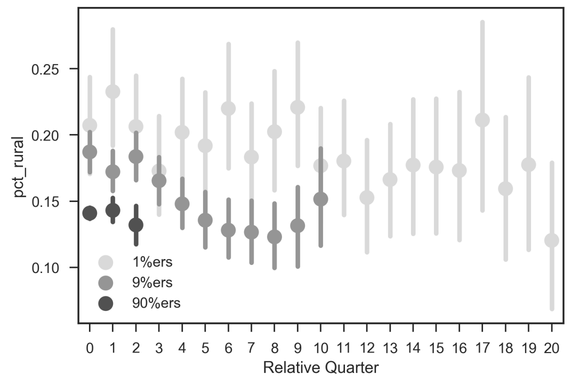

Figure 5.2: Mean std_dist over time, by user class. Error bars show 95% confidence intervals.

find that 1%ers' and 9%ers' average standard distances do not meaningfully vary over time. At first glance, Figure 5.2 may suggest that 1%ers and 9%ers increase their average *std_dist* over their lifespan. However, a closer inspection of the quarterly confidence intervals shows that these changes in means are not meaningfully different from one quarter to the next; the confidence intervals are highly overlapping. By contrast, I do see a meaningful uptick in 90%ers standard distances as their lifespan increases. Note that this figure does not show quarters that exceed the 90th percentile of participation length, because the number of contributors becomes very small.

Although the 95% confidence interval ranges in Figure 5.2 look small and stable over time, I wanted to ensure that individual contributors do not substantially vary their *std_dist* values over time within their group ranges. The potential for this outcome is most salient for 1%ers for two primary reasons: (1) 1%ers contribute most of the content in OpenStreetMap so their geographic behavior has a substantial impact, and (2) in Figure 5.2, 1%ers show the largest confidence interval ranges, conceivably allowing for more individual variation.

To address this question, I did a targeted analysis of 1%ers to evaluate their

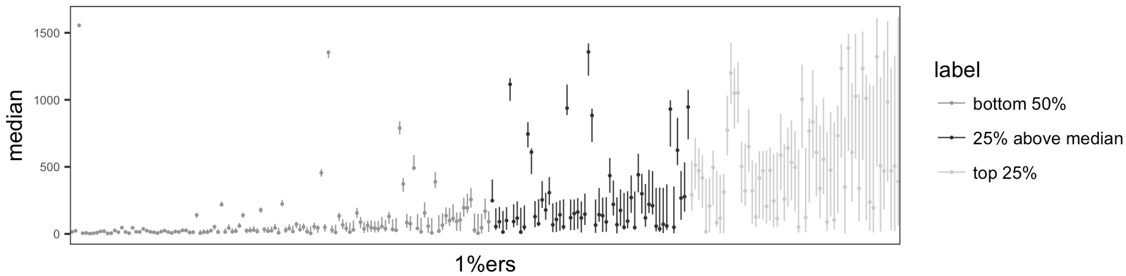

Figure 5.3: Distributions of each individual 1%ers' std_dist. Dots indicate medians, and lines indicate IQR (interquartile range).

consistency over time, the results of which are visible in Figure 5.3. The figure plots each individual 1%ers' *std_dist* distribution, showing the median and interquartile range (IQR) of their *std_dist* in each quarter. The IQR is the distance between the 25th and 75th percentiles of a distribution, or the width of the middle 50% of *std_dist* values here. Individuals are ranked by IQR in increasing order along the x-axis. Critically, shorter lines (smaller IQRs) indicate a higher degree of 'born, not made' behavior with regard to standard distances

The large number of small grey bars on the left side of Figure 5.3 confirms that most 1%ers exhibit 'born, not made' *std_dist* patterns, i.e. their geographic ranges are largely consistent in every quarter. Figure 5.3 also reveals that the higher variance we see in Figure 5.2 is primarily the result of a minority of 1%ers who do not display 'born, not made' *std_dist* patterns. This non-trivial minority exhibits different geographic range patterns across quarters.

It is important to note that the IQR values in Figure 5.3 do not appear to be strongly driven by the number of quarters in which a contributor participates. For instance, a 1%er's *std_dist* IQR and the number of quarters they participate are only weakly correlated (Pearson's r=0.2).

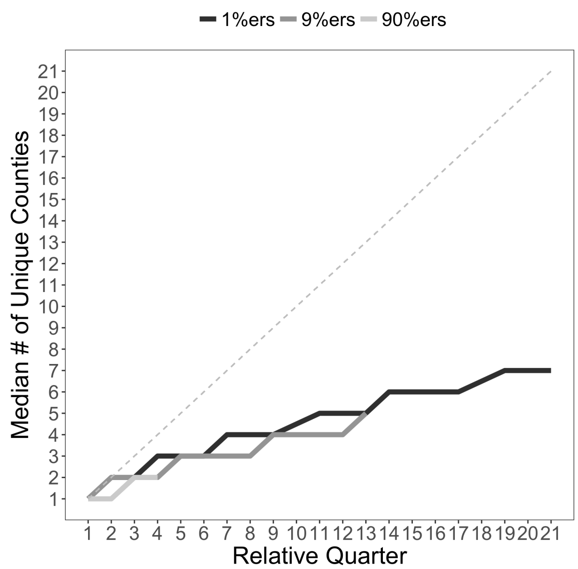

Figure 5.4: Plots the growth of unique plurality_focus counties over time. Each color is a different user class, and the dashed line represents a new plurality_focus county every quarter.

***plurality_focus*:** While *std_dist* characterizes the geographic dispersion of contributor edits, it does not capture where contributors focus. For this, I use *plurality_focus*.

Figure 5.4 plots the median number of unique *plurality_focus* counties over time. Each solid line represents a user class, truncated at the 90th percentile of participation length. The dashed line shows what would occur if the median contributor had a new *plurality_focus* county every quarter.

Figure 5.4 makes one trend clear: while the median contributor does increase the number of counties in which they focus over time, this increase is gradual and substantially less than would be the case if the median contributor focused in new areas each quarter. Intuitively, the median contributor tends to be fairly consistent in where they focus, returning to the same few counties over time. For instance, the

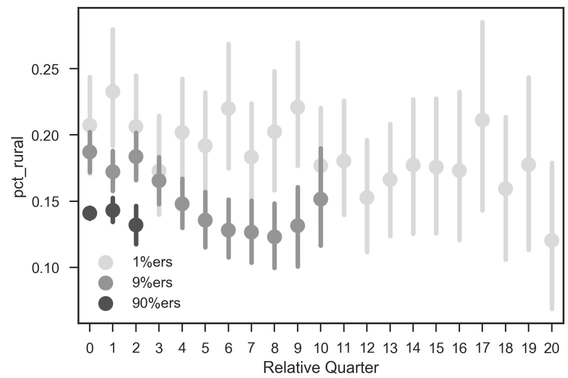

Figure 5.5: Mean pct_rural over time, by user class. Error bars show 95% confidence intervals.

median 90%er participates for two quarters, but has a single *plurality_focus* county on average. The median 9%er participates for five quarters, and this contributor has only three unique *plurality_focus* counties on average. Strikingly, the median 1%er participates for 13 quarters (more than 3 years), and on average has five unique *plurality_focus* counties.

RQ2: Can the spatiotemporal evolution of contributors' focus facilitate systemic coverage biases?

I now turn to my second research question, which. uses the *pct_rural* and *pct_high_poverty* variables to investigate patterns in geographic behavior concerning kinds of places (e.g. poor vs. rich) rather than specific places (i.e. individual counties). I highlight the general trends in these variables as well as impotant deviations from the trends.

5.3.1 Overall Trends

Figure 5.5 (*pct_rural*) shows the mean rate of contributions in counties classified as 5 or 6 on the National Center for Health Statistics urbanness scale. Figure 5.6

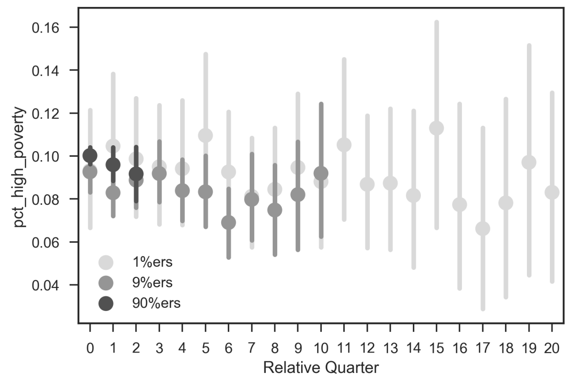

Figure 5.6: Mean pct_high_poverty over time, by user class. Error bars show 95% confidence intervals.

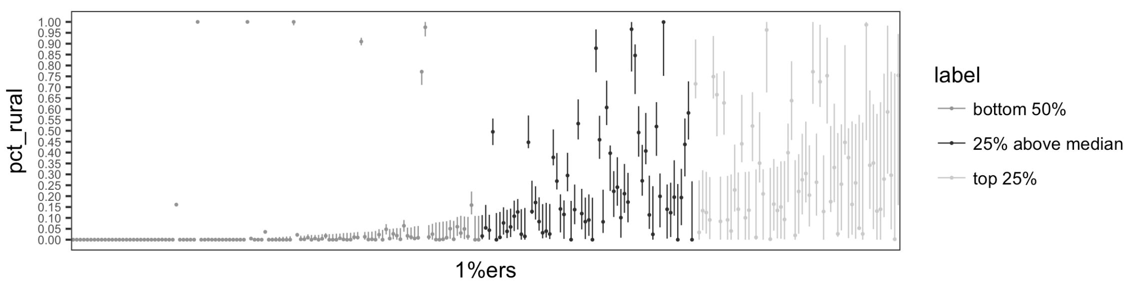

Figure 5.7: Distributions of each individual 1%ers' pct_rural values. The dot indicates the median, and the line indicates their interquartile range.

(*pct_high_poverty*) shows the mean rate of contributions in counties designated as 'high-poverty', according to the US Census. As before, these plots show the 90th percentile number of participation quarters. In both cases, the means of these distributions remain consistent across time for all three user classes, suggesting that most people consistently contribute a relatively small proportion of their edits in rural and poor counties. Even 1%ers, who have the largest standard distance (and thus contribute across larger distances) make less than one fifth of their contributions in rural areas on average, and even fewer in high-poverty areas (and do so consistently across their lifespans).

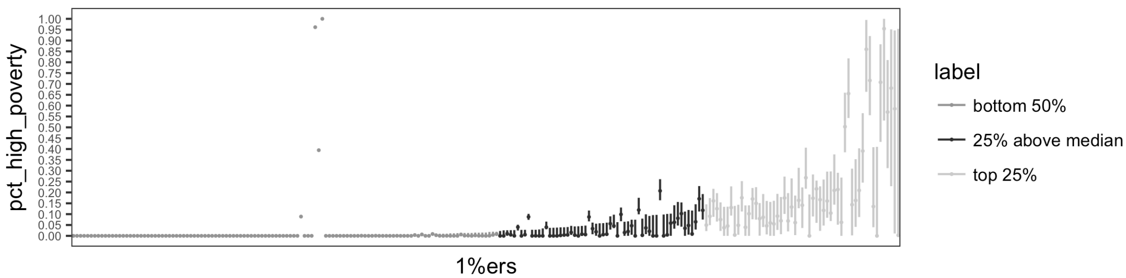

Figure 5.8: Distributions of each individual 1%ers' pct_high_poverty values. The dot indicates the median, and the line indicates their interquartile range.

As before, while the community-level trends are consistent over time, I also wanted to check whether these trends hold at the individual level. I again focused on 1%ers, who have the widest confidence intervals in Figures 5.5 and 5.6 and who contribute the most edits. Figures 5.7 and 5.8 confirm that the majority of 1%ers tend to be quite individually consistent, having persistently low individual median *pct_rural* and *pct_high_poverty* values. The median *pct_rural* IQR is 0.11 and the median *pct_high_poverty* IQR is 0.02, both of which are quite small (on a scale from 0 to 1)[1]. Moreover, the small variation is centered on mostly urban and mostly-non-poor regions, as can be seen by the tendency of the grey lines on the left of Figures 5.7 and 5.8 to be at the bottom of the y-axis.

The results in Figures 5.7 and 5.8 indicate that there is a strong 'born, not made' signal in my *pct_high_poverty* and *pct_rural* variables. In other words, geographic

[1] As was the case above with *std_dist*, we see very weak correlation between the number of quarters a 1%er spends in OpenStreetMap and their IQR (Pearson's $r = 0.09$ and 0.06 for *pct_rural* and *pct_high_poverty*, respectively). # Looking Forward Taking a holistic view on the previous three chapters points to a number of different directions for future work, all of which are motivated by an over-arching theme: mitigating geographic bias. My research here has developed critical insights into crowd members' geographic behavior. I see these studies as laying the groundwork that informs approaches to ameliorating the geographic biases that pervade [24,61,80,81,99,135,137,164] geographic platforms. Some of these future work directions are more open-ended, and present opportunities to develop a new research agenda. In other cases, there are more specific suggestions for future work that follow directly from one of the three chapters. I discuss each of these – open-ended research directions, and more direct follow-up studies – in turn.

biases may be “born, not made’. If contributors start by contributing the large mjaority of their content in urban areas, this trend typically will persist for their entire time in OpenStreetMap. My *pct_high_poverty* variable shows the same result – most contributors (a) do not contribute much content in high-poverty areas, and (b) maintain this trend over time.

5.3.2 Contextualizing *pct_rural* and *pct_high_poverty* values

To put my *pct_rural* and *pct_high_poverty* results into context, I now consider three dimensions against which to compare these results. Specifically I ask if the *pct_rural* and *pct_high_poverty* findings in Figures 5.5, figs. 5.6, 5.7, 5.8 are proportional to what would be expected given (1) the population of these counties, (2) the number of rural or high-poverty counties themselves, or (3) the number of contributors focusing in rural or high-poverty areas.

With regard to county population, according to the United States Census [121], nearly 15% of the US population lives in rural areas, and approximately 14% live in high-poverty areas. Comparing these numbers against Figures 5.5 and 5.6 suggests that the average rate of rural contribution is actually proportional to the population rate in these counties. However, this is not true for my *pct_high_poverty* variable. The average rate of contribution in high-poverty areas is approximately 10%, indicating that high-poverty counties are underrepresented across the board.

Another option to consider is whether these *pct_rural* or *pct_high_poverty* rates are proportional to the number of counties that are rural or high-poverty counties, i.e. maybe there are just fewer of these counties. 63% of counties are rural (have urbanness classes 5 or 6), and 24% of counties are high-poverty (at least 20% of their population is in poverty). Comparing these numbers to the median *pct_rural* or *pct_high_poverty* rates shown in Figures 5.5 and 5.6, the conclusion is clear: in terms of the number of counties, OSM contributors in all user classes are undercovering rural and high-poverty counties. While there may be fewer people in many of these counties, these counties still have road networks, natural features like lakes and rivers, and many other entities that are not directly correlated with population [79] and that typically are mapped in OpenStreetMap.

A third consideration is whether the number of contributors focusing in rural or high-poverty counties is proportoinal to the population of these regions. One important reason to consider this dimension is the effect it may have on content quality. Prior work has shown that people who focus near where they live produce more diverse [175], richer [80], and higher quality [38] content. Unfortunately, Figures 5.7 and 5.8 suggest concerning trends here too. As noted above, 15% of the US population live in rural areas, and 14% live in high-poverty areas. However, Figures 7 and suggest substantially fewer 1%ers focus in rural or high-poverty areas – very few have medians near the top of the y-axis.

Thus far, my results suggest that most contributors – across all user classes – are consistent across time, and contribute in consistently urban and wealthier areas. Further still, rural and high-poverty areas are disproportionately undercovered in comparison to (a) the number of rural and high-poverty counties, and (b) the number of contributors who focus in these areas. Taken together, my results suggest that (a) where contributors focus, (b) the kinds of places they focus in, and (c) the consistency with which this occurs all contribute to the geographic coverage biases shown in prior literature.

5.3.3 Additional Mechanisms of Bias

I noted above that 1%ers participate for the longest period of time, which creates a secondary mechanism facilitating bias – longevity bias. Specifically, people who participate longer contribute longer and because of 'born, not made' trends, contribute in the same places (and kinds of places) longer.

While this trend is intuitive when comparing 1%ers and 90%ers (after all, 1%ers produce most of the content), I wanted to understand how a longevity bias might facilitate socioeconomic and urbanness focus biases. Therefore, I split contributors into two groups, those who tend to be rural-focused (have a median *pct_rural* of at least 50%), and those who tend to be urban-focused (have a median *pct_rural* below 50%). I computed how long each contributor participated, and compared the urban-focused and rural-focused groups. Examining the means of these groups (urban-focused: 1.9 quarters, rural-focused: 1.65 quarters) suggests that urban-focused contributors participate longer, on average. Due to a skewed distribution, I conducted a Wilcoxon Rank-Sum Test which found significant differences between the two groups (z=2.67, $p < 0.01$). I ran the same analysis for my *pct_high_poverty* contributors. Again, the means (non-high-poverty focused: 1.9 quarters, high-poverty focused: 1.55 quarters) suggest that high-poverty focused contributors participate longer, on average. A Wilcoxon Rank-Sum Test also found significant differences between the two groups (z=4.81, $p < 0.001$).

While these findings are not causal – and future work should examine predictors of retention in OSM – they do potentially have implications for the evolution of bias in OSM. Specifically, these results suggest that the bias in where people focus is perpetuated by who remains a contributor. Most people, across all user classes, consistently contribute small amounts of content in rural and high-poverty areas over the course of their time in OSM. People who do focus in rural and high-poverty areas stop contributing earlier than people who focus in more urban, or wealthier areas. This finding potentially has important implications for improving coverage in rural and high-poverty areas, something to which I return in the Discussion section.

5.3.4 Deviations from Trend

Faced with results that suggest that most people consistently contribute in urban and non-high-poverty areas, I sought to better understand contributors who do primarily focus in rural and/or high-poverty areas and the contributions that they make. What I found aligns strongly with what is shown in Figures 5.7 and 5.8. The majority of rural and high-poverty content is not contributed by consistently rural or consistently high-poverty contributors. Figures 5.7 and 5.8 indicate that relatively few 1%ers have high median *pct_rural* and *pct_high_poverty* values, and that many of those who do also tend to have wider IQRs, indicating that they are less consistent over time in terms of the types of places they edit than the median 1%er.

To understand these rural and high-poverty focused contributors in more detail, I use the same metric as above: if a contributors' median *pct_rural* and median *pct_high_poverty* are at least 50%, I consider them rural-focused and high-poverty-focused, respectively.

Beginning with rural-focused contributors, I found that 3,126 people tend to contribute in rural areas, and as a group contribute less than 40% of content in rural areas. There are 27 rural-focused 1%ers (those nearer the top of the Y axis in Figure 5.7), 315 rural-focused 9%ers, and the rest (2,748) are 90%ers. They account for 25%, 11%, and 2% of rural content, respectively (totaling 38% of rural content). Because

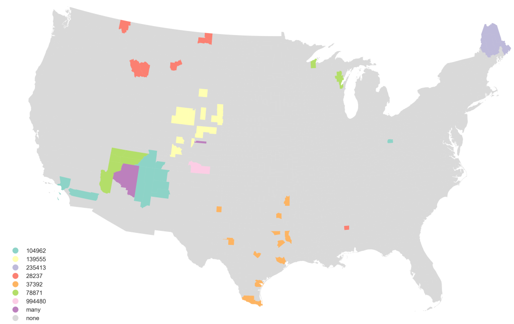

Figure 5.9: All counties for rural focused 1%ers.

1%ers contribute most of the content in OpenStreetMap, I have mapped the plurality focus counties for the seven most prolific rural-focused 1%ers in Figure 5.9. I selected only the seven most prolific to aid in map legibility [179].

There are two primary trends in Figure 5.9: (1) people who contribute in national parks (and national forests), and (2) people who contribute regionally. With respect to the national parks, (a) prior studies have shown that vacation destinations are common locations for VGI contribution [132], and (b) very few people live in counties with national parks. What this suggests is that some of the participants who I termed rural-focused may instead be 'national park-focused', with national parks serving huge numbers of urban visitors. The second pattern in Figure 5.9 involves regional contributors. To take one example, consider the person contributing in northern Maine (in the upper northeast corner of Figure 5.9). This area is very sparsely populated, and yet a single, consistently rural 1%er contributes most their content, over multiple quarters, in those counties. Both groups have implications for recruitment in peer production communities, which I discuss further below.

Turning to high-poverty contributions (Figure 5.10), the trend I observed for rural areas is even more severe. I found that 2,014 people consistently contribute in high-

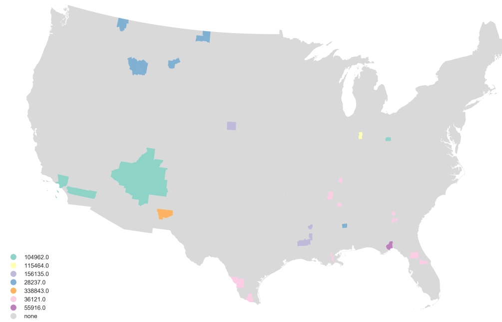

Figure 5.10: All counties for high poverty focused 1%ers.

poverty areas, and as a group contribute slightly more than one-fourth of the content in high-poverty areas. There are 11 high-poverty-focused 1%ers (those nearer the top of the Y axis in Figure 5.8), 126 high-poverty-focused 9%ers, and the rest (1,877) are 90%ers. They contribute 16%, 8%, and 2% of high-poverty content, respectively (totaling 26% of high-poverty content). I have mapped the plurality focus counties for the seven most prolific high-poverty focused 1%ers in Figure 5.10. I again selected only the seven most prolific to aid in map legibility.

The contributors in Figure 5.10 show similar trends to those in Figure 5.9: many of the counties shown contain national parks and forests and a few are contributors who contribute regionally. One example of the first trend is the large teal section in the southwestern section of the map (the area surrounding the Grand Canyon). The counties that contain the Grand Canyon also contain the Navajo Indian Reservation, one of the five most impoverished reservations in the United States [180]. This lends further credence to the idea that some contributors focus in natural parks, and it is likely that these contributors are not contributing in the very impoverished parts of this region. However, there are some contributors who are consistently focused in high-poverty areas. For example, consider Sierra County, New Mexico (reddish), also in the southwestern corner of the map. The person primarily contributing here is focused on high-poverty counties. Residents of Sierra County tend to be quite poor, with a median household income of $25,583, and a per-capita income of $16,667. Another example of a high-poverty area is the more northern county in Texas (pink, central southern section of the map) – Webb County, Texas. Webb County is home to Laredo, the third largest city on the Mexico-United States border. The median household income in Webb County is $28,100, and the per-capita income is $10,179. As before, both examples suggest implications for recruitment that I discuss below.

5.4 Discussion

In this section, I step up a level and discuss the implications of my findings more broadly. This section follows the same structure as the results section. Specifically, we first discuss what my findings mean for my understanding of contributor behavior in peer production systems. Second, I discuss what my findings suggest for the mitigation of urban and socioeconomic coverage biases in peer production systems.

5.4.1 Implications for Peer Production

Standard Distance and Spatial Interaction Behavior

Closely related to my *std_dist* variable is a concept from geography called spatial interaction [88,89,152], which is used to describe 'flow' between regions, e.g., of physical goods [88,89] or people [152]. This process often is modeled with gravity models and characterizes, e.g., the rate at which travel between regions changes as a function of distance and attributes of the regions. The 'cost of distance' aspect of these models is particularly relevant to my findings here.

I find that different classes of contributors (e.g. 1%ers vs. 90%ers) have consistently distinct sizes of geographic range, which presents an important opportunity for future work. Intuitively, these findings suggest that different contributor classes interact consistently differently across distance. Prior geographic HCI [71] work using gravity models has not accounted for contributor class, but doing so may provide for better understanding of the mechanisms behind spatial content production. This may also help support predictions about which areas would receive contributions if, for instance, a concerted recruiting effort were made in rural areas (as is discussed in more detail below).

5.4.2 Mitigating Coverage Biases

My results suggest that 'born, not made' dynamics may naturally facilitate the creation of geographic coverage biases, which are in part enabled by who remains a contributor over time. I next reflect on how my results suggests mechanisms for reducing these biases.

5.4.3 Existing Consistently Rural or High-poverty Contributors

The first intuitive approach to mitigating biases is to examine those participants who do consistently focus in rural and high-poverty areas. After all, these participants are contributing a non-trivial amount of content in rural and high-poverty areas already. As noted above, my results suggest that there are two trends in where these rural-focused or high-poverty-focused 1%ers contribute: national parks and regional areas.

National Parks: Leveraging existing contributors who focus in the counties that contain national parks and forests to address poverty or urban/rural bias is likely to be difficult. Prior work suggests that vacation destinations are common locations for geographic contributions [132], and that people tend to be more aware of the geography in places with which they are familiar [59]. Thus, it is likely that the contributors who focus in counties that contain national parks are not producing content in the rural or high-poverty sections of those counties (although investigating this hypothesis in detail is a good targeted direction of future work).

Regional Focus: The other group of rural- or high-poverty-focused 1%ers, however, may be more promising. These contributors already are focusing their effort in rural or high-poverty areas. I see two implications for design here. The first is simple: find ways to keep these contributors in the community! My results suggest that the longevity of these contributors in OpenStreetMap is less than their peers, and targeting this issue would be one immediate and effective partial solution to coverage biases. Second, my results suggest that targeted recruitment of regionally-focused 1%ers in low-coverage areas could be effective.

5.5 Conclusion

In this paper, I performed the first examination of the spatiotemporal behavior of contributors to geographic peer production communities. I observed that contributors' spatiotemporal behavior is generally consistent throughout their contribution lifespans, both with respect to the geometric structure of contributions and with respect to the types of places to which contributions are made (e.g. urban places vs. rural places). In other words, I saw evidence that there is a strong (but not omnipresent) 'born, not made' tendency in spatiotemporal peer production behavior. More generally, this work sheds light on some of the mechanisms by which the coverage (and coverage biases) of peer produced geographic datasets may occur.

5.6 Open-Ended Research Directions

The over-arching theme of both directions discussed below is mitigating geographic biases in crowd platforms. All of the studies presented in this thesis provide important insights into how crowd members behave geographically, with an eye towards understanding how, and perhaps why, these geographic biases come about.

Therefore, an intuitive next step is developing system-level strategies in order to mitigate disparities in sharing economy availability, or content coverage in VGI crowd platforms. Each of my studies in this thesis speak to a different aspect of understanding of the mechanisms that underpin these geographic biases. That said, developing an understanding of the clear, practicable mechanisms undergirding these biases is a necessary first step towards mitigating these disparities.

5.6.1 Crowd Member Context and Mental Maps

One of the primary reoccurring themes in the discussions near the end of each chapter was the role of mental maps and geographic context in crowd member behavior. In Chapter 3, participants quite clearly discussed how their own mental maps played a role in where they were available to work. On the other hand, in Chapters 4, 5, the role of geographic context in crowd member behavior is more implicit. For instance, in Chapter 4, I noted that there are no technological barriers to VGI crowd members producing content "from their armchairs", but given the existence of geographic biases in coverage, this is clearly not occurring equally. I found that different content types have varying degrees of localness, and hypothesized that the factors driving crowd members' decisions about where to contribute may be self-focus bias [68], or a pattern of contribution preferential attachment. In Chapter 5 I saw a large degree of regularity in the places, and kinds of places, individual crowd members focus their contributions. Taken together, these three studies show that at an individual-level, crowd members contribution behavior or work patterns are likely influenced by different facets of their geographic context (e.g. knowledge, feelings of safety, self-focus bias, topic interest, etc.). At the system-level, however, this seems to manifest as disparities in content quality and coverage, or service availability.

If this hypothesis bears out, one clear research agenda focuses on this individual-level geographic context – can we intervene on the mental maps or context of crowd workers to help facilitate better sharing economy service or VGI coverage? My work points to two different instantiations of geographic context: physical travel, and 'arm-chair VGI production'. Each of these threads open up a new direction of research.

Context for Physical Travel. The first direction of research to intervene on crowd members' geographic context is focused on the sharing economy, where crowd members *do* physically travel to perform tasks, or pick up drivers. Among some recent studies [33,96], my work above points to geographic factors that serve as barriers to sharing economy availability. While I note above that localized recruitment in low-SES, suburban, and non-white areas may go a long way towards mitigating geographic biases, I also discuss some approaches to adjusting crowd members' mental maps, predominantly focused on the perceptions of safety my participants discussed. Beyond this, however, little is known within the HCI literature about what factors drive the geographic decisions of sharing economy crowd members. In other domains, there is a wide body of literature about impact that cognitive biases and heuristics have on human judgement. For instance, Danziger et al. [30] found that judges rulings are much more lenient immediately after a lunch break, rather than immediately before – hunger is an important contextual factor in how judges make decisions. What are analogous concerns when designing crowd platforms that require physical travel, like the sharing economy? Are perception of safety and distance the only salient concerns? What other contextual factors affect the *geographic* decisions sharing economy crowd members make? Do these concerns vary by time of day? Building on my work here, I see a two clear steps of future research. First, developing a more complete understanding of *why* sharing economy crowd members make the decisions they do. This would serve as a stepping stone to build systems and interfaces that intervene on individual-level behaviors, and in turn mitigate the system-level geographic biases discussed above. This agenda has implications for the design of tools for sharing economy crowd members, and may generalize into other forms of 'physical travel crowd platforms' like the observational citizen science platform eBird.

Context for 'Armchair VGI Production'. The second direction of 'geographic context' research only applies in crowd platforms where the work can be done 'remotely', like in peer production crowd platforms. The difference between this context and the physical travel context discussed above hinges on a point I mentioned earlier: there is no technical reason blocking crowd members in OpenStreetMap or Wikipedia from producing content about any place in the world. That is, unlike the sharing economy crowd platforms where being physically present in the place is necessary, OpenStreetMap and Wikipedia have software that facilitates 'armchair' content production. This distinction between being physically present or not is fundamental – physical presence is an implicit form of context in the sharing economy crowd platforms. Indeed, my participants in Chapter 3 discuss perception of safety as one example of these differences. After all, because an 'armchair' contributor in OpenStreetMap is not physically in the place they contribute to, any concerns about physical safety do not obviously apply.

My work in Chapters 4 and 5 shows that context has an impact on geographic behavior, even in 'armchair VGI' crowd platforms. Specifically, different contextual factors like the type of content and familiarity with a (type of) region seem to influence the way crowd members contribute in VGI platforms. For instance, in Chapter 4, I saw that the different content types may change the ways in which 'attractiveness' and distance affect contribution behavior. In VGI production, geographic behavior and the type of content seem to be related. In Chapter 5, I found that over time, OpenStreetMap contributors tend to focus in many of the same places, or other places that are similar socioeconomically or in terms of population density. Geographic behaviors in VGI crowd platforms follow 'born, not made' [128] patterns.

Taken together, the findings of these two studies suggest that set of contextual factors impact VGI production behavior differ meaningfully from those that impact physical travel geographic behavior. Individual geographic behavior tends to be quite consistent over time, but geographic behavior varies meaningfully as the content type changes. Can people's content type preferences be identified? Are people as consistent in the type of content they contribute as they are in the kinds of places they contribute to? If so, informing VGI crowd members of topical work to be done in undercovered regions may be quite fruitful. One fundamental unknown in this space is whether or not the role of distance and attraction are static, for a given type of content. If so, some types of content *may never reach coverage parity* with the current set of contributors. On the other hand, if not, a clear direction of future work is to understand how tools can adjust the effect of attraction or distance. Put another way, if a given content type's 'localness' is not inherent to the type of content, there may be opportunities to benefit from changing how local the production of that content must be.

5.6.2 Localized Recruitment

A second broad research agenda from my work focuses on adjusting the number of local crowd members, whereas the agenda discussed above focused on using existing crowd members to mitigate biases. Put another way, rather than get 'outsiders' to provide sharing economy service or produce VGI content, what would it mean to increase the number of 'locals'? In Chapter 3, where crowd members lived directly impacted where the sharing economy was available. In Chapter 4, different content types are likely to receive different rates of contribution, depending on rate of distance decay for that type of content. In Chapter 5, individual crowd members tend to be fairly regular in the places they focus their contributions, indicating a degree of 'local focus' in how the participate in the crowd.

The value of local participation is only beginning to be understood. Some preliminary studies have found that in VGI crowd platforms, content produced by locals tends to be richer[82], more diverse[176], and higher quality [38]. However, more work is warranted to understand the impact and value of local contributions. Is local content actually better in all cases? If so, should crowd platforms make 'armchair' participation more difficult to do?

A critical component of this research direction is the question of recruitment – is it even possible to recruit new crowd members from everywhere these crowd platforms seek to serve? While I discuss barriers to people from low-SES areas participating in sharing economy crowd platforms, I am unaware of similar work in the VGI crowd platform space.

5.7 Specific Follow-Up Studies

One immediate study that can directly inform the broader 'crowd worker context' agenda described above is a qualitative study focused the reasons that underpin crowd members' geographic behavior. For instance, why to OpenStreetMap contributors decide to contribute in the places they do? Why do UberX drivers focus in some geographic regions? Developing a deeper understanding of crowd members' reasons for their decisions would be quite fruitful.

Further, in all three chapters above, I focus my analyses in a US context, and in some cases in one particular metropolitan area. There are clear follow-up studies to be performed exploring the same questions I studied here, but in different geographic areas (or globally). The geographic principles discussed in Chapter 2 are likely to hold across different regions and cultures, though their particular instantiations may differ. The geographic HCI [71] community would benefit substantially from replications across many different geographic contexts.

With regard to my sharing economy work in Chapter 3, I believe that longitudinal analyses (akin to the one I ran in Chapter 5) that expand the static snapshots in this paper are important directions of future work. The sharing economy is an incredibly fast-moving space: adoption rates are growing both on the consumer and the crowd member side, policy is shifting, and (as noted above) geosociotechnical designs are constantly changing. It is unlikely that geography's role in sharing economy effectiveness will decline. However, the character of geography's role may change as the sharing economy develops.

References

1. Pragya Agarwal. 2001. Walter Christaller, Hierarchical Patterns of Urbanization. *eScholarship*.

2. Florian Alt, Alireza Sahami Shirazi, Albrecht Schmidt, Urs Kramer, and Zahid Nawaz. 2010. Location-based crowdsourcing. In *The 6th Nordic Conference*, 13. https://doi.org/10.1145/1868914.1868921

3. Mike Annany, Karrie Karahalios, Christian Sandvig, and Christo Wilson. 2015. Auditing Algorithms From the Outside: *Auditing Algorithms From the Outside:*

4. L. Anselin. 1988. *Spatial Econometrics: Methods and Models*. Springer Science & Business Media.

5. Judd Antin, Ed H. Chi, James Howison, Sharoda Paul, Aaron Shaw, and Jude Yew. 2011. Apples to Oranges?: Comparing Across Studies of Open Collaboration/Peer Production. In *Proceedings of the 7th International Symposium on Wikis and Open Collaboration* (WikiSym '11), 227–228. https://doi.org/10.1145/2038558.2038610

6. Ofer Arazy, Hila Liifshitz-Assaf, Oded Nov, Johannes Daxenberger, Martina Balestra, and Coye Cheshire. 2017. On the "How" and "Why" of Emergent Role Behaviors in Wikipedia. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (CSCW '17), 2039–2051. https://doi.org/10.1145/2998181.2998317

7. Joseph De Avila. 2016. Airbnb Will Collect Hotel Tax for Hosts in Connecticut. *Wall Street Journal*.

8. Victoria M.E. Bellotti, Sara Cambridge, Karen Hoy, Patrick C. Shih, Lisa Renery Handalian, Kyungsik Han, and John M. Carroll. 2014. Towards community-centered support for peer-to-peer service exchange: Rethinking the timebanking metaphor. 2975–2984. https://doi.org/10.1145/2556288.2557061

9. Yochai Benkler and others. 2016. Peer production and cooperation. *Handbook on the Economics of the Internet* 91.

10. Alan Benson, Aaron J. Sojourner, and Akhmed Umyarov. 2015. Can Reputation Discipline the Gig Economy? Experimental Evidence from an Online Labor Market.

11. Michael S. Bernstein, Greg Little, Robert C. Miller, Björn Hartmann, Mark S. Ackerman, David R. Karger, David Crowell, and Katrina Panovich. 2010. Soylent: A Word Processor with a Crowd Inside. In *Proceedings of the 23Nd Annual ACM Symposium on User Interface Software and Technology* (UIST '10), 313–322. https://doi.org/10.1145/1866029.1866078

12. Michael Bernstein. 15AD. The Future of Work: Working for the Machine. *Pacific Standard*.

13. William T. Bielby. 2000. Minimizing Workplace Gender and Racial Bias. *Contemporary Sociology* 29, 1: 120–129. https://doi.org/10.2307/2654937

14. Mark Bjelland, Daniel Montello, Jerome Fellmann, Arthur Getis, and Judith Getis. 2013. *Human Geography: Landscapes of Human Activities*. McGraw-Hill Education, New York.

15. Charles M. Blow. 2015. How Expensive It Is to Be Poor. *The New York Times*.

16. Jonathan M. Bossenbroek, Clifford E. Kraft, and Jeffrey C. Nekola. 2001. Prediction of Long-Distance Dispersal Using Gravity Models: Zebra Mussel Invasion of Inland Lakes. *Ecological Applications* 11, 6: 1778–1788. https://doi.org/10.1890/1051-0761(2001)011[1778:POLDDU]2.0.CO;2

17. Vince Breneman and Michele Ver Ploeg. 2016. USDA Economic Research Service - Food Access Research Atlas.

18. Stanley D. Brunn, Jack Francis Williams, and Donald J. Zeigler. 2003. *Cities of the World: World Regional Urban Development*. Rowman & Littlefield.

19. US Census Bureau. 1999. Explanation of Race and Hispanic Origin Categories.

20. US Census Bureau. American Community Survey (ACS).

21. Ludwig Burger. Fri Oct 30 21:31:46 UTC 2015. Uber Germany retreats to Berlin, Munich. *Reuters*.

22. John Byrne. 2016. Airbnb rules easily pass Chicago City Council despite vocal opposition. *chicagotribune.com*.

23. R. A. Cochrane. 1975. A Possible Economic Basis for the Gravity Model. *Journal of Transport Economics and Policy* 9, 1: 34–49.

24. Ashley Colley, Jacob Thebault-Spieker, Allen Yilun Lin, Donald Degraen, Benjamin Fischman, Jonna Häkkilä, Kate Kuehl, Valentina Nisi, Nuno Jardim Nunes, Nina Wenig, and others. 2017. The Geography of Pokémon GO: Beneficial and Problematic Effects on Places and Movement. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.

25. Ronan Collobert and Jason Weston. 2008. A Unified Architecture for Natural Language Processing: Deep Neural Networks with Multitask Learning. In *Proceedings of the 25th International Conference on Machine Learning* (ICML '08), 160–167.

https://doi.org/10.1145/1390156.1390177

26. Ryan Compton, Craig Lee, Jiejun Xu, Luis Artieda-Moncada, Tsai-Ching Lu, Lalindra De Silva, and Michael Macy. 2014. Using publicly visible social media to build detailed forecasts of civil unrest. *Security Informatics* 3, 1: 4. https://doi.org/10.1186/s13388-014-0004-6

27. United States Congress. 1964. Civil Rights Act of 1964.

28. Stefany Coxe, Stephen G. West, and Leona S. Aiken. 2009. The Analysis of Count Data: A Gentle Introduction to Poisson Regression and Its Alternatives. *Journal of Personality Assessment* 91, 2: 121–136. https://doi.org/10.1080/00223890802634175

29. Kimberle Crenshaw. 1989. Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics. *University of Chicago Legal Forum* 1989: 139.

30. Shai Danziger, Jonathan Levav, and Liora Avnaim-Pesso. 2011. Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences of the United States of America* 108, 17: 6889–6892. https://doi.org/10.1073/pnas.1018033108

31. Adam Dennett. 2012. *Estimating flows between geographical locations:'Get me started in'Spatial interaction modelling*. Citeseer.

32. Nicholas Diakopoulos. 2015. How Uber surge pricing really works. *The Washington Post*.

33. Tawanna R. Dillahunt and Amelia R. Malone. 2015. The Promise of the Sharing Economy among Disadvantaged Communities. 2285–2294. https://doi.org/10.1145/2702123.2702189

34. Anne-Célia Disdier and Keith Head. 2008. The puzzling persistence of the

distance effect on bilateral trade. *The Review of Economics and statistics* 90, 1: 37–48.

35. Martin Dittus, Giovanni Quattrone, and Licia Capra. 2017. Mass Participation During Emergency Response: Event-centric Crowdsourcing in Humanitarian Mapping. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (CSCW '17), 1290–1303. https://doi.org/10.1145/2998181.2998216

36. Stuart Carter Dodd. 1950. The Interactance Hypothesis: A Gravity Model Fitting Physical Masses and Human Groups. *American Sociological Review* 15, 2: 245–256. https://doi.org/10.2307/2086789

37. Jessica D'Onofrio and Charles Thomas. 2016. Chicago approves new rules on Airbnb, ride-hailing services. *ABC7 Chicago*.

38. Melanie Eckle and João Porto de Albuquerque. 2015. Quality Assessment of Remote Mapping in OpenStreetMap for Disaster Management Purposes. In *ISCRAM*.

39. Benjamin G. Edelman and Michael Luca. 2014. Digital Discrimination: The Case of Airbnb. Com. *Harvard Business School NOM Unit Working Paper*, 14-054.

40. Benjamin Edelman, Michael Luca, and Dan Svirsky. 2015. Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment. *Harvard Business School NOM Unit Working Paper*, 16-069.

41. J. Paul Elhorst. 2010. Applied Spatial Econometrics: Raising the Bar. *Spatial Economic Analysis* 5, 1: 9–28. https://doi.org/10.1080/17421770903541772

42. ESRI. 2016. Location-allocation analysis. *ArcMap Extensions*.

43. Cyrus Farivar. 2016. Chariot for Women, an Uber-like service, set to launch this month. *Ars Technica*.

44. Robin Flowerdew and Murray Aitkin. 1982. A Method of Fitting the Gravity Model Based on the Poisson Distribution*. *Journal of Regional Science* 22, 2: 191–202. `https://doi.org/10.1111/j.1467-9787.1982.tb00744.x`

45. William H Frey and Dowell Myers. 2005. Racial segregation in US metropolitan areas and cities, 19902000: Patterns, trends, and explanations.

46. William H Frey. 2015. Census shows modest declines in black-white segregation. *The Brookings Institution*.

47. Evgeniy Gabrilovich and Shaul Markovitch. 2007. Computing semantic relatedness using Wikipedia-based explicit semantic analysis. In *IJcAI*, 1606–1611.

48. Lee Gaines. 2016. Prompted by luxury treehouse, Schaumburg imposes tax on Airbnb rentals. *chicagotribune.com*.

49. Ruth García-Gavilanes, Yelena Mejova, and Daniele Quercia. 2014. Twitter Ain't Without Frontiers: Economic, Social, and Cultural Boundaries in International Communication. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing* (CSCW '14), 1511–1522. `https://doi.org/10.1145/2531602.2531725`

50. Joel Garreau. 1992. *Edge City: Life on the New Frontier*. Anchor, New York.

51. Yanbo Ge, Christopher R. Knittel, Don MacKenzie, and Stephen Zoepf. 2016. *Racial and Gender Discrimination in Transportation Network Companies*. National Bureau of Economic Research.

52. Yanbo Ge, Christopher R. Knittel, Don MacKenzie, and Stephen Zoepf. 2016. *Racial and Gender Discrimination in Transportation Network Companies*. National Bureau of Economic Research.

53. Nicholas Generous, Geoffrey Fairchild, Alina Deshpande, Sara Y Del Valle, and Reid Priedhorsky. 2014. Global Disease Monitoring and Forecasting with

Wikipedia. *PLoS Computational Biology* 10, 11: e1003892. https://doi.org/10.1371/journal.pcbi.1003892

54. Darren Gergle and Brent Hecht. 2010. The tower of Babel meets web 2.0. In *The 28th international conference*, 291. https://doi.org/10.1145/1753326.1753370

55. M. Gibson and M. Pullen. 1972. Retail turnover in the East Midlands: A regional application of a gravity model. *Regional Studies* 6, 2: 183–196. https://doi.org/10.1080/09595237200185161

56. Ruediger Glott, Philipp Schmidt, and Rishab Ghosh. 2010. Wikipedia Survey - Overview of Results.

57. Mareike Glöss, Moira McGregor, and Barry Brown. 2016. Designing for Labour: Uber and the On-Demand Mobile Workforce. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 1632–1643.

58. Michael F Goodchild. 2007. Citizens as sensors: The world of volunteered geography. *GeoJournal* 69, 4: 211–221. https://doi.org/10.1007/s10708-007-9111-y

59. Peter Gould and Rodney White. 1986. *Mental Maps*. Routledge, London; New York.

60. Simon Gröchenig, Richard Brunauer, and Karl Rehrl. 2014. Estimating Completeness of VGI Datasets by Analyzing Community Activity Over Time Periods. 3–18. https://doi.org/10.1007/978-3-319-03611-3_1

61. Mordechai Haklay. 2010. How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. *Environment and Planning B: Planning and Design* 37, 4: 682–703. https://doi.org/10.1068/b35097

62. A Halfaker, Aaron Halfaker, J T Morgan, and J Riedl. 2013. The Rise and

Decline of an Open Collaboration System: How Wikipedia's Reaction to Popularity Is Causing Its Decline. *American Behavioral Scientist* 57, 5: 664–688. https://doi.org/10.1177/0002764212469365

63. Andrew Hall, Sarah McRoberts, Jacob Thebault-Spieker, Yilun Lin, Shilad Sen, Brent Hecht, and Loren Terveen. 2017. Freedom Versus Standardization: Structured Data Generation in a Peer Production Community. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (CHI '17), 6352–6362. https://doi.org/10.1145/3025453.3025940

64. Mark E Hanson. 1966. *Project METRAN: An integrated, evolutionary transportation system for urban areas*. Cambridge, Mass.: MIT Press.

65. Darren Hardy, James Frew, and Michael F. Goodchild. 2012. Volunteered geographic information production as a spatial process. *International Journal of Geographical Information Science* 26, 7: 1191–1212. https://doi.org/10.1080/13658816.2011.629618

66. Darren Hardy, James Frew, and Michael F Goodchild. 2012. Volunteered geographic information production as a spatial process. *International Journal of Geographical Information Science* 26, 7: 1191–1212. https://doi.org/10.1080/13658816.2011.629618

67. Benjamin Harris and Melissa Kearney. 2014. The Unequal Burden of Crime and Incarceration on America's Poor. *The Brookings Institution*.

68. Brent Hecht and Darren Gergle. 2009. Measuring Self-Focus Bias in Community-Maintained Knowledge Repositories. In *Communities and Technologies 2009: 4th International Conference on Communities and Technologies*, 11–19.

69. Brent Hecht and Darren Gergle. 2010. On the "localness" of user-generated content. 229. https://doi.org/10.1145/1718918.1718962

70. Brent Hecht and Monica Stephens. 2014. A Tale of Cities: Urban Biases in Volunteered Geographic Information. In *Eighth International AAAI Conference on Weblogs and Social Media*.

71. Brent Hecht, Johannes Schöning, Muki Haklay, Licia Capra, Afra J Mashhadi, Loren Terveen, and Mei-Po Kwan. 2013. Geographic human-computer interaction. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, 3163. https://doi.org/10.1145/2468356.2479637

72. Brent Hecht. 2013. The Mining and Application of Diverse Cultural Perspectives in User-generated Content. Northwestern University, Evanston, IL, USA.

73. James Heyman and Dan Ariely. 2004. Effort for Payment A Tale of Two Markets. *Psychological Science* 15, 11: 787–793. https://doi.org/10.1111/j.0956-7976.2004.00757.x

74. Jon Hilkevitch. 2014. Chicago halts Uber try at airport pickups. *tribunedigital-chicagotribune*.

75. Marit Hinnosaar, Toomas Hinnosaar, Michael Kummer, and Olga Slivko. 2017. Wikipedia Matters.

76. Ryan Hughes and Don MacKenzie. 2016. Transportation network company wait times in Greater Seattle, and relationship to socioeconomic indicators. *Journal of Transport Geography* 56: 36–44. https://doi.org/10.1016/j.jtrangeo.2016.08.014

77. Tapio Ikkala and Airi Lampinen. 2015. Monetizing Network Hospitality: Hospitality and Sociability in the Context of Airbnb. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (CSCW '15), 1033–1044. https://doi.org/10.1145/2675133.2675274

78. Panagiotis G Ipeirotis, Foster Provost, and Jing Wang. 2010. Quality

management on Amazon Mechanical Turk. In *The ACM SIGKDD Workshop*, 64. https://doi.org/10.1145/1837885.1837906

79. Isaac Johnson and Brent Hecht. 2015. Structural Causes of Bias in Crowd-derived Geographic Information: Towards a Holistic Understanding. *AAAI Spring Symposium on Observational Studies through Social Media and Other Human-Generated Content*.

80. Isaac L. Johnson, Yilun Lin, Toby Jia-Jun Li, Andrew Hall, Aaron Halfaker, Johannes Schöning, and Brent Hecht. 2016. Not at Home on the Range: Peer Production and the Urban/Rural Divide. 13–25. https://doi.org/10.1145/2858036.2858123

81. Isaac L. Johnson, Connor J McMahon, Johannes Schöning, and Brent Hecht. 2017. The Effect of Population and "Structural" Biases on Social Media-based Algorithms – A Case Study in Geolocation Inference Across the Urban-Rural Spectrum. In *Proceedings of the 35th Annual ACM Conference on Human Factors in Computing Systems (CHI 2017)*. https://doi.org/http://dx.doi.org/10.1145/3025453.3026015

82. Isaac L. Johnson, Subhasree Sengupta, Johannes Schöning, and Brent Hecht. 2016. The Geography and Importance of Localness in Geotagged Social Media. 515–526. https://doi.org/10.1145/2858036.2858122

83. David Jurgens, Tyler Finethy, James McCorriston, Yi Tian Xu, and Derek Ruths. 2015. Geolocation Prediction in Twitter Using Social Networks: A Critical Analysis and Review of Current Practice. In *ICWSM*, 188–197.

84. Leo H. Kahane. 2013. Understanding the Interstate Export of Crime Guns: A Gravity Model Approach. *Contemporary Economic Policy* 31, 3: 618–634. https://doi.org/10.1111/j.1465-7287.2012.00324.x

85. Matthew Kay, Cynthia Matuszek, and Sean A. Munson. 2015. Unequal Representation and Gender Stereotypes in Image Search Results for Occupations. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (CHI '15), 3819–3828. https://doi.org/10.1145/2702123.2702520

86. Bonnie L Keeler, Spencer A Wood, Stephen Polasky, Catherine Kling, Christopher T Filstrup, and John A Downing. 2015. Recreational demand for clean water: Evidence from geotagged photographs by visitors to lakes. *Frontiers in Ecology and the Environment* 13, 2: 76–81. https://doi.org/10.1890/140124

87. Aniket Kittur, Boris Smus, Robert E Kraut, and Robert Kraut. 2011. CrowdForge. In *The 2011 annual conference extended abstracts*, 1801. https://doi.org/10.1145/1979742.1979902

88. Won W. Koo and David Karemera. 1991. Determinants of World Wheat Trade Flows and Policy Analysis. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie* 39, 3: 439–455. https://doi.org/10.1111/j.1744-7976.1991.tb03585.x

89. Won W. Koo, David Karemera, and Richard Taylor. 1994. A gravity model analysis of meat trade policies. *Agricultural Economics* 10, 1: 81–88. https://doi.org/10.1016/0169-5150(94)90042-6

90. Gautier Krings, Francesco Calabrese, Carlo Ratti, and Vincent D. Blondel. 2009. Urban gravity: A model for inter-city telecommunication flows. *Journal of Statistical Mechanics: Theory and Experiment* 2009, 07: L07003. https://doi.org/10.1088/1742-5468/2009/07/L07003

91. Florence Ladd. 1967. A note on "the world across the street". *Harvard Graduate School Education Association Bulletin* 12, 1967: 47–49.

92. Shyong Tony K Lam, Anuradha Uduwage, Zhenhua Dong, Shilad Sen, David

R Musicant, Loren Terveen, and John Riedl. 2011. WP:Clubhouse? In *The 7th International Symposium*, 1. https://doi.org/10.1145/2038558.2038560

93. Airi Lampinen. 2016. Why we need to examine multiple social network sites. *Communication and the Public* 1, 4: 489–493. https://doi.org/10.1177/2057047316681171

94. Sarah Leberstein. 2016. Uber's car leasing program turns its drivers into modern-day sharecroppers. *Quartz*.

95. David LeBlanc. 1999. About Slugging. *Slug-Lines.com*.

96. Lung-fei Lee. 2007. Identification and estimation of econometric models with group interactions, contextual factors and fixed effects. *Journal of Econometrics* 140, 2: 333–374. https://doi.org/10.1016/j.jeconom.2006.07.001

97. Min Kyung Lee, Daniel Kusbit, Evan Metsky, and Laura Dabbish. 2015. Working with Machines: The Impact of Algorithmic and Data-Driven Management on Human Workers. 1603–1612. https://doi.org/10.1145/2702123.2702548

98. James LeSage and R Kelley Pace. 2009. *Introduction to Spatial Econometrics*. Chapman and Hall/CRC.

99. Linna Li, Michael Goodchild, and Bo Xu. 2013. Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr. *Cartography and Geographic Information Science* 40, 2: 61–77. https://doi.org/10.1080/15230406.2013.777139

100. David Liben-Nowell, Jasmine Novak, Ravi Kumar, Prabhakar Raghavan, Andrew Tomkins, and Ronald L. Graham. 2005. Geographic Routing in Social Networks. *Proceedings of the National Academy of Sciences of the United States of America* 102, 33: 11623–11628.

101. Michael D Lieberman and Jimmy Lin. 2009. You Are Where You Edit: Locating Wikipedia Contributors through Edit Histories. In *ICWSM*.

102. New York City Taxi & Limousine Commision. 2016. New York City Taxi & Limousine Commision Trip Data. *TLC Trip Record Data*.

103. Tessa Love. 2016. Study shows Airbnb's growth overcomes hotels. *San Francisco Business Times*.

104. Kevin Lynch. 1960. *The Image of the City*. MIT Press.

105. Xiao Ma, Jeffery T. Hancock, Kenneth Lim Mingjie, and Mor Naaman. 2017. Self-Disclosure and Perceived Trustworthiness of Airbnb Host Profiles. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (CSCW '17), 2397–2409. https://doi.org/10.1145/2998181.2998269

106. Momin M. Malik, Hemank Lamba, Constantine Nakos, and Jürgen Pfeffer. 2015. Population Bias in Geotagged Tweets. In *Ninth International AAAI Conference on Web and Social Media*.

107. Charles F. Manski. 1993. Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies* 60, 3: 531–542. https://doi.org/10.2307/2298123

108. Afra Mashhadi, Giovanni Quattrone, Licia Capra, and Peter Mooney. 2012. On the accuracy of urban crowd-sourcing for maintaining large-scale geospatial databases. In *The Eighth Annual International Symposium*, 1. https://doi.org/10.1145/2462932.2462952

109. Sorin Matei, Sandra J. Ball-Rokeach, and Jack Linchuan Qiu. 2001. Fear and Misperception of Los Angeles Urban Space A Spatial-Statistical Study of Communication-Shaped Mental Maps. *Communication Research* 28, 4: 429–463. https://doi.org/10.1177/009365001028004004

110. Shannon McDonald and Lucas Shubert. 2016. Transportation Conditions and Solutions in Carbondale, Illinois. *Journal of Applied Sciences and Arts* 1, 1.

111. Robert McGill, John W. Tukey, and Wayne A. Larsen. 1978. Variations of Box Plots. *The American Statistician* 32, 1: 12–16. https://doi.org/10.2307/2683468

112. Tom L. McKnight. 2004. *Regional Geography of the United States and Canada*. Prentice Hall, Upper Saddle River, N.J.

113. Miller McPherson, Lynn Smith-Lovin, and James M Cook. 2001. Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology* 27, 1: 415–444. https://doi.org/10.1146/annurev.soc.27.1.415

114. Amanda Menking and Ingrid Erickson. 2015. The Heart Work of Wikipedia: Gendered, Emotional Labor in the World's Largest Online Encyclopedia. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (CHI '15), 207–210. https://doi.org/10.1145/2702123.2702514

115. Tad Milbourn. 2015. In the Future, Employees Won't Exist. *TechCrunch*.

116. Alan Mislove, Sune Lehmann, Yong-Yeol Ahn, Jukka-Pekka Onnela, and J Niels Rosenquist. 2011. Understanding the Demographics of Twitter Users. *ICWSM* 11: 5th.

117. Lewis Mitchell, Morgan R. Frank, Kameron Decker Harris, Peter Sheridan Dodds, and Christopher M. Danforth. 2013. The Geography of Happiness: Connecting Twitter Sentiment and Expression, Demographics, and Objective Characteristics of Place. *PLoS ONE* 8, 5: e64417. https://doi.org/10.1371/journal.pone.0064417

118. Peter Mooney, Padraig Corcoran, and Adam C Winstanley. 2010. Towards quality metrics for OpenStreetMap. In *The 18th SIGSPATIAL International Conference*, 514. https://doi.org/10.1145/1869790.1869875

119. Stuart Moran, Nadia Pantidi, Tom Rodden, Alan Chamberlain, Chloe Grif-

fiths, Davide Zilli, Geoff Merrett, and Alex Rogers. 2014. Listening to the Forest and Its Curators: Lessons Learnt from a Bioacoustic Smartphone Application Deployment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '14), 2387–2396. https://doi.org/10.1145/2556288.2557022

120. Mohamed Musthag and Deepak Ganesan. 2013. Labor Dynamics in a Mobile Micro-task Market. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '13), 641–650. https://doi.org/10.1145/2470654.2470745

121. NCHS. 2010. NCHS Urban-Rural Classification Scheme For Counties.

122. Pascal Neis and Alexander Zipf. 2012. Analyzing the Contributor Activity of a Volunteered Geographic Information Project of OpenStreetMap. *ISPRS International Journal of Geo-Information* 1, 3: 146–165. https://doi.org/10.3390/ijgi1020146

123. Eric Newcomer and Olivia Zaleski. 2016. Inside Uber's Auto-Lease Machine, Where Almost Anyone Can Get a Car. *Bloomberg.com*.

124. Andrew Odlyzko. 2015. The Forgotten Discovery of Gravity Models and the Inefficiency of Early Railway Networks. *Œconomia. History, Methodology, Philosophy*, 5-2. https://doi.org/10.4000/oeconomia.1684

125. Karen Olsson. 2016. The Short-Lived Career of an Austin Uber Driver. *The New Yorker*.

126. OpenStreetMap. Armchair Mapping - OpenStreetMap Wiki.

127. OpenStreetMap. Good practice - OpenStreetMap Wiki.

128. Katherine Panciera, Aaron Halfaker, and Loren Terveen. 2009. Wikipedians are born, not made. In *Proceedings of the ACM 2009 international conference*, 51. https://doi.org/10.1145/1531674.1531682

129. Katherine Panciera, Reid Priedhorsky, Thomas Erickson, and Loren Terveen. 2010. Lurking? Cyclopaths? In *The 28th international conference*, 1917. https://doi.org/10.1145/1753326.1753615

130. Katherine Panciera. 2011. User lifecycles in cyclopath. In *The 2011 iConference*, 741–742. https://doi.org/10.1145/1940761.1940892

131. Ingmar Poese, Steve Uhlig, Mohamed Ali Kaafar, Benoit Donnet, and Bamba Gueye. 2011. IP Geolocation Databases: Unreliable? *SIGCOMM Comput. Commun. Rev.* 41, 2: 53–56. https://doi.org/10.1145/1971162.1971171

132. Adrian Popescu and Gregory Grefenstette. 2009. Deducing Trip Related Information from Flickr. In *Proceedings of the 18th International Conference on World Wide Web* (WWW '09), 1183–1184. https://doi.org/10.1145/1526709.1526919

133. Reid Priedhorsky and Loren Terveen. 2008. The Computational Geowiki: What, Why, and How. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work* (CSCW '08), 267–276. https://doi.org/10.1145/1460563.1460606

134. Reid Priedhorsky, Jilin Chen, Shyong Tony K Lam, Katherine Panciera, Loren Terveen, and John Riedl. 2007. Creating, destroying, and restoring value in wikipedia. In *The 2007 international ACM conference*, 259. https://doi.org/10.1145/1316624.1316663

135. Giovanni Quattrone, Licia Capra, and Pasquale De Meo. 2015. There's No Such Thing as the Perfect Map: Quantifying Bias in Spatial Crowd-sourcing Datasets. 1021–1032. https://doi.org/10.1145/2675133.2675235

136. Giovanni Quattrone, Afra Mashhadi, Daniele Quercia, Chris Smith-Clarke, and Licia Capra. 2014. Modelling growth of urban crowd-sourced information. In *The 7th ACM international conference*, 563–572. https://doi.org/10.1145/2556195.

2556244

137. Giovanni Quattrone, Davide Proserpio, Daniele Quercia, Licia Capra, and Mirco Musolesi. 2016. Who Benefits from the "Sharing" Economy of Airbnb? *arXiv:1602.02238 [physics]*. Retrieved February 25, 2016 from http://arxiv.org/abs/1602.02238

138. Noopur Raval and Paul Dourish. 2016. Standing Out from the Crowd: Emotional Labor, Body Labor, and Temporal Labor in Ridesharing. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (CSCW '16), 97–107. https://doi.org/10.1145/2818048.2820026

139. Karl Rehrl, Simon Gröechenig, Hartwig Hochmair, Sven Leitinger, Renate Steinmann, and Andreas Wagner. 2013. A Conceptual Model for Analyzing Contribution Patterns in the Context of VGI. In *Progress in Location-Based Services*, Jukka M. Krisp (ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 373–388. https://doi.org/10.1007/978-3-642-34203-5_21

140. William John Reilly. 1931. *The law of retail gravitation*. WJ Reilly.

141. William John Reilly. 1931. *The law of retail gravitation*. WJ Reilly.

142. Paul Resnick. 2004. Dynamic ride sharing in the Bay Area. *Paul Resnick's occasional musings*.

143. Dan Rivoli. 2016. Report shows more Uber cars involved in fatal crashes than cabs. *NY Daily News*.

144. Alex Rosenblat and Luke Stark. 2015. *Uber's Drivers: Information Asymmetries and Control in Dynamic Work*. Social Science Research Network, Rochester, NY.

145. Takeshi Sakaki, Makoto Okazaki, and Yutaka Matsuo. 2010. Earthquake Shakes Twitter Users: Real-time Event Detection by Social Sensors. In *Proceedings*

of the 19th International Conference on World Wide Web (WWW '10), 851–860. https://doi.org/10.1145/1772690.1772777

146. Christian Sandvig, Kevin Hamilton, Karrie Karahalios, and Cedric Langbort. 2014. Auditing algorithms: Research methods for detecting discrimination on internet platforms. *Data and Discrimination: Converting Critical Concerns into Productive Inquiry*.

147. Salvatore Scellato, Anastasios Noulas, Renaud Lambiotte, and Cecilia Mascolo. 2011. Socio-Spatial Properties of Online Location-Based Social Networks. In *Fifth International AAAI Conference on Weblogs and Social Media*.

148. Juliet B Schor and Connor J Fitzmaurice. 2015. 26. Collaborating and connecting: The emergence of the sharing economy. *Handbook of research on sustainable consumption*: 410.

149. Juliet Schor. 2014. Debating the sharing economy. *Great transition initiative*.

150. Shilad W. Sen, Heather Ford, David R. Musicant, Mark Graham, Oliver S.B. Keyes, and Brent Hecht. 2015. Barriers to the Localness of Volunteered Geographic Information. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (CHI '15), 197–206. https://doi.org/10.1145/2702123.2702170

151. Patrick C. Shih, Victoria Bellotti, Kyungsik Han, and John M. Carroll. 2015. Unequal Time for Unequal Value: Implications of Differing Motivations for Participation in Timebanking. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (CHI '15), 1075–1084. https://doi.org/10.1145/2702123.2702560

152. Chris Smith, Daniele Quercia, and Licia Capra. 2013. Finger on the Pulse:

Identifying Deprivation Using Transit Flow Analysis. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work* (CSCW '13), 683–692. https://doi.org/10.1145/2441776.2441852

153. Jennifer Stark and Nicholas Diakopoulos. 2016. Uber seems to offer better service in areas with more white people. That raises some tough questions. *The Washington Post*.

154. Judith Steward. 2009. Economic Status. *MacArtuhur Research Network on SES & Health*.

155. John Q. Stewart. 1948. Demographic Gravitation: Evidence and Applications. *Sociometry* 11, 1/2: 31–58. https://doi.org/10.2307/2785468

156. Bob Sullivan. 2013–2013-04-22T12:57:11-0400. For many households, banking is a costly luxury. *CNBC*.

157. TaskRabbit Support. 2016. Which Category Should I Choose? *TaskRabbit Support*.

158. Peter T. Suzuki. 1985. Vernacular cabs: Jitneys and gypsies in five cities. *Transportation Research Part A: General* 19, 4: 337–347. https://doi.org/10.1016/0191-2607(85)90069-X

159. Yuri Takhteyev, Anatoliy Gruzd, and Barry Wellman. 2012. Geography of Twitter networks. *Social Networks* 34, 1: 73–81. https://doi.org/10.1016/j.socnet.2011.05.006

160. TaskRabbit. 2016. TaskRabbit.

161. B. Taylor. 2006. Price Ceilings. *economics.fundamentalfinance.com*.

162. Rannie Teodoro, Pinar Ozturk, Mor Naaman, Winter Mason, and Janne Lindqvist. 2014. The motivations and experiences of the on-demand mobile workforce. In *The 17th ACM conference*, 236–247. https://doi.org/10.1145/2531602.

2531680

163. Jacob Thebault-Spieker, Loren G. Terveen, and Brent Hecht. 2015. Avoiding the South Side and the Suburbs: The Geography of Mobile Crowdsourcing Markets. In *CSCW 2015*, 265–275. https://doi.org/10.1145/2675133.2675278

164. Jacob Thebault-Spieker, Loren Terveen, and Brent Hecht. 2017. Towards a Geographic Understanding of the Sharing Economy: Systemic Biases in UberX and TaskRabbit. *ACM TOCHI*.

165. Chicago Tribune. 2016. Chicago shooting victims. *Chicago Tribute: Crime in Chicagoland*.

166. Alessandro Venerandi, Giovanni Quattrone, Licia Capra, Daniele Quercia, and Diego Saez-Trumper. 2015. Measuring Urban Deprivation from User Generated Content. 254–264. https://doi.org/10.1145/2675133.2675233

167. Vijaysree Venkatraman. 2010. Time to Hire a Housekeeper? *Science*. https://doi.org/10.1126/science.caredit.a1000056

168. Morten Warncke-Wang, Vivek Ranjan, Loren G Terveen, and Brent J Hecht. 2015. Misalignment Between Supply and Demand of Quality Content in Peer Production Communities. In *ICWSM*, 493–502.

169. James O. Wheeler and Ronald L. Mitchelson. 1989. Information Flows among Major Metropolitan Areas in the United States. *Annals of the Association of American Geographers* 79, 4: 523–543. https://doi.org/10.1111/j.1467-8306.1989.tb00275.x

170. Wikipedia. 2015. Wikipedia:Introduction. *Wikipedia*.

171. Business Wire. 2016. Green Dot and Uber Announce ”Uber Checking by GoBank” | Business Wire.

172. Spencer A. Wood, Anne D. Guerry, Jessica M. Silver, and Martin Lacayo.

2013. Using social media to quantify nature-based tourism and recreation. *Scientific Reports* 3. https://doi.org/10.1038/srep02976

173. Tse-Chuan Yang, Aggie J. Noah, and Carla Shoff. 2013. Exploring Geographic Variation in US Mortality Rates Using a Spatial Durbin Approach: Mortality and Spatial Durbin. *Population, Space and Place* 21, 1: 18–37. https://doi.org/10.1002/psp.1809

174. Ke Zhang and Konstantinos Pelechrinis. 2014. Understanding Spatial Homophily: The Case of Peer Influence and Social Selection. In *Proceedings of the 23rd International Conference on World Wide Web* (WWW '14), 271–282. https://doi.org/10.1145/2566486.2567990

175. Dennis Zielstra and Hartwig Hochmair. 2012. Using Free and Proprietary Data to Compare Shortest-Path Lengths for Effective Pedestrian Routing in Street Networks. *Transportation Research Record: Journal of the Transportation Research Board* 2299: 41–47. https://doi.org/10.3141/2299-05

176. Dennis Zielstra, Hartwig H. Hochmair, Pascal Neis, and Francesco Tonini. 2014. Areal Delineation of Home Regions from Contribution and Editing Patterns in OpenStreetMap. *ISPRS International Journal of Geo-Information* 3, 4: 1211–1233. https://doi.org/10.3390/ijgi3041211

177. Alexander Zipf and Dennis Zielstra. 2010. A comparative study of proprietary geodata and volunteered geographic information for Germany. *Conference on Geographic Information* 1: 1–15.

178. 2009. The big sort: Why the clustering of like-minded America is tearing us apart. *Choice Reviews Online* 47, 04: 47–2303–47–2303. https://doi.org/10.5860/CHOICE.47-2303

179. 2012. Cartographic guidelines for public health.

180. 2017. Reservation poverty. *Wikipedia*.

181. Apple Maps | OpenStreetMap Blog.

182. OpenStreetMap Statistics.