Shared Electric Scooters and Transportation Equity:
A Multivariate Spatial Regression Analysis of Environmental Factors on Revealed Travel Behavior and Mode Shift Potential

by

Bernard Maxwell Arnell

B.A. Psychology
Bates College
Lewiston, Maine (2013)

SUBMITTED TO THE DEPARTMENT OF URBAN STUDIES AND PLANNING IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER IN CITY PLANNING
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2019

Signature of Author:_____
Department of Urban Studies and Planning
May 23, 2019

Certified by:_____
Professor Jinhua Zhao
Professor of City and Transportation Planning
Thesis Supervisor

Accepted by:_____
Professor of the Practice Ceasar McDowell
Co-Chair, MCP Committee
Department of Urban Studies and Planning

Shared Electric Scooters and Transportation Equity:
A Multivariate Spatial Regression Analysis of Environmental Factors on Revealed Travel Behavior and Mode Shift Potential

by

Bernard Maxwell Arnell

Submitted to the Department of Urban Studies and Planning
on May 23, 2019 in Partial Fulfillment of the Requirements for the
Degree of Master in City Planning

ABSTRACT

The past year has seen the emergence of the shared electric scooter: a new form of micro-mobility in the United States. Electric scooters embody many of the historic trends and contradictions endemic to innovation and marginalization in the transportation space. Researchers and policy-makers are faced with a number of unanswered questions about who travels on scooters and what factors might influence when and how they are used. Hexagonal spatial binning feeds a series of spatial lag models to identify explanatory environmental and demographic variables for trip characteristics. These models reveal that employment density and the location of rebalance points are among the strongest indicators of scooter activity overall, but show significant variation when models are subsetted by time of day. The Communities of Concern framework, adapted from the Association of Bay Area Governments, provides a regionally-sensitive index of relative marginalization to undergird the analysis. These findings are placed in context with transportation justice and broad outlines of current e-scooter policy. Additionally, a framework for examining mode shift potential using Open Trip Planner is discussed, and preliminary results are described. Use patterns are also analyzed by time and in relation to contemporaneous weather.

Thesis Supervisor: Jinhua Zhao
Title: Associate Professor of City and Transportation Planning
Thesis Reader: Eric M. Huntley
Title: Technical Instructor of GIS, Data Visualization, and Graphics
Thesis Reader: Peyman Noursalehi
Title: Postdoctoral Associate

Acknowledgements

The creation of this work was the result of the combined efforts, brilliance, and unending intellectual and emotional generosity of a vast network of people, far too many to name. If you are not mentioned below, thank you.

I wish first to recognize the tremendous work and guidance provided by my thesis committee. Jinhua, your incisive feedback and clear sense of exactly what this thesis needed at every stage was indispensable, and I am honored to have worked with you. Peyman, thank you for being my unflagging ally in the fight against uncooperative code at all hours. Eric, the methodological and theoretical insights you always have at the ready pushed my work to the next level after every conversation. Thank you all.

To my advocate inside Gust, I am so thankful for your tireless work on my behalf, and I am truly honored by your trust in me.

Kate, you are unbelievably wonderful, brilliant, and have managed to figure out how to support this thesis at every stage from 3,000 miles away. I love you.
Jessie, you are, without a doubt, the best sister. I've been so happy to have shared a city with you for the last two years, and your tenacity and fortitude are true inspirations.
Mom and Dad, it almost goes without saying that I would not be who I am today without your endless love, kindness, and belief in my ability to (eventually) accomplish whatever was in front of me.

To my DUSP and JTL community, I am so grateful to have had the chance to get to know you over the past two years. I have learned more than I could have possibly imagined; we have been friends, road trip companions, staunch academic antagonists, roommates, 4 AM work compatriots, board gamers, advocates, and collectives. Thank you.

Contents

Figures

Chapter 1 – Context and Theoretical Framing

Introduction

Shared electric scooters dominated the news in cities across the United States in 2018. People took millions of rides on these new devices, and policy-makers and transportation planners have faced huge questions about how people use scooters, who uses them, when they use them, and where they go, among many others. This thesis is an attempt to address parts of that larger conversation.

I begin by setting the context of a world with cities on the rise, but beleaguered by environmental concerns. Next, I explore the evolution in our transportation systems that have created a moment ripe for innovation. I then detail the factors of equity, marginalization, and justice that cross-cut the first two domains and guide this research.

Following this introduction, I provide a brief overview of the analysis steps to provide context for detailed explanations of the data sources. After thoroughly describing the source material, I provide an in-depth description of the Methodology, including the iterative history of the final procedure. I then present the results in a mixture of tables, maps, and graphs, contextualized and explained in the following Discussion section. Next, I describe a framework for future micro-mobility mode shift research. Finally, I draw inferences from these results and discuss possible implications, as well as limitations, future research directions, and policy ramifications and recommendations.

Climate Crisis

We live in a deeply interconnected world undergoing rapid change. For the vast majority of human history, most people have lived in smaller rural communities (Clark, 1951). While cities have existed across the world for thousands of years as centers of culture and commerce, most societies were still primarily rural by population. Over centuries, technology evolved to facilitate life in closer quarters (Boserup, 1981). From ancient devices like the aqueduct and paper to industrial revolution steam engines, many prerequisites to action have been redefined. While there has been no globally coordinated goal, the general trend has been to innovate in a way that reduces the work or resources required by the end user (Hekkert, Suurs, Negro, Kuhlmann, & Smits, 2007). However, this does not imply a reduction of human labor or resources across the system required to facilitate the invention. The reapportionment of work

and resource needs is a major aspect of human development, and continues to be an important factor in the modern world (Daly, 1996).

The global population has grown exponentially, made possible in part by significant advances in baseline disciplines like agriculture and medicine (D. of E. and S. A. United Nations, 2015). People are able to live longer, farm more efficiently, and distribute that food around the world to sustain populations scattered across the earth. Technological advances have rapidly increased quality of life for billions to an unprecedented degree at an unprecedented rate. The rapid spread of specialized knowledge has increased with the global penetration of the internet, and the transference of ideas and technology is no longer mediated by physical distance (D. of E. and S. A. United Nations, 2015).

However, while we have advanced as a species along many axes of quality of life, there are caveats that significantly undermine the narrative of universally beneficent progress. Massive environmental degradation has been the direct outcome of many of the systems developed to reduce individual work by the end user. This destruction and disruption of the non-human world has taken place across the globe with little foresight or consideration of what the downstream effects (in time and geography) might be (Bartlett, n.d.). Localized effects on non-human populations have included loss of habitat and extinction, and the non-local effects of this mode of human development may be even more significant. Human activity has significantly altered the climate of the Earth, with unpredictable but generally drastic results. Extreme weather events now occur more frequently, sea levels are rising, and consistent weather patterns that humans have relied on for millennia to underpin agricultural cycles are full of aberrations. The effects on the planet, of which humanity is just a subset of the affected parties, have been and will continue to be dire.

It has become clear that immediate and decisive action is required to stave off global climactic catastrophe. The recent UN report (I. S.-P. P. United Nations, 2019) is just the latest in a string of urgent signs (Intergovernmental Panel on Climate Change, 2018) that action must be taken on many levels. Regulation of industry is one of the most important aspects of the needed steps forward, in such a way that these regulations shape the choices available to other businesses and consumers (Tietenberg, 1990). The production of resource-intensive materials and objects is a major contributor to greenhouse gas emissions, but this activity also underpins many of the major technology advances previously described. Electronics, especially small portable ones, rely heavily on uncommon materials such as rare earth metals that are mined in a destructive and energy intensive way (Khan & Kushler, 2013). Once devices are produced, they

consume a substantial amount of electric power, which brings this discussion to the next major area for environmental reforms.

Historically, the production of power has been based on the combustion of chemically potent materials, which release energy that is used to generate either heat (as in a steam engine) or direct kinetic energy (as in an internal combustion engine). These fuels have usually been "found" materials, naturally existing substances that hold a great deal of chemical potentially energy in the form of hydrocarbons. When combusted in the presence of oxygen atoms, these long hydrocarbon molecules split and re-bond in an exothermic reaction, releasing significant amounts of energy that had been stored as bonds within the concatenated hydrocarbon molecules. In the context of a discussion on global environmental catastrophe, the derivation of the majority of the world's energy from hydrocarbon combustion has two main implications (Ashford & Caldart, 2008) .

The first implication is declining availability of easily available complex hydrocarbons. Coal mining and drilling for oil or natural gas have massive environmental ramifications even before the desired substance is combusted (Desai, 2002) . Hydrocarbons in even moderate quantities above what typically occurs on the surface of the earth is poisonous to most life. These extractive practices usually end up increasing the localized concentration of these substances, even if measures are taken to mitigate this effect.

The other major impact of a global reliance on fossil fuels is the release of carbon dioxide, a direct chemical byproduct of the exothermic combustion reaction that makes hydrocarbons an effective fuel in the first place. Carbon dioxide has been part of the atmosphere on Earth for eons, and nearly all fauna exhale carbon dioxide as a byproduct of normal life. However, the global concentration of this chemical has been increasing rapidly since the onset of the industrial revolution, with destabilizing effects on the atmosphere (Sagoff, 2007) . Carbon dioxide is a greenhouse gas, and traps heat from the sun within the Earth's atmosphere more than the standard mix of atmospheric components. The excess heat in the atmosphere serves to not only increase the global average temperature, but also to modify global weather patterns and circulatory currents that are driven by the relative heat of the air and water.

While industrial scale power production is a primary use of fossil fuels, two other uses for these substances interface more directly with most of our lives: the creation of plastics and synthetics, and transportation. The internal combustion engine is dominant in many domains, especially those that prioritize scale. Cars and trucks, let alone ships and airplanes, nearly all run on the direct combustion of hydrocarbon fuels. In the United States, the transportation sector produces the largest share of greenhouse gases. This sector division includes industrial

transportation, but personal, non-commercial transportation still makes up a significant share (Lawrence Livermore National Laboratory, 2019). The US has debated a number of policies related to decreasing the environmental impact of the transportation section on the national level (Nivola & Crandall, 1995), but it appears increasingly clear that, at least in the United States, these policies must pursued at the city or state level (Burns, Jordan, & Scarborough, 2013).

Critical Geography

This work is deeply informed by a core tenet of critical technology discourse: technology is not neutral, and is inherently political (Stiegler, 2010) . Technology is a broad category here, encompassing not just physical objects like a wristwatch but also concepts and practices, including foundational concepts like language. Political does not describe electoral politics, but rather the idea that a technology is imbued with the policy preferences of those who create and define its place in society. When viewed through this lens, the interconnected technologies that shape the modern world can be seen as an amalgamation of expressions of policy and cultural priority, aggregated and reconstituted in ways that can reinforce or counteract the inherent politics of the underlying components.

In some cases, an object has an intentional designer, who may design features and functions of the technology to reflect what they view as the most important attributes of the technology (Stiegler, 2010) . These may or may not align with what users of the technology view as the most important attributes. Furthermore, these design choices serve to reinforce the biases inherent in the design team. For example, for decades, Band-Aids were sold primarily in a cream-tan color, described as neutral, or “skin-toned” (Malo, 2013) . As we know, this color is only skin-tone for the culturally dominant segment of the population. By implicitly portraying white skin as the default option, this marketing and product choice served to enforce a racist hierarchy across society, placing the needs of non-white individuals below those of the white population. The ways in which other parts of society reorient to meet the needs of a group of people can be seen an effective proxy for their standing in that larger society. The manufacturer is not the only actor bearing responsibility for politics of this technology. Even if Band-Aids in other skin tones had always been available, the dominant cultural narrative of whiteness-as-default could be reinforced by store owners who would choose to only stock one color. The choices made by the manufacturer and the store serve to propagate the politics of these actors, even if driven by neoliberal capitalist imperatives. A technology ostensibly designed for healthcare becomes a tool of white supremacy. In other cases, such as with language, the

"technology" (specific codes of phonemes understood to represent shared meanings) evolves in a distributed fashion, and the form and routines of language become reflective of cultural priorities.

In critical geography and critical urban studies, we reflect particularly on the technologies that define. These technologies include property rights, borders, and techniques of exclusion. Edward Soja, in the landmark Seeking Spatial Justice, (Soja, 2011) describes in detail how so-called public spaces in the United States are designed or managed with a particular subset of the public in mind.

These core understandings of critical geography run through Lynn Staeheli and Don Mitchell's The People's Property (Staeheli & Mitchell, 2008) . This book posits that "close attention to the people *in* spaces and those who occupy and use property can help us understand the ways in which politics, power, and publicity are constituted and enacted in the formation of political communities; these configurations of community hold different possibilities for citizenship, democracy, and justice." (Staeheli & Mitchell, 2008). This call to critical interrogation of the meanings of public space and the technologies that fill and define it are core motivators for this research.

Transportation

To lay out policies intended to address the fundamental social and environmental disasters in the United States transportation sector, one must have a grasp on the functional landscape. This consideration of the context of mobility and the built environment in the U.S. is critical to understanding the role and implications of the new technological and social phenomenon that is the focus of this paper. The United States has a long and convoluted history of mobility innovations, many of which have been exported as cultural norms across the globe with wide-ranging results. As stated, a crucial aspect of critical urban theory is the recognition that technology has politics. When viewed through the lens of critical urban theory, the functional attributes of the transportation system become representative of the political and social priorities of systemic racism and patriarchy, a non-severable aspect of understanding how transportation and societal organization evolve in tandem.

In pre-colonial North America, the traction complex was not a deeply embedded part of social organization (Jacobsen & Eighmy, 1980). Animals that were harnessed for transportation and strength did not emerge until horses arrived in the Americas with the first European ships. Horses spread across the continent far faster than colonizing settlers, and soon became an important part of Native American cultures and society. Along with the wheel, horses were the historically dominant innovation in land transportation beyond walking and running.

Various combinations of these two technologies remained preeminent along a primary dimension of human innovation. This dimension is best summarized as attempting to reduce the work and resources required by the end user to achieve a similar effect, or in a slight reformulation, achieving a stronger effect using the same level of work or resources (Hekkert et al., 2007) .

In the transportation context, the steam engine was a revolutionary advance on this concept. The steam engine used static fuel, burned to create heat, which in turn boiled water to create steam. Gas exerts increasing pressure on its container as its temperature increases, and this force drives pistons and axels. Steam power revolutionized a number of industries, but the portability advantage over other harnessed power sources of the time, notably mills driven by hydropower, was particularly advantageous for transportation applications. The steam-powered locomotive was the most wide-spread example of this new transportation application.

The railroad traction complex radically reshaped societies across the globe, and this effect was particularly profound in the United States (Cronon, 1992) . New cities and towns were soon linked by a mode of transportation that was able to travel faster and further than any land transportation (Rodrigue, 2017). However, unlike marine or riverine transportation, the number of people who could use this new mode was tightly limited by the inherently exclusive nature of railroad tracks. A high personal resource barrier to entry for rapid transportation was not new; horses, stagecoaches, oceanic ships, and the like were all far more available to those with wealth. However, they generally made use of the same basic infrastructure, dirt roads and ports, that were used by people on foot or in dinghies. The “transportation commons” was used by rich and poor alike.

The growth of train travel reshaped American society along many intersectional axes of analysis, and these shifts have been well-documented elsewhere. As this research focuses primarily on the implications of a new mode of transportation within a city (rather than between cities), many of these major structural changes, such as the advent of time zones, lie outside the realm of our analysis. However, as rail technology matured and became both smaller and more reliable, some transportation lines began to be constructed within, rather than between, cities. Like the majority of railroads being constructed between cities, these projects were planned and constructed primarily by private companies and corporations chartered for this purpose (Warner, 1978). Many of the first urban rail lines used vehicles on rails that were still pulled by animals. There were some steam powered lines, but electrical distribution and electric traction supercharged urban rail and streetcar construction in the early 1900’s.

Public Space

Before outlining the implications of transportation capabilities of new streetcar systems on populations and U.S. society in general, I want to explore the radical shift in the demarcation of public space that accompanied this building boom. I spoke before of the "transportation commons." By this, I refer to the idea that public space belongs to everyone, and people are equally entitled to its use (Ostrom, 1990). Of course, this lofty goal has always been unmet (Soja, 2011). Especially in the United States, systems of white supremacy and patriarchy have always made public spaces more accommodating to white men, while presenting unpredictability and danger to women and people of color. The *de jure* neutrality of public space, celebrated by de Tocqueville and foundational to our conception and organization of society, has yet to be realized *de facto*. Public urban space in United States before the spread of the streetcar was closed to some publics, but it remained more neutral along the axis of wealth. As described earlier, rapid or comfortable movement modes used by the wealthy generally shared a public right of way with anyone else who was using it.

Streetcar systems took many forms in different parts of the country. In some cities, companies constructed and operated the whole network, existing at least partially in publicly owned spaces. As with state governments and inter-city rail permits, local governments authorized or chartered construction activity to modify the public space to meet the unique needs of a railroad. Primarily, this including the laying of tracks and platforms. The stringing of power lines often followed after the shift away from horse-drawn vehicles towards electricity. Local governments had compelling reasons to acquiesce to these applications from streetcar companies. This technology was touted as a significant economic development strategy, a way to draw new residents, a way to increase city efficiency, and a way to make streets cleaner and improve public health, to name a few common justifications. In many cases, these claims were accurate and for the most part realized (Warner, 1978). However, there was an oft-unacknowledged public cost for these gains.

These rail tracks were not typically space-exclusive. While there was no streetcar present, the street space was generally open to be used for walking, carts, or some other pre-existing mobile uses. However, these corridors of nominally public space had been designated for priority use by the customers of the private streetcar company. Yes, theoretically anyone accepted as an active member of public society, which in the early 1900s in the U.S. was generally limited to white men, could purchase a ticket a partake equally of this public space. This is an early iteration of the free market understanding of government's role in transportation – discrimination based on wealth is a core component of the capitalist system,

and therefore poses no problems when applied to the transportation space (Staeheli & Mitchell, 2008).

One might plausibly argue that it was the responsibility of the government to grant this use of public space, despite its disruptive impacts on some other public space users, because the benefits that would accrue to the whole greatly outweighed the negative impacts. I do not claim that local governments across the country erred in permitting the construction and operation of private competitive streetcar companies. I wish to highlight the partial privatization of public space into a semi-public space to which access was limited by time, as the arrival of a streetcar would necessitate the relocation of whoever had been standing on the tracks.

Streetcars proliferated in cities across the United States, providing the ability for users to quickly move between distant locations without significant personal investment of time or resources. In this regard, the electric streetcar was yet another step in the innovation progression described above. Individual work required to accomplish a similar or improved outcome was lowered, and people could move through urban areas faster and cheaper than ever before. To create a dedicated customer base, many streetcar companies moved into homebuilding, working in tandem with their rail planning to create neighborhoods centered around the ability to walk to a streetcar stop. Residents of these new suburbs were promised more space and a higher quality of life while retaining the ability to quickly access the commercial hub in the city center.

In the United States, this vision of promise was not available to everyone. Many of these new communities had racial restrictions on who was permitted to move in. People needed the financial liquidity to move homes. Under the dominant patriarchal system, many women were expected to stay at home to take care of children, and in this new semi-urban form, this activity was more spatially isolated than it had been in denser urban conditions. Spatial isolation, especially that which was brought upon by a relocation to a new part of the city without existing social connections, frequently resulted in greater social isolation.

A well-known challenge of systems in both the transportation context and beyond is significant path dependency (Nivola & Crandall, 1995). This is the notion that each choice, movement, or evolution of a system shapes the choices available at the next decision point. Sometimes these choices are well recognized and the immediate implications might be clear, but frequently the decision points are obscured, or the magnitude of the adoption of a strategy may not be apparent until much later. Actions, by their nature, preclude the realization of other universes of possibility, and the many possible paths available to both a single individual and a society as a whole are possible because of the prior path decisions. As an example, a region that

invests in coal mining has an incentive to keep this industry relevant. The region does not see previous coal investments of both physical and human capital as sunk costs, as significant additional costs would be incurred by abandoning the industry. While the region could theoretically choose to abandon this economic strategy, the previous policy choices present continuing to invest in coal production as the best and easiest path forward. This choice set is mirrored to a degree on the individual level by residents of this region.

When considering large composite systems not under the conscious control of a single actor or group of actors, the idea of path choice becomes diffuse. The evolution of that system is sometimes thought of as the sum of the choices of the component individuals. In this frame, we assume that most individual choice is reflective of the preferences and value system of that individual, and that the system behavior is therefore somewhat representative of the composite value system of those within it. In some form, this is a major component of the theory of democratic governance. A natural outgrowth of this intuitive formulation is the assumption that the evolution of that system reflects the best path for both the system as a whole, and for the people within it. However, an examination of the broad transportation and land use paradigms in the United States is at odds with that conclusion.

At least two well-recognized phenomena are ignored in this formulation. The first is the tyranny of the majority, and the second is the assumption that individuals are rational actors, and understand and act in their best interest. In a Benthamist utilitarian frame, the best path forward is the one which satisfies the largest number of individuals. A simple voting system acts on this frame, with a yes vote for marginal satisfaction carrying the same weight as a no vote for deep dissatisfaction. Important nuance is added to the utilitarian frame when the analysis goes beyond a binary conception of satisfaction, and registers the degree of utility that each party in the analysis derives from a projected outcome. Broadly, this is the idea behind a cost-benefit analysis, and is frequently used in the U.S. when examining the potential effects of a new policy. Cost-benefit analysis has its own set of flaws outside of the scope of this discussion, but is most appropriate in situations where a system has some degree of central control (Pearce, Atkinson, & Mourato, 2006). The evolution of the U.S. transportation sector was certainly shaped by policy decisions, but the diffuse nature of development policy made this more akin to a rolling vote, where people or towns or states act in service of current understanding of preferences, unfurling across time.

This uncoordinated nature of the development of the United States' transportation network increases the cross-pollination of path dependency as described above, but also amplifies the majoritarian dominance endemic to the utilitarian frame. This functions on the

individual level as well as the municipal level. Individuals typically understand what they choose to be the best choice given their current situation, and in transportation considerations, the external resources and infrastructure available to them. All of these considerations are shaped by the choices of their neighbors, community at large, and region, for the moment setting aside the impact of messaging from non-adjacent parties, such as firms or the media, on what form personal utility maximization might take. The tyranny of the majority amplifies these choices when governments look to the expressed preferences of individuals to guide their policy choices. If this occurs through the democratic voting process, the largest voting bloc will define the expressed preferences of the whole group, sidelining the interest of any minority faction (Madison, 1787). The path dependency of these transportation and infrastructure choices then shapes the preferences available at the next decision point, further minimizing the ability of the larger entity to account for the expressed preferences of any group save the majority.

In the United States, the concept of a voting majority was not simply numeric. A plethora of legal and *de facto* methods were used to enforce white supremacy and ensure that the "majority" preference was the preference of the dominant class, regardless of the population balance of the locale in question. Systematic and intentional marginalization of minority communities took and continues to manifest in many forms, from redlining to voter suppression. Particularly relevant to this transportation story is the decision of cities to prioritize the creation of infrastructure to serve the newly prevalent private automobile over both streetcars and over space used for walking or other, less expensive modes, exacerbating this marginalization on the basis of race and class. Racialized controls on wealth generation make automobiles less available for purchase by minority Americans. Patriarchal oppression made the automotive world a far more open place to men than women, further limiting women's ability to access society. The Americans with Disabilities Act recognized the extent to which full participation in society had been unrealized for people living with differences in physical ability. Age discrimination continues to be prevalent as well.

Personal Mobility in the Early 20th Century

Despite our modern association with paved city streets as the domain of cars, it was actually early bicycle enthusiasts who lobbied for smooth, paved surfaces on which to use their new vehicles. The bicycle, also known as a velocipede, had gained popularity in the late 1800s in the United States. The bicycle provided substantially increased mobility potential, improving speed while reducing human energy output (di Prampero, 1986). Furthermore, use of a bicycle

did not require external energy sources. In terms of maintenance, cost, and resource use, bicycles were significantly less demanding than horses. Horses provide significant advantages over bicycles in many domains, especially outside of cities. Within the urban domain, the convenience of an inert personal mobility technology marked a major improvement for many people.

This mobility revolution was especially important for women. The implications of this technology on women in cities has recently been described by Sarah Hallenbeck (Hallenbeck, 2016). Famously, Susan B. Anthony claimed in 1896 that the bicycle had "…done more for the emancipation of women than anything else in the world" (Hallenbeck, 2012). The contributions of the bicycle in societal liberation of women were not limited to the functional aspects of mobility, along the transportation innovation axes of output and resources described earlier. Flexible mobility meant that women could gather and associate on their own terms, a critical element of the growing suffragette movement.

While today we view this as an unequivocal societal good, whether or not this was a positive development was hotly contested at the time. The patriarchal objections to increased mobility for women via bicycles were often couched in an unfounded but medicalized fear of the effects of the bicycle on the female anatomy and psyche. There were also concerns that mobility would prompt promiscuity and a neglect of the household tasks that were deemed the sole province of women at the time. The ensuing moral panic was a pervasive feature of political discourse through much of the 20th century. However, bicycle manufacturers recognized the demand for this technology, and began marketing directly to female consumers. Establishment of female bicycle clubs and societies made progress towards "regendering" the bicycle (Hallenbeck, 2016).

Paved roads in cities and towns replaced packed dirt or cobblestones, and these smooth surfaces led to faster and more comfortable travel for cyclists. The American Cycling Association led this national lobbying effort. This new infrastructure network also opened the door to other devices that used similar ground interfaces.

In 1913, the Autoped Company of America began producing a motorized personal scooter (Mansky, n.d.). The form factor of this scooter was much the same as modern iterations, with the exception of a compact gasoline motor. Over the next few decades, the Autoped waxed and waned in popularity, particularly in urban areas and beach destinations. Just as the bicycle revolution of previous decades, the adoption of this new machine generated alarm in some quarters for the same reasons it was celebrated in others. Namely, it provided freedom and flexibility for those who were often denied those liberties. The motor scooter gained prominence

among suffragettes, and number of prominent women were public champions of these vehicles, including Amelia Earhart, pictured below.

Figure 1.1 – Amelia Earhart with June Travis and a motorized scooter – 1935 (Wild, n.d.)

However, within a few years of their arrival, the wide viability of these personally owned scooters as a transportation paradigm began to wane. The reasons for this decline may have been diverse, but a major factor was the same set of policy decisions that sidelined bikes, kneecapped streetcars, and pushed walking to the public periphery: the rise of cars (Mansky, n.d.).

Local governments exert significant control over the allocation and disbursement of publicly owned space. The pattern of reserving public space for particular modes of transportation that had gained prominence during the spread of streetcars had continued as streets were paved for use by bicyclists, and later scooters. This marked a departure from streetcar allocation in terms of the vehicle ownership model. Cyclists needed to own their own bicycle to take advantage of this newly prioritized slice of public space in the intended fashion.

However, the speed that cyclists and Autopeds traveled, below 20 miles per hour and in a form factor not much larger than a human body (and significantly smaller than an animal-drawn vehicle), did not categorically exclude use of the street by pedestrians.

The non-exclusive nature of street allocation shifted dramatically as automobile technology improved in the late 1910's and early 1920's. Advocates of the automobile wanted to be able to use this technology to its fullest technological capacity at the time, with the eternal goal of achieving a better outcome with the same amount of work. Here, the desired improved outcome was higher speed, and automobile advocates successfully pushed cities across the country to raise the speed limit on city streets. With permitted speeds far above those of a bicycle or a scooter, let alone a pedestrian, these paved spaces were suddenly unsafe for anyone not inside an automobile. Soon, this reservation of street space for exclusive automobile use was codified in law. Other users were often relegated to the sides of streets, on sidewalks separated by the newly prevalent grade differential curb.

It is important to note that automobiles, while they were able to move faster and achieve a "better result" than walking, reduced only the immediate work required by the end user, and did not reduce the overall resource use required to achieve this result. Rather, the consumption of resources for this activity was widely dispersed, falling on the shoulders of people and communities across the city and globe, as described earlier in the section on the Climate Crisis.

The physics of simple human locomotion vary greatly by age, gender, and disability status. Normal gait walking speed of adults age 20-59 average approximately 80 meters per minute (about 3 miles per hour), with a standard deviation of about 10 meters per minute. The table below is excerpted from (Waters, Lunsford, Perry, & Byrd, 1988)

Group	Velocity (m/min) Normal	Slow	Fast
Children (6–12 yr)			
F	68.29	53.55	88.48
	9.19	8.02	12.40
M	70.72	57.18	87.59
	8.02	6.29	11.79
T	69.64	55.55	87.96
	8.57	7.28	11.96
Teens (13–19 yr)			
F	73.16	57.23	98.15
	8.96	8.14	16.19
M	73.41	56.49	98.96
	11.58	10.83	15.43
T	73.28[c]	56.89	98.53[c]
	10.17	9.39	15.69
Adults (20–59)			
F	77.67	37.01[b]	99.36[b]
	10.71	14.27	12.22
M	81.58	47.65[b]	110.44[b]
	9.41	16.02	12.65
T	79.76[c]	42.76[c]	105.57[c]
	10.16	16.03	13.55
Seniors (60–80 yr)			
F	71.83[b]	48.28	85.37[b]
	10.04	10.53	9.65
M	76.64[b]	49.64	96.71[b]
	9.01	11.01	9.93
T	73.55[c]	48.91[c]	89.52[c]
	9.93	10.65	11.14

[a] Mean and 1 SD.

The energy required to travel at these rates again varies by the categories described above. Self-powered human locomotion by definition draws energy from internal chemical sources. As described previously, people sought to reduce this energy usage by either leveraging power sources outside of their own body or increasing the efficiency of energy expenditure. External sources of energy included animals, combustible chemical fuel such as wood or hydrocarbons, or wind. Methods to increase the efficiency of energy expenditure included technology like wheels or gearing, friction reduction technology such as boats or skis, or energy dissipation mitigation technologies like rubber tires or engine cooling mechanisms.

The distinction between gasoline powered scooters and early gasoline-powered cars lies primarily in the physics endemic to each vehicle type. Both are powered by internal combustion engines, but the power needed to propel the significantly larger and heavier car is proportionally greater than the requirements of the scooter. This isn't initially a revealing comparison, as scooters and automobiles have different passenger capacities. To understand the energy implications of this form factor divergence, we must examine energy unit per passenger mile. By this measure, it is clear that a scooter is more efficient, and therefore the theoretically stronger option when comparing modes by time and resources required. So why did people across the country opt for cars?

The answer to this question is complex, and not all of it is explained by raw consumer preference. However, there were a number of features of automobiles that positioned them both then and now as attractive to consumers. Additionally, many of the drawbacks of this new mode

were not apparent at the time. This is especially true with regard to the automotive sector's role in global climate change.

Bunten and Rolheiser (Bunten & Rolheiser, 2019) have classified various transportation modes along three axes: Capacity, Range, and Route Flexibility. For example, a train would score highly on Capacity and Range, but very poorly on Flexibility. A bicycle would score low on Capacity and slightly better on Range, but reasonably well on Flexibility. In this frame, the advantages of a car become more apparent. Even early automobiles provided relatively high scores on all three, whereas scooters traded Capacity and some Range for a small boost in Flexibility. Scooters were certainly more appropriate for some applications, but the advantage further waned when the costs were taken into account. According to reporting by the Smithsonian, at the debut of the Autoped, it was priced at $100 compared to $400 for a Model T. Clearly, the car was still far more expensive, and the gap grows even larger when fuel costs are considered, but they remain within the same order of magnitude.

Even as the car became dominant in the U.S., the motorized scooter did not entirely disappear. It gradually changed shape, first with the addition of a seat and gradually with additional bodywork, to become the Vespa-style conveyance that was most strongly associated with the word "scooter" until fairly recently (Birtchnell, Harada, & Waitt, 2018). These have been most popular outside of the U.S., particularly in cities (Bishop, Doucette, Robinson, Mills, & McCulloch, 2011). This descendant of the scooter is prized for its versatility, small form factor, relatively low costs, and simple maintenance needs.

The other major contributing factor in the accelerating dominance of the private car in the U.S. context was the growth of the "concrete commons" (J. Coughlin, 1996). This concept groups a loose assemblage of parties who stood to gain from a expansion of concrete, including car companies, quarries and mines that supplied raw materials, fuel producers, and an enormous web of interconnected commercial entities. The Autoped scooter company would rightly have been grouped here as well, but its interests were not central to the orientation of the concrete commons because scooters simply required the production, and therefore sale, of fewer raw materials. It is important to note that while there have long been trade groups in these industries and communication about priorities and preferred policies, the concrete commons does not refer to a specific, conscious organization. Rather, it is a diverse group with aligned interests, who stand to gain through concrete and car-oriented policies.

Intensive lobbying by the concrete commons was pervasive and has been well documented. The lobbying efforts were directed not just at governmental decision makers, but at the wider population, selling the idea of a freedom-suffused lifestyle mediated entirely by

automobile travel. The notorious "World of Tomorrow" pavilion at the 1939 World's Fair was a particularly clear example of how groups that stood to gain from the prevalence of the car had a major role in shaping the public vision of the future. As more communities began to embrace a car-oriented future, the ability of residents to opt out of this lifestyle gradually decreased as businesses and services became more spatially diffuse. The path dependency of opting for a car in one context continued to create the appearance that the private car was the best option for each consecutive decision about who to allocate funding and public space (J. F. Coughlin, 1997).

As with any historical analysis in the U.S., there is not a single narrative that encapsulates the experience of people of all classes, races, and genders. Private car ownership, especially in the early and middle of the 20th century, was predicated on the possession of some measure of wealth. Widespread employment discrimination meant that the financial means to buy a car were far rarer among populations of color. Home mortgages became a major way to build wealth, but nearly all districts with major minority populations were "redlined" by the Home Owners Loan Corporation (HOLC). HOLC was a governmental agency responsible for providing information to banks regarding the expected viability of mortgages and investment, but it made these determinations on boldly discriminatory criteria. Populations of color could not receive a mortgage, banks ignored these neighborhoods for investment, and cities used these determinations to restrict public spending as well. As the residents of these redlined zones were intentionally shut out of the capitalist wealth generation process, cities and states tore these neighborhoods down to build highways, justifying this choice on the often dismal quality of life that was a direct result of racist government action.

The new suburban developments outside cities, where this car-oriented vision truly dominated, were usually built under racially restrictive housing covenants, which prevented houses from being sold to people of color. Despite invalidation of these covenants by the Supreme Court in 1948 (Shelley v. Kramer), the presence of these clauses in titles and deeds continues to this day. While they cannot be enforced, their continued prevalence is indicative of the *de facto* embrace of this practice by white homeowners, banks, and realtors in the second half of the 20th century.

The division of society along racial and wealth lines had tremendous ramifications for the ways in which people were able to move through the world. With cities and states prioritizing private car owners, public transit services in most cities saw funding decline. Many streetcar lines were removed entirely, often at the behest of car companies. The road space, which was initially semi-privatized but often non-exclusive, was allocated for use by private

automobiles and the occasional public bus. The barrier to entry for this nominally public space was significantly higher than it had been for streetcar service.

The removal or contraction of public transit services such as bus lines had cascading effects on those who had depended on these service. Moving within a city freely was no longer as feasible or predictable. As jobs, especially higher paying office and white collar professions, moved to the suburbs (following the bulk of their workforce), the ability of those who were excluded from the suburbs to reach these jobs was minimized. Even if a bus line technically ran to a suburban office park, the frequency and reliability of that bus might make reliance on this mode precarious. Underinvestment in bus service resulted in a lower quality of service, which diminished ridership, which in turn led to continued reduction of budgets for public transportation. This downward spiral was disastrous for communities of color that had been intentionally marginalized through earlier government action and widespread discrimination.

While owning a car became significantly cheaper during this period (1960s – 1990s), transportation barriers remained even for people in cities who could afford one. Credit practices continue to discriminate based on factors explicitly linked to the acknowledged history of racist economic and land use policy, including renter status and occupation. Overt racism has also been shown to play a large factor in determining creditworthiness or hiring, with similarly positioned applicants of different races receiving divergent results (Bertrand & Mullainathan, 2003).

Overall, we see a grim picture for the prospect of mobility equity and justice in cities across the country. The trends and forces at play were relatively stagnant in the transportation sector for much of the latter half of the 20th century, but the past 15 years have brought us to an inflection point in the transportation paradigm. Technologically, this shift has been made possible primarily by the spread of networked mobile technology, specifically small but powerful batteries, cellular data networks, and geographic positioning systems. Socially, this shift has been accelerated by factors including a renewed interest in urban life and policy by affluent former suburban residents, revived urban business districts, and the rise of the green commons as a political force, at least on the urban level.

Mobility Revolution

In early 2018, the first electric scooter sharing service in the United States placed dozens of devices on the streets of Santa Monica, California. Shaped much like a traditional kick powered scooter that had been used primarily as toys for centuries, these scooters contained GPS hardware, some basic data connection capabilities, and a small, battery powered electric

motor. This new form of mobility solution quickly spread to cities across the country and world, with over thirty million rides taken in the US in the first year alone.

These scooters arrived in a mobility landscape that has undergone seismic shifts in the past decade, after a long period of technological and political stagnation. When first proposed, car-sharing was a revolutionary idea. People could rent cars for short, flexible periods of time from distributed rental locations that have no physical infrastructure other than a sign. Individuals and households who used their car only occasionally could now explore the possibility of moving away from owning their own car, or could consider shifting from two cars to a single personal vehicle, with the convenience of on-demand car usage filling the gap. Zipcar faced a number of regulatory obstacles, as cities and states tried to figure out where existing laws applied to this slightly, but not radically, new form of transaction. Questions of insurance and liability, inspections, and zoning were all fiercely debated. A particularly contentious issues was if and how this service differed from traditional rental car agencies. Zipcar and other similar services spread gradually, but car sharing services can now be found in nearly every major US city and are common in other parts of the world as well.

About a decade later, Transportation Network Companies (TNCs) made an appearance. These companies provided an app that could connect people who wanted to get to a destination with a driver who would use their personal car to drive them, and the app would facilitate communication and payment. For riders, these apps provided a relatively frictionless way to summon a ride to a destination on their phone, and for drivers, these apps offered a flexible way to turn an asset (a personal car) into a remunerative device on their own schedule. TNCs were able to demonstrate that a digitally controlled model of fleet mobility could use strategic trip sharing and routing to reduce the number of car trips that would be needed to get everyone to their desired destination. Ride hailing was seen as a huge turning point for urban and suburban mobility (Clewlow & Mishra, 2017; Conway, Salon, & King, 2018).

However, this new mode was not available to the population on an even basis. The technologies described previously as prerequisites for access, namely smartphones and mobile data, are disproportionally absent among low-income communities. As a result, these new modes that might dramatically increase these residents' ability to move are often out of reach. This pattern risks replicating and further engraining existing transportation disparities (Groth, 2019). This paper provides an important critical perspective on the ways in which this mobility revolution are a replication of existing social structures and a reification of the politics endemic to technologically-mediated transportation access.

It has become apparent that many of the discrimination issues that were evident in taxi fleets persist in ride hailing (Moody, Middleton, & Zhao, 2019). Discrimination in ride hailing means that this new frontier of mobility is not available evenly (Ge, Knittel, MacKenzie, & Zoepf, 2016), and the boundaries of this access merely follow the wealth and racial boundaries endemic to the neoliberal capitalist system.

While TNC technology has added a great deal of mobility flexibility for some people, it has actually served to increase overall vehicle miles traveled (Chen, Mislove, & Wilson, 2015). The "independent contractor" model of employment that typifies the gig economy, of which Lyft and Uber are prime players, has resulted in a huge workforce outside of the traditional employment protections like health insurance that were the outcome of over a century of determined advocacy (Hall & Krueger, 2016). The local and global externalities of ride hailing and ride sharing are still the subject of intense scrutiny from policymakers, academics, and corporate interests (Shaheen, 2018), but for all the fanfare of this "new mode", people are still traveling inside a traditional, internal combustion (or sometimes hybrid) car (Salvucci, 2019).

In the past three decades, cycling has seen a resurgence in US cities as a standard mode of transportation (Dill & McNeil, 2013). This trend has accelerated in recent years as cities build larger and more robust infrastructure networks that address the needs of these significantly more vulnerable road users, as compared with automobile drivers and passengers. Cities have begun to acknowledge studies indicating that people who bike to a commercial district tend to visit more frequently and spend more per month, and have found strong economic and environmental rationales for incorporating bike planning into infrastructure projects of all sizes. Protected bicycles facilities are now in place or under construction even in traditionally car-oriented US metropolitan areas (NACTO, 2019). As demonstrated most effectively by the city of Seville, Spain (Marqués & Hernández-Herrador, 2017), isolated fragments of bicycle infrastructure have far less impact than fully-connected networks. Even a single discontinuous link in a safe network can discourage hesitant riders from making a trip on a bicycle (Lowry, Furth, & Hadden-Loh, 2016; Marqués, Hernández-Herrador, Calvo-Salazar, & García-Cebrián, 2015).

The bicycle infrastructure investment boom has happened in tandem with a surge in the visibility of the bicycle lobby in local politics (Hoffmann, 2016). However, despite the wide utility of these safe streetscapes, the most visible members of these activist groups are often reflective of socio-economic and racial divides that already exist for the distribution of political power. The whiteness of the bicycle lobby has strengthened the associations of bicycle

infrastructure construction with gentrification (Larsen, Student, El-Geneidy, & Yasmin, n.d.), a topic explored further below.

The spread of bike share systems has facilitated some of this growth in bicycle traffic, as these systems remove the barriers or ownership, maintenance, and parking that discourage many people from using bikes (Fuller et al., 2011). Dockless bike share systems and companies have expanded the reach of bike share while reducing the planning footprint, and in parts of the world, dockless bike share accounts for a huge number of trips per day (Ursaki & Aultman-Hall, 2016).

Pedal assist electric bikes, also known as e-bikes, have also arrived on the market with widespread availability in recent years (Bishop et al., 2011). These bicycles have a small electric motor that reduces the effort required to move oneself forward, but still requires some pedaling motion in addition to the effort required to steer and balance. By making it easy to maintain a high speed without high levels of exertion, these e-assist bikes have opened new possibilities for using bicycles to commute longer distances (Dill & Rose, 2012; Fishman & Cherry, 2016; Fyhri & Fearnley, 2015; Kroesen, 2017; Popovich et al., 2014).

Considerations of equity and access have been mandated in transportation planning decisions since the mid 1990s, but huge disparities in access to opportunity and public investment persist. Most public transportation systems have not been fundamentally reworked since the middle of the last century, and low-income communities and communities of color are disproportionately locked in transit deserts, with infrequent or unreliable service (Handy & Clifton, n.d.; Lindsey, Maraj, & Kuan, 2001). Recent waves of active transportation investment in bicycle infrastructure or improved sidewalks have often passed over low-income communities (Braun, 2019), due to a variety of factors including structural racism that rewards types of public participation that are far easier for those who already possess significant societal clout (Ursaki & Aultman-Hall, 2016). Despite transportation investments of all kinds exhibiting high levels of induced demand, arguments about these investments have often rested on fallacious and racialized claims of potential ridership (Hankey et al., 2012). While one should debate the possible causes of imbalanced demographics in bicycle use, statistics from cities across the country show that people who ride bicycles for commuting are disproportionately white, male, and medium to high income (Hoffmann, 2016). Within this reality, there is an argument that continued investment in bicycling infrastructure as opposed to strengthened public transit is one more type of inequitable transportation plan (Garrard, Rose, & Lo, 2008).

Part of this transportation milieu is that public transit infrastructure across the United States is aging and nearing the end of its useful life. Decades of underinvestment in maintenance

is the norm rather than the exception, and transit agencies across the US are scrambling to address major deficiencies as they begin to impact service (Zureiqat, 2010). Younger generations are also moving more to cities and are buying fewer personal cars (Hamre & Buehler, 2014). This phenomenon has various explanatory factors, including lower relative generational disposable income, awareness of critical environmental harms, or personal preferences (C. Jones & Kammen, 2014). However, it is clear that the densest parts of the US are getting even denser, and the geometry and logistics of traditional automobiles have made this mode of travel impractical or uneconomical for many (Dwyer & Hu, n.d.; Moya-Gómez & García-Palomares, 2016).

Finally, the shifts in mobility have already caused significant upheaval in planning departments across the country, as agencies try to adapt their traditional organizational infrastructure to new realities. Cities often struggle to meaningfully incorporate new types of data to improve predictive planning, especially in transportation planning. Traditional, federally-mandated demand models are still used to underpin analyses for all types of infrastructure, but because cities cannot parse meaning in the reams of GPS traces and social organization data thrown off by their residents, the model scenarios and human realities continue to diverge (Favaretto, De Clercq, & Elger, 2019).

The environmental implications and imperatives of this mobility shift were described in the *Three Revolutions in Transportation* report (Shaheen, 2018). The three revolutions refer to three factors which, taken either together or independently, have the potential to radically reshape transportation: Electrification, Automation, and Sharing. While Automation appears increasingly far away and dangerous following a brief period of optimism (Litman, 2019), we have continued to make headway on Electrification and Sharing.

Scooters

Into this tumult, scooters have spread with unprecedented speed. E-scooters rely on a host of relatively mature technologies combined in an innovative way, including electric batteries and motors, handheld smartphone GPS, and QR readers. These scooters are emerging at a time when people have seen the rise and spread of bike lanes, and the idea of going on the road outside of a car may not be quite as alien as it might have been a decade ago. The self-locking character of the scooters means that users can consider trying them a single time for a journey without worrying about having to park it or prepare for the trip. The ubiquity of ride-hailing apps means that even people who would not consider public transit could potentially use these scooters and not commit to traveling in this mode longer than they want. People have

cited the active and environmentally sustainable character of the transportation mode as a possible reason to try scooters. In the popular press, many others cite convenience or cost as the factors that convinced them to try a scooter (Arellano, Eng, & Fang, 2019).

While there has been a great deal of excitement about this technology and the potential for improved mobility, there has also been a significant amount of backlash. In some cities, companies have competed on "coverage" by attempting to blanket the landscape with their own brand of scooters, with the goal that their network is the most convenient. However, the result of this has often been an unsolicited blanket of scooters left in the public right of way. Because they do not need to be locked to racks like bicycles, riders are free to park them nearly anywhere with little accountability. Some companies have deposited scooters into cities without alerting city officials, or even against the express wishes of city governments. Community groups have often decried the scooters as a danger to pedestrians and to drivers, claiming that inexperienced riders move erratically in ways likely to cause a crash. Relative to previous years, emergency rooms have seen increases in the number of scooter related injuries. However, these popular narratives omit important context.

With regard to the number of scooter injuries, mediating the raw count of injuries by the miles travelled on this mode would do a better job of placing the potential danger of scooters in the broader mobility context. In the United States, most are inured to the constant mortal toll exacted by automobiles, both through direct crash injury and from tailpipe emissions. In light of this, one should place the relative dangers of scooters in context with the relative dangers of automobiles. The safety questions around this new mode are beyond the scope of this research, but bear further examination provided with the proper framing.

A typical scooter ride on a scooter sharing service might have the following chain of events. First, a potential rider decides that they could travel to their destination using a scooter. This could be prompted by seeing a scooter on the street and downloading the app, or they might already have the app, and open it to locate a scooter near them. Once the user reaches the scooter with the app open, they scan the QR code with their app to unlock the scooter. To ride the scooter, users turn a throttle on one of the handlebars while balancing on the baseboard. The scooters can accelerate up to 15 miles per hour, and are slowed by a lever brake on the handlebars. Riders are instructed to ride in the bike lane, to avoid the sidewalk, and to ride safely. When they reach their destination, riders move the scooter onto the furniture zone of the sidewalk, and tap End Ride on their phone to relock the scooter. Some companies require that riders send a picture of the scooter locked to validate that it is parked correctly. Pricing is usually $1.00 to unlock the scooter, then $0.15 per minute of use. Most providers offer a form of a low-

income program for customers who already receive federal or state benefits. In recent weeks, some major companies have begun to vary this pricing model by city or time of day.

Figure 1.2 – Shared Electric Scooters, Photos by Author. Brookline, MA. Clockwise from left: A scooter on a street corner; printed guidance on the scooter for riders; a view of the handlebars, showing accelerator tab and brake lever; footboard showing pricing information and kickstand; detail of rear tire.

Chapter 2 – Literature Review

Literature Review

There are a few core concepts and aspects of literature that will guide my research. As discussed in the previous chapter, critical urban theory is the motivating lens through which I am pushed to interrogate the implications of new mobility technologies. The second is the concept of accessibility in transportation. Laid out convincingly by Sheller and Urry in 2006 (Sheller & Urry, 2006), this posits that transportation is essential to the modern conception of self. The ability to move through both physical space and cultural space is at the core of full participation in society. This concept has been recognized and enshrined in the United Nation's Sustainable Development Goals (Weiss et al., 2018). The ability to control one's own movement provides the personal power and self-determination needed to seek preferred employment, engage with family, find entertainment, and find social connection (Boarnet, Giuliano, Hou, & Shin, 2017). This theoretical framing on my research also implicates the degree to which inequitable mobility power can shape and cramp the lives of marginalized individuals and communities (Grengs, Levine, Qing Shen, & Qingyun Shen, 2010).

The third key concept is equity. Equity is the idea that rather than providing people of all backgrounds and experiences the same supports towards some desired end, policy and society should ideally provide differentiated supports that take into account historic and present conditions of marginalization and power (Lee, Sener, & Jones, 2017).

Another important concept is the idea of transportation mode split and mode shift. This is a conceptualization of the percentage of people who use different methods of transportation, such as a bicycle or a private automobile, to move about during the course of their day.

My analysis builds on the work of other scholars. In the traditional framing, exemplified by the work of Ben-Akiva (Ben-Akiva, Lerman, & Lerman, 1985), transportation choices are made in three steps. First, when considering making a trip for the first time, most people select a destination. Once the destination has been identified, people consider what travel modality they would like to use to there. Usually, the options are private car, public transportation, or sometimes walking or biking. Finally, people decide on the route they will take to get there in their chosen travel mode(Miskeen, Alhodairi, & Rahmat, 2013). Most travel demand "four step" models use a version of this decision tree, while adding a spatial dimension to simulate trip

decisions from a simulated population dispersed with origins and destinations across a region (Ermagun, Lindsey, & Hadden Loh, 2018). These models are used in every metropolitan area in the United States. However, many of them are relatively antiquated and do not consider the rapidly shifting transportation tools that populations now have at their disposal. These new tools include trip planning applications on phones, ride hailing, and bike share, among others. A number of papers have recently explored how to consider ride hailing in this traditional model by calculating relative utilities of travel time for this new mode (Ferrara, Liberto, Nigro, Trojani, & Valenti, 2019). My work will build on this trend and expand the scope of consideration to include scooters, laying out a framework for future mode shift research. I will be drawing from a number of methodologically related works examining the rise and potential for e-bikes, including (Fyhri & Fearnley, 2015), (Fyhri, Heinen, Fearnley, & Sundfør, 2017), (Cairns, Behrendt, Raffo, Beaumont, & Kiefer, 2017), and others. I propose using a choice based sample methodology to assess revealed preferences, despite possessing only the positive choice components of the experiential universe (Imbens, 1992). Wilson et al. (2015) provide a useful methodology to estimate levels of TNC activity within this proposed framework.

One guiding piece of research on the topic of scooters is the report titled *The micro-mobility revolution: The introduction and adoption of electric scooters in the United States*. (Clewlow, Regina R, 2018) This is a whitepaper based primarily on a survey conducted over 7000 individuals in seven major US cities. This data shows high levels of interest in this emerging technology, and importantly, shows parity in this level of interest across genders and income levels. This pattern is not replicated when examining cycling, as seen in Garrard (Garrard et al., 2008). A body of research has emerged showing that differences in the use of micromobility depend on a wide set of factors, the importance of which can vary between genders (Krull, 2018).

Understanding who is captured in biking data is both critical and underexamined (Hoffmann, 2016). There is a growing body of work addressing this issue, particularly on publicly owned or co-operated bike share systems. (Ma, Ji, & Yuan, 2019) focus their analysis on understand who might be using bike share and why. They use demographic factors, including population counts, income level, education level, and car ownership. This is supplemented with land use data, including public transit stop density, distance to the central business district, and Points of Interest. Their analysis took the form of a Geographic and Temporally Weighted Regression, which is an extension of an Ordinary Least Squares regression that accounts for spatial relationships. This is a particularly relevant model, as it allows each coefficient to vary

over time. The spatial distribution of services is a major component of transportation equity, and will be central to my analysis.

One emerging mode that bears a strong resemblance to shared electric scooters are dockless bicycles. While they have now been primarily superseded by electric-assist bikes (NACTO, 2019) , dockless bicycles spread rapidly during 2016 and 2017 (Rahim Taleqani, Hough, & Nygard, 2019). They provide a good base for examining past methodologies and findings. Zamir et. Al (CITE) used logistic regression with a random forest model to split the user base of docked bike share in Washington D.C. between "member" and "casual" groups. They discovered that users of dockless bike share seemed to follow use patterns more similar to that of a "casual" user rather than a "member" user. This finding has implications for how people may encounter and use dockless micro-mobility.

Research by Pu et al. (2019) uses a semi-parametric geographically weighted regression to understand what factors might influence ridership of bike share systems. They divide their analysis into trip origins, trip destinations, and a subset analysis of users who purchase a full day pass. They find that the geographically weighted regression performs better than a simple Ordinary Least Squares model, though this result is to be expected based on understandings of spatial autocorrelation. Xu et al. (2019) used a geographically weighted Poisson regression to understand what factors might associate with trip generation. Their findings indicate that points of interest and mode split factors, including public transit use, contribute strongly to the model fit.

A recent study (Qian & Jaller, 2019) examines bikeshare in Chicago with a focus on use patterns in disadvantaged communities. They use income levels and populations of color as the signifiers for a disadvantaged community, and find significant differences between these two groups of geographies through a negative binomial regression model. Specifically, they find a significant reduction in the number of trips taken from stations in disadvantaged communities. Furthermore, they find that annual members who start trips in from stations in disadvantaged communities make longer and more expensive trips than the city-wide average.

Wang et al. (J. Wang & Lindsey, Greg, 2019) develop a broader index of neighborhood disadvantage and use k-means clustering to group the trips taken by residents of each neighborhood type. They find that members who live in disadvantaged communities use bike share more frequently, have origin / destination patterns that are more widely dispersed than average, and take longer trips. The methodology leverages mixed-effect regression analysis.

In modeling travel behavior, there has been a surge of interest in using Open Street Map data (Wasserman, Rixey, Zhou, Levitt, & Benjamin, 2019). Wasserman et al. estimate the Level

of Traffic Stress ("Low-Stress Bicycling and Network Connectivity," 2017) using Open Street Map data. They find a high degree of similarity between the digital model and a similar assessment conducted by hand.

Research Question and Hypothesis

What factors serve as effective predictors of the characteristics of shared electric scooter trips in three U.S. cities? Both human factors and environmental variables are explored in a series of spatial lag regression models. We have framed the null hypothesis as a situation in which population density is the only factor influencing the number of trips that occur, with no variations in trip characteristics such as average trip distance across an urban area. We expect to reject this null hypothesis.

Trip characteristics under examination are the total number of trip origins and mean trip distance for trips that originate in a zone.

The general hypothesis is that concentrations of employment are a strong predictor of electric scooter trips, but trips also begin and end outside of central business districts. I further hypothesize that measures of societal marginalization will serve as strong predictors of where scooter trips are less likely to occur.

I hypothesize population density will have a significant, positive effect on the total number of trip origins and destinations. I do not hypothesize an effect of population density on trip distance.

I hypothesize that job density will have a significant, positive effect on the total number of trip origins and destinations. If shown, this will provide support to the claim that job centers are hubs of transportation activity, including electric shared scooter activity. I further hypothesize that the magnitude of this effect will increase during the PM peak block.

I hypothesize that the Communities of Concern Index will have a significant, negative effect on the total number of trip origins and destinations. If shown, this would support the

claim that electric scooters, as currently deployed, may be replicating some features of existing transportation inequities. If demonstrated in the Spatial Lag models, this result does not claim that electric scooters are inherently inequitable, and supports claims only about historical deployment and use captured in the analysis data.

I hypothesize that the distance from the central business district will have a significant, negative effect on the number of trips that begin in a zone, but it will have a strong positive effect on the distance of trips that begin there. This finding would support the claim that rides are more likely to be generated in areas close to the densest set of amenities, and that people may use scooters to reach these areas. Additionally, this variable will serve a normalizing effect on other predictive independent variables.

I hypothesize that the density of street network in a zone will have a positive predictive effect on the number of trips that begin in a zone. If found, this will support the claim that potential places to ride are a partial determinant of where rides may occur.

I hypothesize that the median household income will have a weak positive effect on the number of rides that occur, with no impact on the length of those rides. If found, this result would support the claim that scooters may be more likely to be used by those with higher levels of disposable income.

I hypothesize that the number of rebalance points will have a highly significant impact on the number of trips that occur. The number of times that a scooter has been placed in a zone through non-transportation activity is an important normalizing variable. If a strong predictive effect is found, this would support the claim that decisions about how to deploy the scooter network have a significant effect on where rides occur. The methodology for deriving this input is covered in Chapter 3.

Beyond the spatial lag model, I hypothesize that weather, both temperature and precipitation will be strongly correlated with scooter trip starts, with temperature showing a positive relationship and precipitation showing a negative relationship. If found, this evidence would support the claim that the use of scooters is highly dependent on changeable and uncontrollable environmental factors.

We predict time of day to be a crucial determinant of how these factors might interact, but within the scope of this analysis, our use of time as a predictor variable is limited. We hypothesize that in the three analysis cities, the number of trips will be highest in the evening.

We expect to see small peaks at traditional commute hours, but these will not be at the same magnitude as traditional transportation usage peak hours.

We hypothesize that revealed travel behavior will be generally longer than idealized bicycle routing. We generally expect most trips via scooter to have a faster routing via both car and bicycle. Furthermore, we expect a small portion of trips taken on scooter to be completed faster than the O/D pair could be linked by transit at the appropriate time of day.

Chapter 3 – Data

Mobility Data Specification

The data used in this analysis is comprised of the complete record of one company's scooter activity in three focus cities: Nashville, Tennessee; San Diego, California; and Portland, Oregon. I intend to refer to the company as "Gust", for General Unaffiliated Scooter Trips. The data is structed in the Mobility Data Specification schema, developed by the City of Los Angeles to enable public sharing of new mobility data in a form that allowed planning and regulatory access while protecting the privacy of individuals and sensitive business information of companies.

CITY	TOTAL TRIPS (AFTER OUTLIER TREATMENT)	DATES
NASHVILLE	419,667	May 25, 2018 – Feb 28, 2019
SAN DIEGO	1,982,757	September 1, 2018 – Feb 28, 2019
PORTLAND	226,912	July 22, 2018 – Nov 20, 2018

Table 3.1 – Total Number of Trips by City, Date of Analysis

For the purposes of this analysis, data was cut at the end of February, 2018. Future research should test the estimated models on new data from March and beyond.

The Mobility Data Specification is split into two halves, the Provider schema and the Agency schema. The Agency schema is designed for public agencies to share information about current conditions, services, and policies with mobility providers, and to track compliance with regulations and permits in real-time. This includes detailed vehicle permitting and registration, which providers are able to do entirely through the API. Within the Agency schema, there is a large focus on tracking vehicle events, or status changes. These include significantly more detailed parameters about customer activity than are present in the Provider schema, including separate flags for a customer reserving a vehicle and starting a trip, as well as canceling a trip. This also provides documentation for when a city agency picks up a vehicle, or when a provider

picks up a vehicle, and specifies the reasons for this activity, including rebalancing, maintenance, charging, or compliance.

The Agency schema formerly delivered more precise telemetry data about each vehicle, including position, speed (if available), and percent charged. However, these telemetry points are not grouped by trips, and are returned as points from an Agency API request. The Agency API also tracks the current valid service area as a detailed polygon object, giving the city the ability to track compliance spatial regulations. The Service Areas request allows cities to monitor use in different areas of a city that might require special attention or different policies, like a marginalized community or a central business district. This is a critical ability for cities to manage mobility vehicles active in their city, as outlined in Chapter 3 (Mobility).

Generally, the Agency schema is designed to manage regulatory compliance and deployment, around topics such as vehicle registration, territory, speed, deployments, and general presence of vehicles in the public right of way.

The Provider schema is intended for providers of mobility services to share information about their operations in a detailed and anonymous way, and contains two endpoints. Each trip furnished by a provider creates a record in the Trips endpoint, which can be tracked and stored by relevant public agencies for use in planning, compliance, and research. Each state change of the device creates a record in the Status Changes endpoint. These state changes include beginning to charge, start of a ride, removal for maintenance, and others detailed below.

A trip record contains the following information: the provider ID, the provider name in cleartext, the device ID (within the data system), the vehicle ID (issued by the company and visible on the vehicle), a vehicle type (scooter, bike), a propulsion type (electric, human, gasoline, hybrid), a unique trip ID, a start time, and end time, coordinates reported to the finest level of detail available saved as points (formatted as a json file), the approximate level of GPS accuracy for reported locations, the duration of the trip in seconds, and the approximate distance in meters. There are additionally three columns that are optional under the MDS provider schema, but are included in the Gust dataset. These are the standard cost (what would be paid under the generic pricing scheme), the actual cost (what the user paid), and a user-submitted photo (accessed through URL), ostensibly taken for parking verification.

Field	Type	Required/Optional	Comments
provider_id	UUID	Required	A UUID for the Provider
provider_name	String	Required	The public-facing name of the Provider
device_id	UUID	Required	A unique device ID in UUID format
vehicle_id	String	Required	The Vehicle Identification Number visible on the vehicle itself
vehicle_type	Enum	Required	See vehicle types table
propulsion_type	Enum[]	Required	Array of propulsion types; allows multiple values
trip_id	UUID	Required	A unique ID for each trip
trip_duration	Integer	Required	Time, in Seconds
trip_distance	Integer	Required	Trip Distance, in Meters
route	GeoJSON FeatureCollection	Required	See Routes detail below
accuracy	Integer	Required	The approximate level of accuracy, in meters, of Points within route
start_time	timestamp	Required	
end_time	timestamp	Required	
parking_verification_url	String	Optional	A URL to a photo (or other evidence) of proper vehicle parking
standard_cost	Integer	Optional	The cost, in cents, that it would cost to perform that trip in the standard operation of the System
actual_cost	Integer	Optional	The actual cost, in cents, paid by the customer of the *mobility as a service* provider

*Table 3.2 – Mobility Data Specification Schema (*A data standard for Mobility as a Service Providers who work within in the City of Los Angeles*, 2018/2019)*

The provider ID and provider name are straightforward, and we have only data from Gust, and cannot make meaningful comparisons between MDS data from other providers.

For purpose of analysis, the device ID is a better metric than a vehicle ID. A company could theoretically transfer the customer-facing ID number to a new vehicle, or could have duplicates of the same vehicle ID spread across different cities. The vehicle ID is usually a short string of letters and numbers that a user can enter on their phone to unlock the device if the QR interface fails to work. It is also a way for customers to reference the vehicle in reporting problems or concerns, and must be human-legible. The device ID is a Universally Unique Identifier (UUID) code, meaning that there are, with near certainty, no duplicates of this code

not just within MDS, but within any system. This makes it possible to track the digital components of the scooters over their lifetime. However, it is theoretically possible that during maintenance, the scooter brain could be moved into a new or different physical body, which could be a relevant consideration during an analysis of vehicle durability.

The vehicle type and propulsion type work together to identify the characteristics of the vehicle. The vehicle type has only two possibilities, scooter or bike. As the "bike" value would encompass both traditional bicycles and electric-assist bicycles, the propulsion type column must be used in conjunction to get a better idea of the specific vehicle. It is unclear how more novel vehicle typologies fit into this form, such as a bike form factor without pedals, entirely electrically powered. Providers have indicated that they are actively exploring other form factors, but at this time Gust deploys only electric powered scooters.

The trip ID is also a UUID, and will be carried into all data tables tied to each trip record. This is the primary key for all analyses, marking each trip as unique and serving as the join factor in the data cleaning and manipulation.

The start and end times are saved to thousandths of a second, and include a time zone. These times refer to the start of a trip, rather than the beginning of a reservation. Gust treats these events as the same, and does not allow advance reservation of a device.

The route column is stored as a JSONB Feature Collection object in a single column of a trip record. The Feature Collection contains a series of points, identified through latitude and longitude, each with an associated timestamp. The MDS protocol requires devices to record their location to this record as frequently as the company receives it, so this is the most accurate record of vehicle location. An initial analysis of the available Gust records shows one point recorded every 30 seconds from the beginning of the ride, with a final location reported at ride end. It is possible that this rate increases with the introduction of newer scooter models, but the baseline appears to be every 30 seconds. From these point collections, I will calculate a linestring geometry roughly corresponding with the route.

All reported points are somewhat limited by GPS accuracy, recorded in the next column. An initial analysis shows a common accuracy figure of 152 meters, indicating that the reported point lies in an area of uncertainty with a radius of 152 meters. This is a sizeable area of uncertainty, but there is reason to believe that points may generally be more accurate than this figure. The MDS indicates that the Accuracy column should be the "Approximate accuracy, in meters, of points within Route." Each route contains at a minimum two points, but generally a number more on the order of 200 unique points. However, accuracy is a single, composite

statistic. It appears to be a reasonable assumption that this statistic represents a minimum value, rather than a calculated average. This has been confirmed by a data engineer at Gust.

The trip duration is a number of whole seconds, calculated by subtracting the timestamp for the first reported location from the timestamp for the final reported location.

The trip distance is reported as an integer number of meters. This is created by calculating the linestring distance from each consecutive point on the route, as reported by the GPS.

The standard cost column seems to be a simple calculation following the general pricing schema for Gust, which is $1 to start a ride, and $0.15 per minute of ride time. A subset analysis revealed that in all examined cases (approximately 20,000 trips, not evenly distributed by date), the standard cost column is equal to the actual cost column. This is disappointing, as there were a number of discount codes that Gust publicized to entice new users, which appear to not be reflected in the data. Additionally, there do not appear to be any trips taken by members of the equity programs, at least using this metric. The implications and current status of this metric are covered in the Chapter 6.

The final column for each trip record is a link to a parking verification photo. These links are present for about 65% of trip records, based on a subset analysis. A manual review of approximately 20 photos shows that they vary widely in quality and utility, and many are blurry pictures of the ground. However, there are also a number of photos of Gust scooters parked in various locations. I have not saved these photos, and do not intend to analyze them. The URL contains what appears to be a an additional randomly generated UUID, which does not visibly link it to either the trip or the user. Analysis of these pictures is outside the scope of this research.

Obtaining and Processing MDS Data

The data partner on this research, Gust, provided access to their complete MDS Provider databases for three cities. The process of downloading and transforming this data into a useable format was time intensive and non-linear. The implications of the hurdles encountered will be covered in the Discussion chapter, and this section will outline the process used for this report.

To facilitate rapid ingestion of MDS data, I adopted a Github project repository hosted by the City of Santa Monica. The code in this repository was primarily authored by Kegan Maher, an employee of the City. (*Services for working with MDS Provider data, built as runnable Docker containers.*, 2018/2019) A static fork of this repository at the time of this

research is hosted on the author's Github. This project wraps a number of Python scripts written to work with MDS data into virtual machines, runnable through the Docker environment. Docker is the core product of Docker Inc., and is a software environment that facilitates the use of small, easily configured virtual programming environments with the ability to execute code, interface with external objects, and hold significant amounts of cached data in an efficient fashion. In the Docker context, these virtual machines are referred to as "containers."

A "virtual machine" simulates some of the aspects of a normal, physical computer but exists entirely within a separate physical computer. They are often used in programming and research because they provide inexpensive flexibility to define environmental attributes of the computing environment. A virtual machine can be as complex as running an entire Windows environment, or as simple as a single Python kernel and parser. The resources available to virtual machines are limited by the computer they are housed within. The main advantages of using virtual machines as compared with running processes in the local environment are threefold: experimental code or configurations can be run without any potential spillover to the main operating system; the competition for processing resources within a virtual machine is more limited, so some types of code may run more efficiently; and virtual machine configurations can be shared and replicated on different host machines so that code executes within identical parameters no matter the host context.

The MDS-Provider-Services repository available from Santa Monica was crucial for this project, as initial attempts to run MDS code and schema parameters on my local machine encountered repeated and varied errors. The use of containerized code allowed me to sidestep these issues, but presented its own set of problems. The repository provided 6 containers to address potential or likely tasks when working with MDS. The *Server* container emulates a Postgres database configured to hold MDS schema tables. The *Client* container contains a current installation of PGAdmin, the preeminent PostgreSQL database management environment. The *Ingest* container packages a set of Python scripts and allows easy definition of their parameters, then these scripts are executed in sequence to extract MDS trips from the appropriate databases (in this case, Gust's database) and upload them to the *Server* container. The *Analytics* container contains some basic functions to determine availability of vehicles included in the *Server* container. The *DB* (database) container provides functions to update the version of existing MDS databases, modify attributes of MDS files such as trip geometries and timestamps, and define focus geographies. Finally, the *Fake* container provides the ability to generate randomized trips that conform to the MDS schema, for use in testing or comparative analysis.

The analysis in this paper makes use of the *Server* and *Ingest* containers. However, initial data pulls using the Docker environment resulted in data loaded in such a way that it was accessible only through the *Client* container. However, attempts to export or extract this data from the virtualized database to a local secured server were unsuccessful, as the virtual machine had no available interface or open ports with which to connect to the external file architecture.

Following this setback, discussions with my reader led me to focus on accessing the virtual server from the local file system. After a number of attempts at remapping ports, the *Server* container was configured to be visible by a local instance of PGadmin. This local instance was used to interface with the *Server* container, and the data loaded in this virtual machine.

The MDS Provider request headers are formatted to be maximally useful for a city with multiple active providers, and the *Ingest* container reflects this. It requests trips once from each provider identified with a unique "Provider ID" field in MDS. The research situation for this project required a different approach, as the Provider ID field was identical across all three cities. Accessing data from each city required a unique Docker instance, configured with the appropriate private API Token to be passed in request headers from the *Ingest* container. This posed no major issues, but necessitated a wipe and reload of the *Server* module before data from the next city could be downloaded. Due to programmatic limitations on the destination server and the processing capabilities of my local machine, the decision was made to dump the database and reupload to the server for SQL manipulation.

I attempted to download monthlong chunks of the Trips table as Comma Separated Value files, to reupload to the final secure local database. Initial tests of this technique were promising, with a 100 row subset successfully exported and reuploaded. I then exported over 120GB of data from the virtual *Client* in this fashion, with the goal of immediately uploading to the local secure server. However, the JSON column containing all of the geographic information for each trip posed a serious problem. The JSON schema mandates a different combination of quotation and escape characters than are allowed in the CSV format defined by PGadmin upon export. Despite a number of attempts to address and reformat these large CSV files using regular expressions, Python+Pandas, R, PostgreSQL, and even Microsoft Excel, all of the data obtained with this workflow (12 days of continuous download and export) was unrecoverable.

Following this significant setback, I adopted a strategy of manipulating the Trips table with PostgreSQL to add a geometry column. This would allow me to connect to the server through QGIS, and export and save the data as ESRI Shapefiles, a far more space-efficient and reliable format for geospatial data. The SQL queries to accomplish this task are described in the following section.

The configuration of the *Ingest* module was difficult due to the scale of data to be downloaded. The module was designed for tackling data periods covering a few hours or a few days, but downloading months of data was beyond the capacity of the virtual machine parameters. A very thoughtful component of the *Ingest* routine was the concept of backfill requests, in which the process requests data from each point in time twice, in partially overlapping time intervals. This was done to ensure that requests returned trips that started but did not end within a particular window. This technique makes the most sense when request tiles are relatively short, on the order of 30 minutes. However, the efficiency of the request process is halved under this system, and longer request tile definition can accomplish the same goal. I settled on sidestepping the including backfill function, and instead manually defining overlapping requests. The request windows were 6 to 8 hours long depending on the city and time of day, and overlap with each other by 1 hour on either side.

To manually define these request windows, I wrote a set of Python scripts with two major components. The original *Ingest* options allowed for manual definition of dates and times of interest, query rate limits, table type (Trips or Status Changes), and pre-upsert (transferring records to end of the existing database table) staging behavior, among other things. After experimenting with different techniques to define the request parameters, I programmed a series of mutating strings that contained these parameters spaced appropriately across a single day (and partially into the next one). Each string was then sent sequentially as a Bash command, meaning that the Bash command line interface would interpret and execute each string. The commands instructed the Bash terminal to initialize an instance of the *Ingest* container to pull trips from a single time window, validate the received data against the MDS schema for structural inconsistencies (and discard any non-passing records), and upsert these records to the MDS database in the *Server* container. Following the upsert, the container instance would end and wipe its cache. The Python script would then wait 5 seconds before sending the next terminal command. The second component took these sets of mutating strings and looped them over each day in a defined date range.

These scripts improve upon the performance of the base *Ingest* module in a few ways relevant for this project. In essence, the scripts take the parameters that would have been fed to a single *Ingest* container and instead divide them into small chunks, run sequentially by a series of hundreds of *Ingest* containers. Initializing and then destroying hundreds of containers proved to be more efficient and durable than demanding similar performance of a single *Ingest* container. By sharply limiting the time interval requested in each *Ingest* container, the temporary storage demands of the virtual machine were drastically reduced. Each instance

transferred validated data to the same *Server*, but transferring only one time window at a time resulted in better recovery from connection errors, as partial data would be safely inside the *Server*. The manual definition of each time window allowed variable parameters depending on the time of day the request covered. Rate limits were required during daytime and evening time blocks, but were extraneous for overnight requests. The use of request rate limits not only eases the burden on the Gust server, but paces the receipt of payloads in a way that provides more consistent and less processor-intensive processing. This flexibility was not possible in the original *Ingest* options. Finally, the lightweight nature of each instance meant that multiple scripts covering different date ranges could run in parallel, with up to seven containers requesting data and upserting to the same database.

The final modification to the Docker container model is the destination of the upsert actions. By modifying the port definition in the Docker environment and configuration files, the upsert actions were aimed at the local secure server, preconfigured with empty tables in MDS format.

American Community Survey

The United States is obligated by the Constitution to collect basic data on the population once every ten years, in a process known as the decennial census. The American Community Survey, known as the ACS, is a data collection program administered by the Census Bureau, a branch of the United States Department of Commerce. The program began in the early 2000s with the goal of addressing the ongoing needs of the federal government for reliable national data on a wide variety of characteristics. One of the main goals of the ACS was to supplement the data collected in the decennial census. While the decennial census collects a full population, the ACS using sampling methodology to create population estimates while only collecting data from a small subset of individuals. In this context, "population" refers to the statistical concept of a dataset which includes direct observations of the entire universe of interest. The three most salient differences between the decennial census and the American community survey are data collection methodology, topics covered, and data release schedule. These are summarized in the table below.

	Decennial Census	*American Community Survey*
Data Collection Methodology	Complete Population- Every person is directly observed, resulting in the definitive record of the universe of interest	Representative Sampling- Small number of observations in each geography, statistically scaled to create population estimates
Topics	Limited Scope- Baseline measures include number of individuals per residence, home ownership/rental status, ethnicity, age, and sex.	Wide Scope- Topics cover exhaustive range of information about people, including but not limited to income, vehicle ownership, previous residence, insurance coverage, language, and many other topics.
Data Schedule	Infrequent- Collected once on the first year of a decade (2010, for example), the preparation and the processing take multiple years on either side of the collection year.	Annual- Data is collected every year, and is two different products are released for each year, 5 Year and 1 Year Estimates. Detail is below.

Table 3.3 – Census and ACS Methodology

The ACS collects data every year, and uses the yearly data to publish two different categories of estimates. The 1 Year Estimates include only data for the sampled year. The total count of individuals sampled in the most recent year for which data has been released, 2017, was 5,014,176 (Table B00001, ACS 2017 1YR). This figure represents a portion of total individuals in the United States, which the ACS from the same year estimates to be 325,719,178 (Table B01003, ACS 2017 1YR). This indicates an approximate 1.5% sample from which 1 YR estimates are drawn. This can provide adequate statistical reliability for larger geographies, but when looking at smaller areas such as census tracts and block groups, the margin of error on these estimates can increase significantly. Furthermore, smaller geographies are more susceptible to the effect of participant non-response as individuals represent a much larger share of the total sample set for smaller areas.

To address the high margin of error while still meeting the goals of high geographic precision and manageable sampling requirements, the ACS produces five year estimates as well. The ACS 5YR is a moving measure of the estimate in a geography created by using the samples from the current year and the previous four combined. For instance, the ACS

2016 5YR estimates are built on data collected in 2012, 2013, 2014, 2015, and 2016. While the population of a geography may increase over this period, the sample size increases at a great rate, yielding a sample percentage typically between 7% and 10%, depending on the geography in question. These estimates are not simple averages, but include trends visible in the year over year, and leverages geographic similarities to create stronger estimates for the current year. Additionally, the 5YR includes data for the block group level, whereas 1YR estimates are only available for geographies with more than 20,000 residents.

For the purposes of this analysis, references to the ACS refer to the 5YR estimates, unless otherwise specified. For some more basic statistics, such as population estimates, 2017 ACS estimates are used. However, 2017 data is not yet available for all block group level topics, and as a result some data sources leverage the ACS 2016 estimates.

In our analysis, variables drawn from the American Community Survey calculated at the block group level include population density, and median household income. Population is sourced a total estimate of the population, and is based on 2017 ACS data, table B01003. Percentage minority and median household income data at the block group level is not yet available for the 2017 ACS, so 2016 ACS data was used. Due to the spatial scale of the analysis and the importance of these axes of study, geographic precision took priority over currency. In fact, 2016 ACS results are still considered current in policy creation, and the AMI calculation has a four year lag.

Communities of Concern

While one-dimensional demographic measures are an important way to examine a region, it can be difficult to draw nuanced conclusions from features like median income. One of the central motivations for this research is to understand how scooter trips and societal marginalization might interact. Equity concerns can be difficult to define, and even more difficult to quantify. As one way of investigating intersectional measures of marginalization, I have adopted the Communities of Concern rubric from the Association of Bay Area Governments. Last revised in 2017, this data schema and policy framework is deeply integrated into planning efforts across the Bay Area. It brings together a number of datasets that may be axes of marginalization, including minority population, income, limited English proficiency, zero-vehicle households, seniors above 75, people with disabilities, single parent households, and severely rent burdened households.

The Communities of Concern (CoC) designation is based on identifications of concentrations of the CoC factors. This is done by aggregating regional data on each factor, and identifying census tracts for which the factor rate exceeds one half standard deviation for the regional data. If a tract exceeds one half standard deviation, that factor is flagged as concentrated for the census tract in question. For a tract to be designated a CoC, it must have a concentration of low-income households. If that concentration threshold is met, it must either also have a concentration of minority residents, or have a concentration of at least three of the remaining six factors. This is visualized in the table below.

Low-Income (< 200% of Federal Poverty Level)	Required for CoC designation
Minority Population	Concentration with Low-Income concentration marks CoC
Limited English Proficiency	At least three concentrations and Low-Income concentration marks CoC
Severely Rent Burdened	
Seniors Over 75	
Population with Disability	
Single Parent Households	
Zero-Vehicle Households	

Table 3.4 – Communities of Concern Qualification Factor Rubric

Additionally, the Association of Bay Area Governments extends the analysis to designate tracts of High, Higher, and Highest concern, indicating factor concentrations more than 1 standard deviation (Higher) and 1.5 standard deviations (Highest) above the regional mean. This underscores the assumption that assessing marginality is not simply a binary process, and that the degree of concentration is an indication of the extent of marginalization.

Categorized Level	Index Value	Criteria Equation
Not a Community of Concern	0	$g_{zvh} \leq \bar{x}_{zvh} + \frac{s}{2}$
High Concern	1	$g_{zvh} > \bar{x}_{zvh} + \frac{s}{2}$
Higher Concern	2	$g_{zvh} > \bar{x}_{zvh} + s$
Highest Concern	3	$g_{zvh} > \bar{x}_{zvh} + \frac{3 * s}{2}$

Table 3.5 – Communities of Concern Factor Equations

Most regions have their own methods of conducting a federally-mandated equity analysis, but the clear focus on a broader understanding of the multi-faceted nature of marginalization makes the CoC rubric well suited for these purposes. The Communities of Concern framework has not been without controversy, recently due to its involvement in the CASA Compact, a housing policy group convened by the Bay Area's Metropolitan Transportation Commission. However, the modifications put forward under the new rubric of "Sensitive Communities" served primarily to reduce the intersectionality of the measure, and also set absolute cutoffs rather than relatives concentration cutoffs. In early iterations of the CoC designation, census tracts were designated if data indicated their factor rate was greater than a set percentage. For Minority Population, the cutoff percentage was 70%, and the Low-Income percentage was 30%. However, these cutoff percentages were often misaligned with prevailing rates of the disadvantage factors in the region, and while some census tracts were still designated, this cutoff methodology was somewhat arbitrary. The old cutoff table is included below.

Proposed Communities of Concern Framework for Plan Bay Area 2040

Disadvantage Factor	*% Regional Population*	*Concentration Threshold*
1. Minority	58%	70%
2. Low Income (<200% Federal Poverty Level - FPL)	25%	30%
3. Limited English Proficiency	9%	20%
4. Zero-Vehicle Household	10%	10%
5. Seniors 75 Years and Over	6%	10%
6. People with Disability	9%	25%
7. Single-Parent Family	14%	20%
8. Severely Rent-Burdened Household	11%	15%
Definition – census tracts that have a concentration of BOTH minority AND low-income households, OR that have a concentration of 3 or more of the remaining 6 factors (#3 to #8) but only IF they also have a concentration of low-income households.		

Table 3.6 – Deprecated Version of the Communities of Concern Framework with Set Thresholds

For the most recent update to the CoC methodology, ABAG converted to using standard deviations from the regional mean as a measure of concentration. This is significantly more appropriate for a number of reasons, primarily that this switch makes the measure a

stronger indicator of relative disadvantage concentration. Furthermore, the method eliminates a potential opening for political manipulation of the rubric. Importantly for this research, it means that because the concentrations are relative to the focus area, the standard can be translated to other regions without methodological adjustments. Policy implications and uses of Communities of Concern are discussed in Chapter 6.

Longitudinal Employment Household Dynamics

The Longitudinal Employment Household Dynamics (LEHD) data is a comprehensive dataset tracking employment, residential location of workers, and home to work travel in the United States. This data product is administered by the U.S. Census Bureau, and is released on an annual basis. As of the time of this paper, the most recently available data is from 2015. In contrast to the American Community Survey, the LEHD data is assembled primarily through existing state administrative records, though the data is then supplemented and refined by incorporating information from tax records, business surveys, and other Census products . As per its name, the Longitudinal Employment Household Dynamics data leverages persistent data over time to develop highly detailed projections of residency and employment.

The LEHD is not the only data source to report on employment factors, but LEHD data is available at the census block level, the smallest possible geographic division. This provides the unique ability to compare both regional level statistics with hyperlocal data. Census block level data is not available through the American Community Survey, which disaggregates only to the census block group level. In urban areas, census block groups may contain nearly 200 census blocks (Graham, Kutzbach, & McKenzie, 2014). The geographic specificity is critical for spatial analyses in urban areas, and was a significant factor in selecting this data as an appropriate analysis component.

For each census block record, the LEHD categorizes data across a range of factors. Three factors of interest are gender of employee, the wage category of the job and the division of workers in the census block by age. Based on existing literature, jobs by industry may be a relevant factor, especially technology-oriented jobs (El Zarwi, Vij, & Walker, 2017). Additional categories in the LEHD include, ethnicity of employee or resident, further industry disaggregation and whether a job is considered a “primary job”. The LEHD also tracks the number of federal employees in a geography.

The LEHD does not provide a measure of central tendency for its data, and instead provides numbers of employees in each category. While a weighted average can be computed,

the ability to understand a detailed distribution of the factor in question is hampered by the limited number of histogram categories. For both age of employee and wage, the LEHD provides only three categories.

The LEHD claims that the data is available at the census block level, and while that is technically correct it elides the reality that census block data is full coverage. In practice, the three cities in this analysis had reported data for approximately 40% of census blocks. However, based on the spatial distribution of gaps and the small polygon size, interpolation is an appropriate choice.

I used Inverse Distance Weighting with a decay factor of 2 to create an employment surface. This closely matches the rough plume file provided by the LEHD, but with far more fine-grained variation.

The Inverse Distance Weighting function is below:

$$z_p = \frac{\sum_{i=1}^{n}\left(\frac{z_i}{{d_i}^p}\right)}{\sum_{i=1}^{n}\left(\frac{1}{{d_i}^p}\right)}$$

Local Climatological Records

Obtained from the National Oceanograpic and Atmospheric Administration (NOAA), this dataset provides hourly reports on local weather conditions from automated weather stations in each city. The stations used in this analysis are all located at the international airport for the city, which were all either directly adjacent to or within the selected analysis area. Each record in this table corresponds to a single hour, with data available for the previous fifteen years. Each weather station uploads data to the NOAA data warehouse automatically, and it becomes available for download within the same day. This dataset is therefore an excellent source of regularly formatted current and historic weather data. In the analysis at hand, the only columns examined are Dry Bulb Temperature, which corresponds to the air temperature as normally reported, and total precipitation within the past hour. Columns not analyzed at this

time, but which may be of interest in future work include humidity, cloud cover, wind speed, and Wet Bulb Temperature, which may be incorporated into a heat index.

Road Network

The use of a road network as a possible explanatory variable is critical, as we are examining behavior that takes place almost exclusively within the public right of way. Road networks in this analysis were drawn from the TIGER geographic database, using the most recent (2017) vintage. While this national database may miss some very recently constructed roads, it is centralized and consistent, and can be adapted easily across cities or regions in the U.S.

As the length of the road network is a useful variable insofar as scooters can travel along roads, removal of roads that do not permit scooters was appropriate. However, detailed data on vehicle prohibitions is not available in a centralized location. However, any road lines identified as Interstates were removed from the road length calculation. They were, however, included in any potential car routing operation.

Chapter 4 – Analysis

Ride Characteristic Analysis

The first step in this examination is to conduct a straightforward set of non-geographic descriptive analyses on the data. As described in the prior section, all of these analyses will be initially conducted on a 10,000 row subset of the MDS database. The results are disaggregated by city.

Trip distance analysis will be based on the Trip_Distance column reported within the MDS data. This is an integer representing the number of meters between each GPS reporting point, summed over the ride. This figure can be subject to errors, and is cleaned as described in the Outlier Treatment section below.

There are many additional metrics by which to measure the characteristics of each ride. While they are not included in the present analysis, a number of them have been explored and generated during the process of data cleaning. They will be incorporated in future iterations of this research, and are described in Chapter 7.

Trip directness is a measure of route efficiency, common in routing analyses. I will be creating two simple versions of this that can be calculated for each trip without referencing a streetmap, for use as a baseline and to allow for intercity comparisons. I will calculate a raw distance (in meters) from the origin and destination coordinates, labeling this as D_{crow} (as the crow flies).

In general, distance calculation from pairs of coordinates will use the following equation:

$$D_{seg} = 2\sin^{-1}\left(\sqrt{\left(\sin\left(\frac{lat_{x-1} - lat_x}{2}\right)^2\right) + \cos(lat_{x-1})\cos(lat_x)\left(\left(\sin\frac{lon_{x-1} - lon_x}{2}\right)^2\right)}\right)$$

Implemented using the *sf* library in R, and the *st_distance* function, I vectorize this function to work on an array and return values as a new column. This uses the Haversine formula to return great circle distances along the surface of the earth, rather than calculating a straight line between two points. The distances will be returned in meters, which aligns with the units in use in the MDS schema.

For D_{crow} , I will use the origin and destination points, with no intervening route points. To calculate the reported distance based on route points, I will implement the preceding distance formula as a summation.

$$D_{route} = \sum_{i=0}^{n} D_{seg}$$

This will provide a sum of the straight line distances between each reported route coordinate. This number will be longer than (or in a few cases equal to) D_{crow} , but will better reflect the path taken by the rider.

I must also validate the reported Distance column, which I will do by calculating a ratio between D_route and Distance.

$$\frac{D_{route}}{Distance} = Directness_B$$

If Dir_B is consistently equal to 1, the reported distance is calculated in the same way. If Dir_B is consistently less than 1, the Distance column is likely a better representation of the actual traveled distance, reflecting better detail than the route coordinates.

$$\frac{D_{crow}}{D_{route}} = Directness_C$$

$$\frac{D_{crow}}{Distance} = Directness_A$$

These three ratios will serve as measures of route directness.

Time Variables

All analyses will be run on the full data set for each city. However, following identification of peaks in early data exploration, we additionally decided to run our analyses segmented by time of day. To allow for human legibility and increased applicability to potential policy outcomes, the time bins were selected with breaks on round numbers, rather than statistical inflections in the ride start distribution. The time bins are presented in table 4.1.

Time Bin	Start Time	End Time
Morning (Weekday)	00:00:00	10:30:00
Midday (Weekday)	10:30:01	15:30:00
PM Peak (Weekday)	15:30:01	18:30:00
Evening (Weekday)	18:30:01	23:59:59
Weekend	00:00:00	23:59:59

Table 4.1 – Time Bin Boundaries and Definitions

The data in this study evolves over time, as additional scooters are added to a city, weather changes, and people become more or less comfortable with the new mode. Tracking changes in general descriptives over this time span is important, as use patterns likely evolve. However, the scope of this research did not permit full incorporation of this research aim. Future iterations and extensions of this work will explore the questions and analyses described below.

$$Perc_{change} = \frac{rides_x - rides_{x-1}}{rides_{x-1}}$$

Outlier Treatment

The identification and treatment of outlier data is a crucial step in any analysis. With such substantial database sizes, this proved difficult as there was no feasible way to visually

check all rows. As such, outlier cleaning was a scripted process, though the definition of outlier boundaries was supervised.

One major issue in this data set is the dependence on a well-functioning GPS device in each scooter. The MDS data is calculated and extrapolated from the GPS reported location and the internal device clock, and a brief misalignment of either of these can result in extremely messy data. Initially, I wrote a cleaning script to address out of line points during the linestring transformation. This script worked along each geolocated point and timestamp in series, measuring the distance and time interval between each point and the previous point in the trip. If the straight line distance traveled over the reported interval was such that the scooter would have had to be traveling at over 30 miles per hour to move in this way, the second point was removed. The scooters have a limited top speed of 15 miles per hour, and the straight line calculated elides any corners, meaning that if every reporting device is working as intended, the speed should never rise above 15 miles per hour, with a tendency to indicate lower speeds if turns occurred in real life between GPS reports. However, even after allowing for the possibility of a downhill acceleration to 20 miles per hour, a low cap resulted in elimination of 30.2% of the points on each ride, an average drawn from the 10,000 row test set in Nashville. Raising the limit to 30 miles per hour reduced point removal to 9.1% of points within the test set.

An important caveat to this cleaning script was the stated minimum accuracy of points within an MDS trip. Throughout the dataset, Gust reports minimum accuracy as within 152 meters of the actual location. This is a significant margin of error. While most points appear to be on streets and in logical paths, this leeway reduced the efficacy of detailed cleaning scripts, and a decision was reached to use the uncleaned, directly drawn data for routes, origins, and destination points.

One common error was a single errant GPS report that placed the scooter up to 7000 miles away for a period of 6 seconds within a single ride. While the cleaning script above sometimes dealt effectively with these errors, it was not consistently successful. Additionally, the distribution of scooter trip distances appears to have a very long tail. Initially, a test was run identifying trips with distances above 3 standard deviations from the distance mean. However, this grouping captured well over 15% of the trips, many of which appeared to be true uses of the scooters, based on visual inspection. After calibration, it was decided to flag trips above the 98th percentile for distance as potential outliers. Similarly, trips with distances under the 2nd percentile were flagged for potential outlier treatment on the other end of the distribution.

To address concerns of errant GPS points, aggregate spatial boundary calculations were run to create flags for potential outliers. Trip records were given three potential boundary-based

flags, set as logical vectors. If a trip's origin lay outside the boundary of the analysis area, it was flagged, with the same process for the destination point. If a trip's original reported path traveled outside the boundary of the analysis area, it received a flag as well. Naturally, if an origin or destination point was reported outside of the boundary, the route also traveled beyond the boundary, meaning that these trips received two flags. This aligns with the goal of identifying trips that may throw off the analysis of within-city travel, and was not considered a problem.

The cleaning script was incorporated in outlier treatment by comparing the number of vertices present in the original trip with the cleaned trip line, and generating a difference column. If a trip had a "vertex difference" value above the 95th percentile, it was flagged for possible removal, based on the messiness of the GPS reporting.

Once these flags had been created, a column was created that summed the flags on each trip. If a trip was marked with three or more outlier flags, it was dropped from subsequent analysis.

Hexagonal Binning

To compare spatial datasets of different types and sizes, a common and well-established technique is to convert to rates, which in this case would be percentages or densities per unit of area. Two primary paths present themselves: the use of existing geographic boundaries or spatial binning.

While a conversion to factor rates appropriate for comparison can be done using existing administrative boundaries, there are trade-offs between analysis scale and human legibility. City boundaries are usually understood, but are too large for detailed spatial analysis. Neighborhoods within cities could provide promise, but neighborhood boundaries issued by cities are often subjects of debate, and often remain too large to reflect significant nuance. Smaller administrative boundaries such as census block groups typically have little real-world significance. This is not universally true, and exceptions exist when the border is formed by a geographic feature such as a river or highway.

$$density = \frac{n_{observations}}{area_{geography}}$$

Another issue with using existing geographic boundaries is the widely variable land area that objects of the same type may cover. While the conversion to rates solves this problem on a computational level, the problem of visual distortion remains. The varied shapes of census block groups lead to drastic changes in rates in unpredictable ways.

The other major option for spatial organization is spatial binning. This type of analysis has been common in physical sciences for a long time, as it has a number of advantages for categorizing disparate data sets. In particular, ecology and environmental studies researchers have leveraged this approach when dealing with topics that either fall in areas with few administrative boundaries, or do not respect administrative boundaries. Studies of flora or migratory patterns of fauna are classic examples of traditional uses of spatial binning (Lewin-Koh, n.d.).

Physical data is another common use of spatial binning, typical in the form of raster images. When most people imagine a digital image, they are thinking of a raster image comprised of a grid of square pixels. Pixel images are the most common form of two-dimensional spatial binning. In the creation of a digital image, each pixel is created through the average of a color spectrum present in the real world. Typically, each pixel holds a value for red, green, and blue. These wavelength intensities are proportionally combined to create a single color that fills the entire pixel. If the size of each pixel, or cell in the grid, is very large, a number of input features must be combined to generate a single value. The image, or grid, loses fidelity and can be hard to interpret. However, if a grid cell becomes too small, the possibility that a grid cell represents either statistical or visual noise increases, depending on the accuracy of the detection mechanism. In a camera, this is a function of the imaging equipment. In a geographic spatial bin, this is a function of the spatial fidelity of the data to be binned.

In an image, bins are shaped as square tiles. Squares are an efficient and appropriate choice for a number of reasons, include ease of calculation and efficient tessellation. They also allow for efficient reference to a two dimensional table, as they perfectly align in rows and columns. This indexing ability is critical in a number of high intensity processing environments, and makes them well suited for applications that demand rapid production and display of the binned grid, such as digital video or computer graphics. And yet, in certain contexts the reliable square bin has recently lost ground to the hexagonal bin, which presents a number of possible advantages.

Square bins suffer from a lack of internal representativeness. After calculation of input attributes, the entire bin is presented as the same value. This new value is an appropriate representation of the center of the square, as the calculation is typically a measure of central

tendency of some sort. However, the further one travels towards the edge of the bin, the less representative the presented value becomes. One can imagine the process of calculating a square bin value as taking the measure of a micro-distribution along two axes. The further one moves from the center of the curve, the more likely that the underlying value is not well-represented by the summary statistic. This is exacerbated even further when you consider the corners of a square bin, each of which contains a tail from both micro-distributions. Finally, consider that the corner of each bin is spatially adjacent to three other corners, each of which is similarly unrepresentative, multiplying the obfuscation that has the potential to result with the use of square bins.

--Illustration

To ensure that every point in the bin was as well represented as possible, one would need to bin on a circle. Corner effects would be eliminated, and a point could only lie at the tail of a single possible curve of computation. Unfortunately, circles do not tessellate, and can therefore not effectively cover a plane in a complete grid. Circles would be ideal because of their infinite number of mathematical vertexes; just as tessellated triangles would have even greater corner distortions than squares. The regular polygon with the highest number of vertexes that tessellates is the hexagon. The use of a hexagonal grid is the most appropriate way to reduce misrepresentation at the edges of a spatial bin.

If one goal of the binned data is visual communication of the underlying information, hexagons hold another advantage over squares. Human vision works by gradually building complex shapes out of a series of lines at various angles, which come together to form edges and eventually objects. The larger a continuous edge or line, the more it helps our brain process and understand the visual field. This is an efficient processing strategy, but leads to a pitfall when presented with square gridded visual data. The divisions between columns and rows, which stretch in an unbroken straight path across the whole image, may begin to dominate our comprehension of the visual information. Hexagon grids do not present continuous straight edges or unbroken rows of bins, and are far less susceptible to the same sort of visual permutation (Strimas-Mackey, n.d.).

Proportional Split

To shift spatial data from the original geometries to hexagonal bins, I relied on a proportional split calculation. In the same way that hexagonal bins act to summarize the more detailed data from the underlying geography, so too do geographies like census block groups.

The proportional split is a weighted average, finding a representative value for the hexagonal bin in a few steps. This process is also depicted in the equation below.

The first step is to trim all geometries, both the census areas and the hexagonal bins, by any water features that cover part of their area. This is appropriate because all statistics that we have measured are related only to activities on land. This clipping was completed for census geographies before calculating densities, but must be completed for the hex geometries as well. The same water boundary shapefile, sourced from the TIGER File USGS database, is used in both operations.

Next, the hexagonal grid is used like a cookie cutter or mesh slicer, in an operation called an Intersection. This operation is performed using the Postgis standard, to comply with MDS schema guidelines. The Intersection operation divides the input features into smaller features along the intersection boundaries, but each new feature contains the same data values as the original feature. This means that each small bit of a divided census block group contains the rate information for the whole census block, which is appropriate considering that the rate information was calculated to best represent the census block group as a whole.

Next the area in square meters of each new geometry is calculated. If the new calculated areas were summed, grouped by the original census block, they should equal the area of the original feature, and we use this method as a validation of the calculation. The area of this small shape is then divided by the area of the hexagon (minus any water features) that it sits within. This operation results in a percentage of the hex area occupied by a portion of the original census block group.

This percentage is then multiplied by the factor rate for the census block group, representing the portion of the final hex value that comes from each underlying census block group. This set of operations is calculated concurrently for all features in the dataset, so that each Intersected geography holds a percentage of the hex area it occupies multiplied by the factor rate.

Finally, these products are summed within each hex, providing a single factor value produced by weighted spatial average. The sum of the percentage coefficients for the input factor rates is 100%, and this is confirmed for each hex as another method of validation. The equation is displayed on the following page.

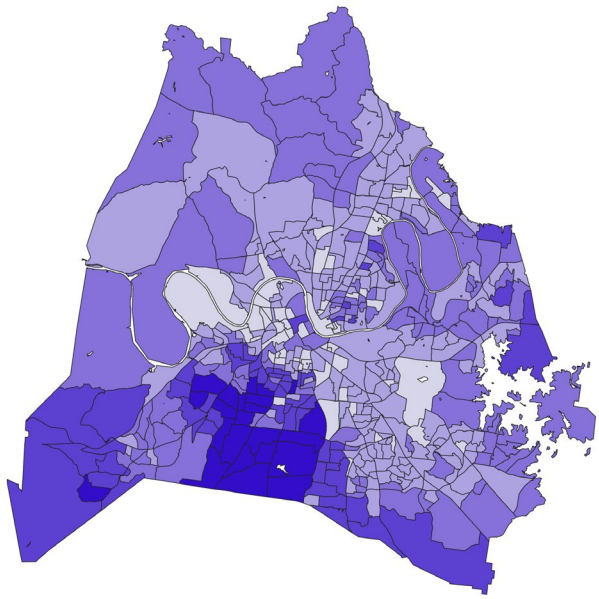

Figure 4.1- Nashville Median Household Income by Census Block Group – Darker is Higher

$$Density_{hex} = \sum_{i=1}^{g} \frac{area\left(\left(shape_g \overset{clip}{\Longleftarrow} shape_{water}\right) \overset{intersect}{\Longleftarrow\!\!=} shape_{hex}\right)}{area\left(shape_{hex} \overset{clip}{\Longleftarrow} shape_{water}\right)} * Density_g$$

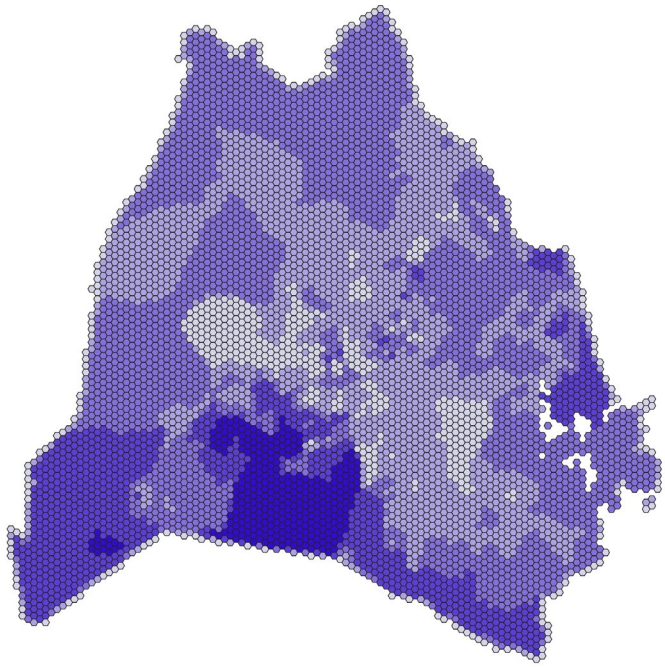

Figure 4.2 – Nashville Median Household Income after Hexagonal Spatial Bin Transformation – Darker is Higher

Hexagonal Binning for Generated Geographies

The above section covers the spatial binning procedure for existing polygon geometries, such as census block groups. A different procedure is used to bin for geometries that arise from point or line data sources. For our analysis, the most common point data type is ride origins and destinations, but these are treated differently in the regression analyses and are not grouped here. The relevant point data includes the location of transit stations and other possible points of interest.

To assess the degree to which these point features attract spatial activity, buffers will be calculated using a network analysis procedure. A simple geographic Buffer operation takes an input distance *x*, then creates a shape with an edge consistently *x* units further away from the edge of the original geometry. If the original geometry is a point feature, the buffer takes the shape of a circle with radius *x*. This type of simple buffer is both useful and easily calculated, but is not appropriate for all circumstances. The assumption of the simple buffer is that proximity to the input geometry is equally relevant from all angles. However, in urban environments, this is plainly not the case, and can be clearly illustrated with an example.

Imagine a point feature located in the middle of an intersection. We wish to generate an area that a hypothetical person could access within 100 meters. If the person proceeds along the four legs of the intersection, they encounter no problems and the edge of the buffer should be 100 meters away from the origin point. However, if they attempt to advance at a 45 degree angle, their intended path would take them straight into a space occupied by a building. Instead, the buffer is calculated by proceeding along open network segments, branching at each opportunity, and expanding outwards following the street network until the buffer distance has been reached. The outside edges of this branched network are then connected and collapsed into a polygon. In this way, a network analysis buffer far more accurately captures the area that is a 100 meter walk from the input point, rather than using a simple radial buffer.

Future iterations of this work will include network analysis buffers for transit stations. These will be calculated using the Open Street Map network files, starting from the transit station coordinates provided by the city's GIS portal.

For transit stops, two overlapping network analysis buffers would be generated. Scooters can be ridden directly to the operator's destination, but the beginning of a trip frequently requires a short walk to locate the nearest available scooter, as identified on the app. Therefore,

this analysis will generate a smaller network analysis buffer to be used when examining destinations than when examining trip origins. For destination calculations, the size of the buffer would be 50 meters, while for origin calculations, the size of the buffer would be 250 meters.

The creation of these buffers results in a series of polygons. Unlike census block polygons, these buffers may be overlapping, reflective of the dense nature of transit stops in some urban areas. In this proportional split, each polygon will be intersected with the hex grid, the area of the split section divided by the total area of the hex to create a percentage, and then all polygon percentages within a hex will be summed. With overlapping polygons, this may result in percentages above 100%, which is in line with our analysis.

The street network is a line shapefile, and cannot be conceptualized as taking up some percentage of a hex. Therefore, the spatial binning operation looks somewhat different. The city street network file is intersected by the hexagon grid, resulting in millions of small line segments, each tied to the underlying hex. The length of each segment is then calculated, and the length of all segments in a hex is summed. This operation results in value representing the total distance of road network within a hex, and because this value is tied to the grid cell, can be used in the regression analysis.

Rebalance Points

An important consideration in analyzing where scooter trips occur is the recognition of the prerequisite to scooter travel: a scooter must be present. There are two possibilities that could lead to a scooter being available in a location, ready to use on a trip. The first option is that the previous rider ended their trip, parked their scooter, and left it available for use by the next person. This point to point, non-directed distribution of the fleet is critical to all sorts of micro-mobility fleet models. However, a scooter may also be placed at a specific location without having ended a trip there. This can happen after someone has charged it overnight and deploys it in the morning, or it can be actively moved from one location to another. These are both *rebalancing* deployments, and represent the biggest lever with which Gust or the city can determine where people have the option of choosing a scooter for their trip.

The goal of this analysis is to assess what factors might lead to scooter use, and the active choices by a policy-maker to place a scooter in a location muddles our assessment. However, by using the number of rebalance points in the regression, we isolate the environmental and demographic factors, clarifying the conclusions and interpretation of the regressions.

The characteristics of rebalancing points is an additional interesting question, and serves as the dependent variable in an additional set of analyses. If rebalance rates are tightly correlated with other environmental and demographic factors that do not emerge as significant predictors in the regression, this may be an indication that what may appear to be differences in use by location are more due to the availability of scooters.

Rebalance locations are not provided in the MDS Trips data. The MDS Status Changes data schema contains rebalancing points and more, but this table was not part of the MDS schema until partway through the study period. Additionally, the formatting of this table has changed significantly over time, and cleaning data so it is both backwards and forwards compatible was not a practical task. For this study, rebalance points were calculated through a series of geometric calculations.

First, the tables of origin and destination points by city were loaded and stacked, so they were a single column alternating between origin and destination. The type of point was noted in a parallel column. The table was then grouped by device_id (unique to each scooter) and ordered by event time within those groups. The distance from each point to the previous point in this table was then calculated, and saved in a new column. Following this operation, destination points were dropped, leaving origin points and the distance that each origin point was from the previous reported destination of the same device.

If an origin point was less than 50 feet from the previous destination point, it was exported as a "stay" trip. If the origin point was greater than 50 feet but less than 500 feet away from the previous destination, it was considered "local", and not a rebalancing trip. The reasons here are twofold. The first is that the allowable error for GPS points in MDS is significant (152 meters), and GPS errors of over 50 feet were not uncommon, and may not represent a moved scooter. The second is that based on news reports and information provided by Gust, it is not uncommon for individuals to move scooters some distance without activating a ride. The scooter is not operational and the wheels are locked, but the possibility of scooters moving in this fashion, combined with the GPS error potential, led to the creation of this gray zone of "local" movements. If an origin point was over 500 feet from the previous destination point for that device, the trip from that origin is considered to be on a rebalanced scooter. The coordinates and time information for the rebalance event are drawn from the origin of the trip.

These points are aggregated by sum into the intersecting hex, and that count is used as a factor in the spatial regression.

Network Distance to Central Business District

Initial exploration of the data revealed heavy clustering near downtown business districts, and we decided to include a score for each hex of the distance to the center of the city. This is common practice in many modeling studies. All three of our cities have significant water features, so it was not appropriate to use Euclidean distance to generate this score. Because Open Trip Planner was already a configured piece of the analysis toolbox (described below), we used this as our network distance routing engine.

The central business district (CBD) centroid was identified by selected the three adjacent hexes in the urban core with the highest job density. A centroid was then calculated for these three hexes, as well as for every hex in the city. A routing batch, using the "walk" mode to get the most direct distance, was returned from OTP, with the origin at each hex centroid to the CBD centroid.

Open Trip Planner "walk" works well for parks, so we hoped to limit the number of hexes without a network distance to the CBD. However, we did find some where that was the case. To address these gaps in our variable, we used *st_distance* to derive the Euclidean distance from each hex point to the central business district centroid. We then calculated a ratio of the network distance to the Euclidean distance for each hex that had both present, giving us a range of ratios from slightly greater than 1 to under 3, depending on the city. Initially, we had intended to use the average of these ratios from neighboring hexes as a scale factor for the Euclidean distance. However, this operation proved to be unstable due to clumps of missing network distances (and therefore ratios). We then estimated a linear model of the Euclidean distance to network distance to return a simple coefficient and intercept. These were used to transform the Euclidean distance into value to fill missing CBD network distance hexes.

Proximity to Public Transit

Using GTFS provided coordinates for public transit stops, a circular buffer was calculated around each stop point with a radius of 500 meters. This is a standard rule of thumb for transit stop access (Canepa, 2007). Each stop point generated its own buffer, most of which overlap with neighboring buffers. Following the buffer operation, these polygons were intersected with the hexagonal grid. This operation cuts the circular polygons with the edges of the grid, and adds a categorical variable to the buffer shapes indicating which hex surrounds each new polygon. Next, the area of each segmented buffer polygon is calculated, and divided by

the standard area of a hex to transform this value into the percentage of a hex that is within 500 meters of the transit stop point. Finally, these percentages are grouped by containing hex and summed. Despite summing percentages, some hexes, especially in the CBD or in transit junctions, have values of up to 40, the equivalent of 4000%. This indicates better transit access, as an individual in that hex would theoretically be within walking distance of 40 different transit stops (as an average across the hex). This variable shows moderate collinearity with Job Density, but was included as it represents an important policy consideration.

Spatial Lag Regression

As part of the exploratory analysis, simple multiple linear regression models were created using the factors described above. These factors are listed in Table 4.3

Factor	*Measure*
Community of Concern Index	Value for Hex
Job Density	Jobs per Hex
Population Density	People per Hex
Median Household Income	Median Income within Hex
Street Length	Length of Road within Hex (meters)
Rebalance Points	Number of Rebalanced Trips per Hex
Network Distance to CBD	Length of Walking Route to CBD (meters)
Near Transit	Sum of Hex Percentages within 500m of Transit

Table 4.2 – Spatial Lag Regression Factors

The multiple linear regression was conducted both as an exploratory analysis, partially to identify variables that may be of use in more complex models. Multiple linear regression is a powerful tool, but because it does not operate in a spatial context, linear regressions for spatial attributes often suffer from acute spatial autocorrelation. Spatial correlation is an underlying assumption of much of this work, and allows us to proceed with spatial binning as a valid technique. However, the same correlation between variables makes attempting to use each as an independent regression record flawed. The simplest way to observe the presence of spatial autocorrelation is to create a plot with our linear regression residuals on one axis, and the mean of the residuals of that object's spatial neighbors on the other axis. If these points form a cluster with a clear trend line, it is likely that the model does not fit appropriately. Residuals will always be present, but if residuals in certain areas of the map are tightly correlated, this implies that the linear model does not fit this region all in the same direction. In the scooter linear model, we see underfitting in the busiest and least busy areas, with the linear model performing well only in moderately active hexes.

A more rigorous way to test for spatial autocorrelation is the Moran's *I* procedure. This test compares the residual distribution of the model in question with itself, and detects correlation between neighboring values by calculating the slope of the residuals. The *I* value is typically between -1 and 1, but when operationalized in *R*, this test returns a *p*-value, from a test to determine if the distribution of residual correlation is statistically distinct from a normal distribution. In all cases, linear models of scooter activity and hex characteristics showed spatial autocorrelation, verified by a Moran's *I* test. With this hypothesis verified, we advanced to spatial lag models.

Spatial lag modeling is a modification on a linear model, but incorporates a major source of spatial error as an explanatory, independent predictor variable. Before running the fit operation, a spatial lag operation determines the spatial lag values of the dependent variable, which it includes as a hidden predictor. In this context, that spatial lag value might be the degree of correlation between the number of scooter trips taken in one hex with the number taken in all neighboring and nearby hexes. Importantly for our purposes, this neighbor operation does not simply factor in the six (in our case) adjacent cells, but factors in values from hexes two more levels away. Each hex considers a wide universe of neighbors in determining its degree of hidden correlation. This makes the operation computationally intensive, but the result is a model that adjusts well to both rapid and gradual value changes over a series of geographies.

In *R*, spatial lag regressions were run using the *spdep* package, and the lagsarlm() function. The outputs from these model appear to be formatted like a standard regression, but because of the interrelation of the input factors, they cannot be interpreted in the same way. The estimate coefficient for each factor represents only the effect of this value on values within its own geometry. However, this level of interpretation misses a large part of the value of spatial lag models, as each geometry has additional impact when you consider the impact on other geometries in its neighbor network. To understand both the direct and indirect effects, we run an impacts() function from the *spatialreg* package. This function provides a simple display of the magnitude of the effect, broken down by Direct and Indirect effects.

Spatial Model Calibration

Correct calibration of the spatial lag model is required to ensure legible output. Furthermore, calibration and scaling of input and dependent variables is important to avoid recursion errors and simulation errors due to mismatched value ranges. To this end, each final spatial lag model was rerun with logarithmically scaled variables. The one exception is the

Community of Concern Index factor, as index factors are not appropriate targets for logarithmic scaling.

Logarithmic scaling distorts the raw value of the input variable, but if all variables are mutated in this fashion, the output is significantly more legible. In a logarithmic output, each unit of independent variable coefficient represents the change one would see in the dependent variable given a 1% change in the value of the independent variable in the environment.

Weather Analysis

The independent variables in this analysis were Air Temperature and Hourly Precipitation. Air Temperature is akin to the reported temperature, rather than a constructed heat index of how the air "feels". Temperature was reported in whole degrees Fahrenheit. Hourly Precipitation is a raw count of the millimeters of rain recorded within the past hour. Each of these variables was construed as an ordered factor level, to allow for dependent variable grouping for summarization functions.

The dependent variables are the median distance of rides in each category, and the trip count rate per category. The trip count rate per category is calculated as the total number of trip counts in hours with the appropriate temperature or precipitation measure, divided by the total number of hours that recorded that weather reading and also contained at least one trip.

Open Trip Planner

Open Trip Planner was developed with the intention of providing an accessible way for public transit agencies to provide the public with a trip planning platform. It has been open source since its inception. Tri-Met, the public transportation agency in the Portland region, led the development of this platform in collaboration with other public agencies, academic researchers, and private companies. There were distinct advantages to an open source model of providing digital infrastructure, and a number of agencies were active contributors to the project. In particular, public agencies were excited by the idea that residents of their region would work to improve the product and code as it needed to be tailored to their local distribution. This would be both open, in line with the open data movement, and also less expensive than hiring a contractor to build out a trip planning web service. This type of web service was in great demand in the period before the dominance of Google Maps. However, the addition of integrated public transportation to the schedules of most personal trip planners reduced the priority of this service for most public transit agencies. While Open Trip Planner is still active on GitHub, most public agencies have reduced backing to the platform.

However, Open Trip Planner holds a great deal of potential as a research tool. Commercial routing engines such as Google Maps have strict API caps and rate limits. To run routing through these services can become pricey, or requires limiting the data collected to carefully selected instances. However, Open Trip Planner can be initialized in a local instance, and supports unlimited queries. If significant amounts of RAM are allocated to Open Trip Planner, it can return requested routes in bulk. Furthermore, as the configuration of OTP is fully customizable, unique features can be incorporated to reflect the parameters of the research at hand.

For this research application, we have created a simple mode comparison by travel time to understand how revealed travel behavior on scooters might have looked if the traveler had chosen a different mode. This analysis is currently exploratory. Chapter 7 lays out a framework for a significantly more robust use of Open Trip Planner to estimate mode shift potential.

Chapter 5 – Results

Time of Day

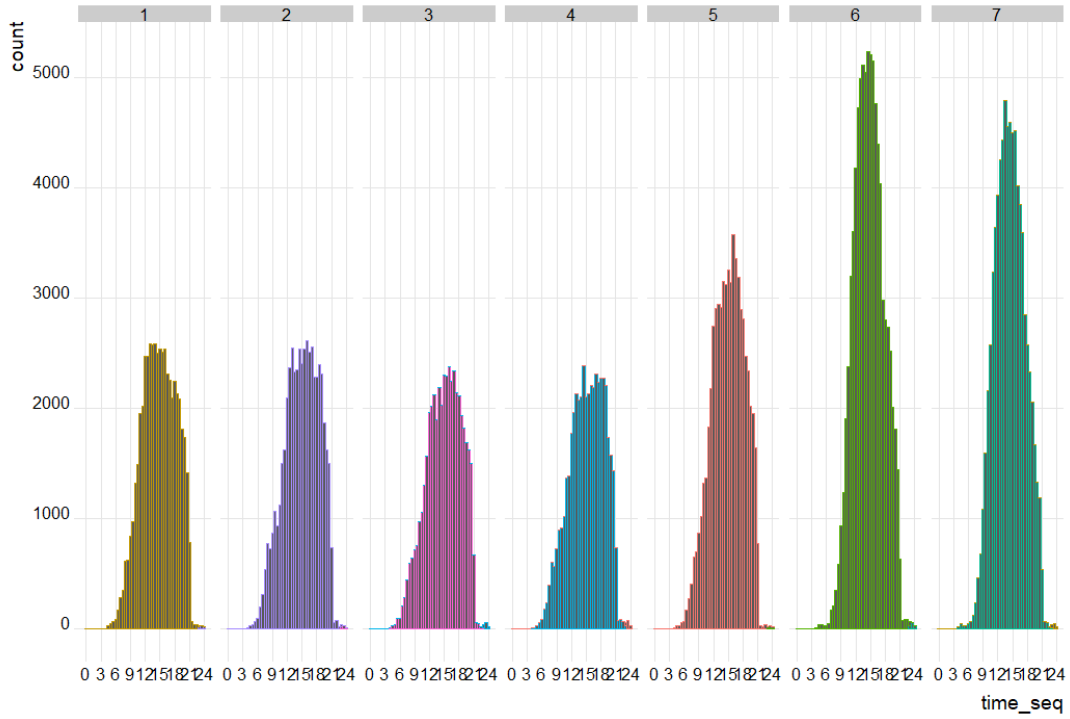

Figure 5.2 – Nashville Trips by Time of Day and Day of Week. Bin Size = 15 minutes

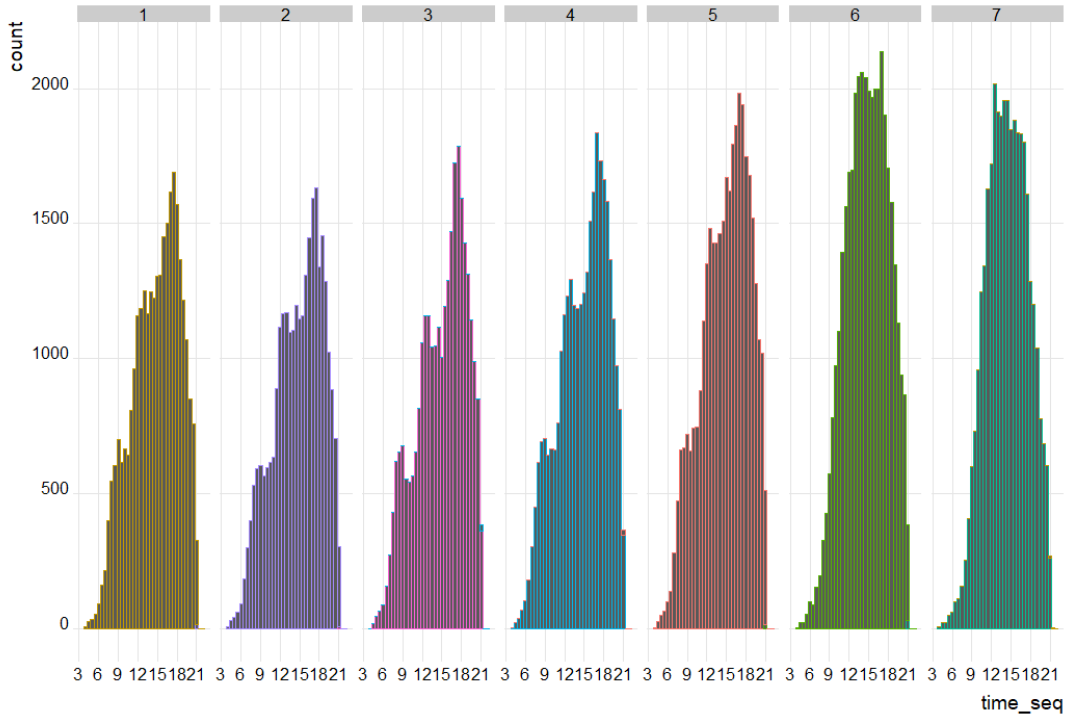

Figure 5.1 – Portland Trips by Time of Day and Day of Week. Bin Size = 15 minutes

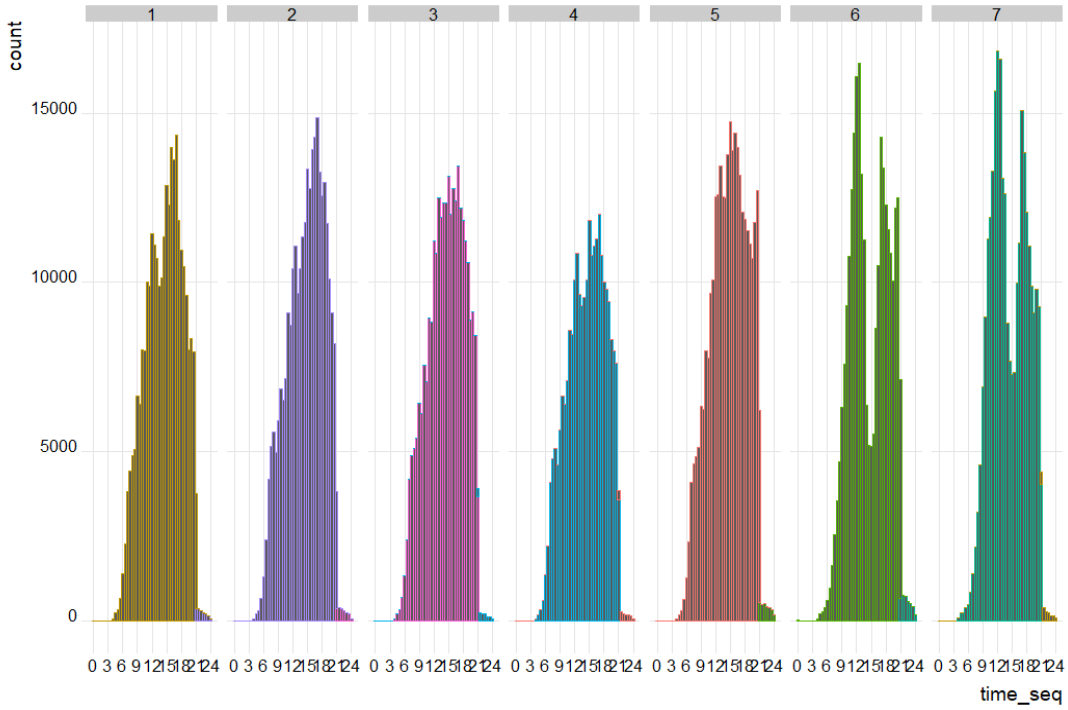

Figure 5.3 – San Diego Trips by Time of Day and Day of Week. Bin Size = 15 Minutes

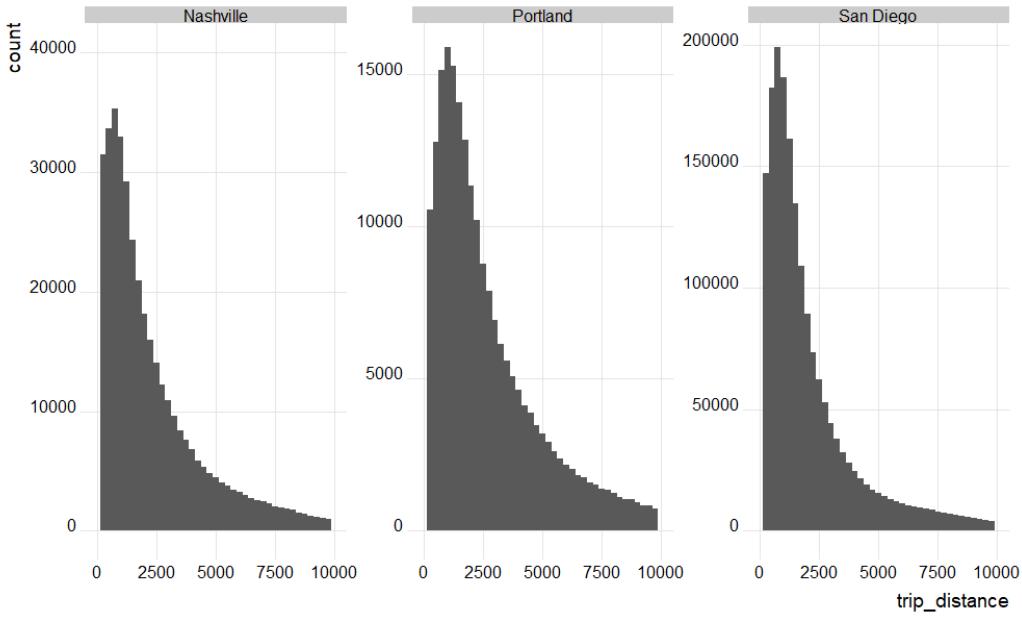

Figure 5.4 – Trip Distance Histogram. Bin Size = 250 meters

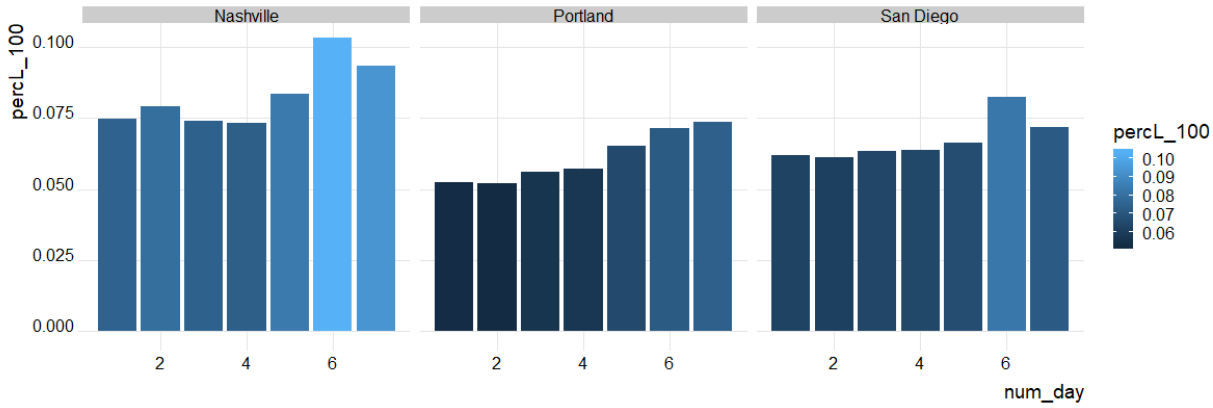

Figure 5.5 – Percentage of Trips Under 100m, by City and Day

Net Change Maps

Nashville

Table 5.1 – Nashville, Number of Trips by Time Bin

Nashville	*Morning*	*Midday*	*PM Peak*	*Evening*	*Weekend*
	00:00-10:30	*10:31-15:30*	*15:31-18:30*	*18:31-23:59*	*00:00-23:59*
Trips	28,830	109,631	73,536	46,086	161,584

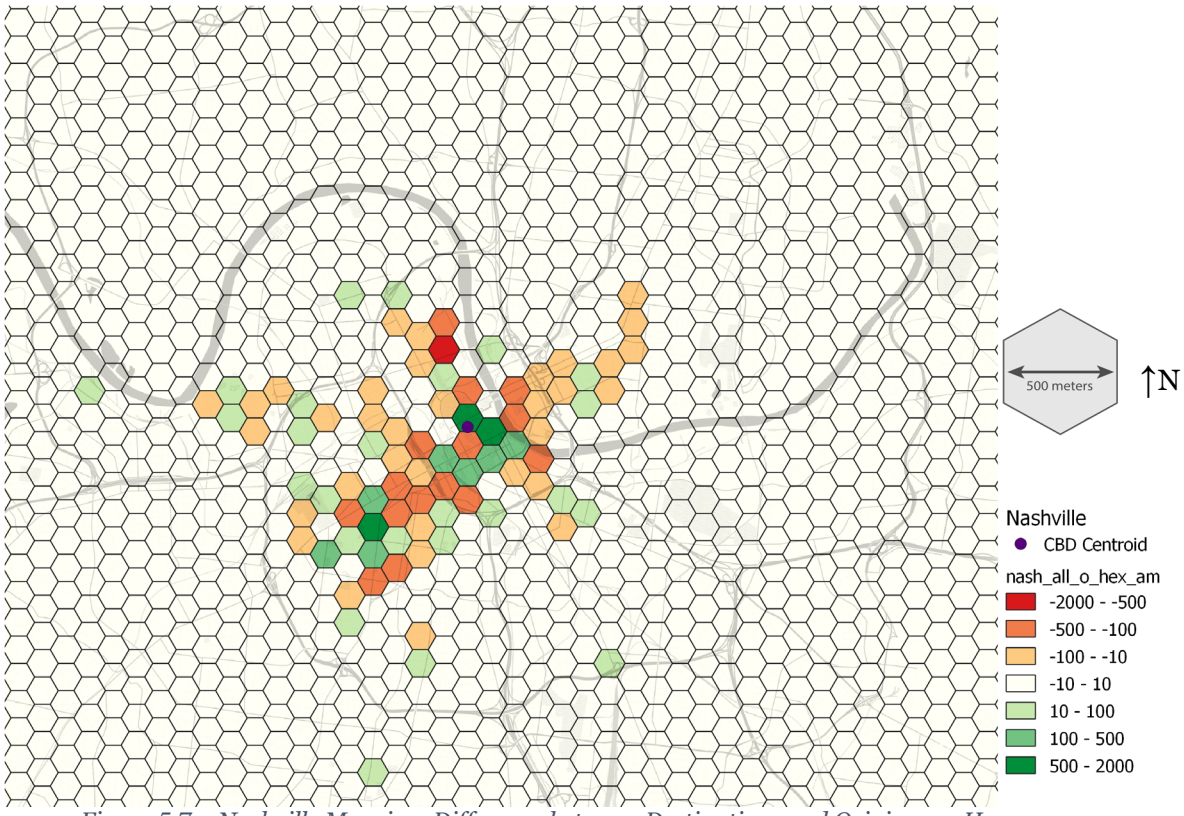

Figure 5.7 – Nashville Morning, Difference between Destinations and Origins per Hex

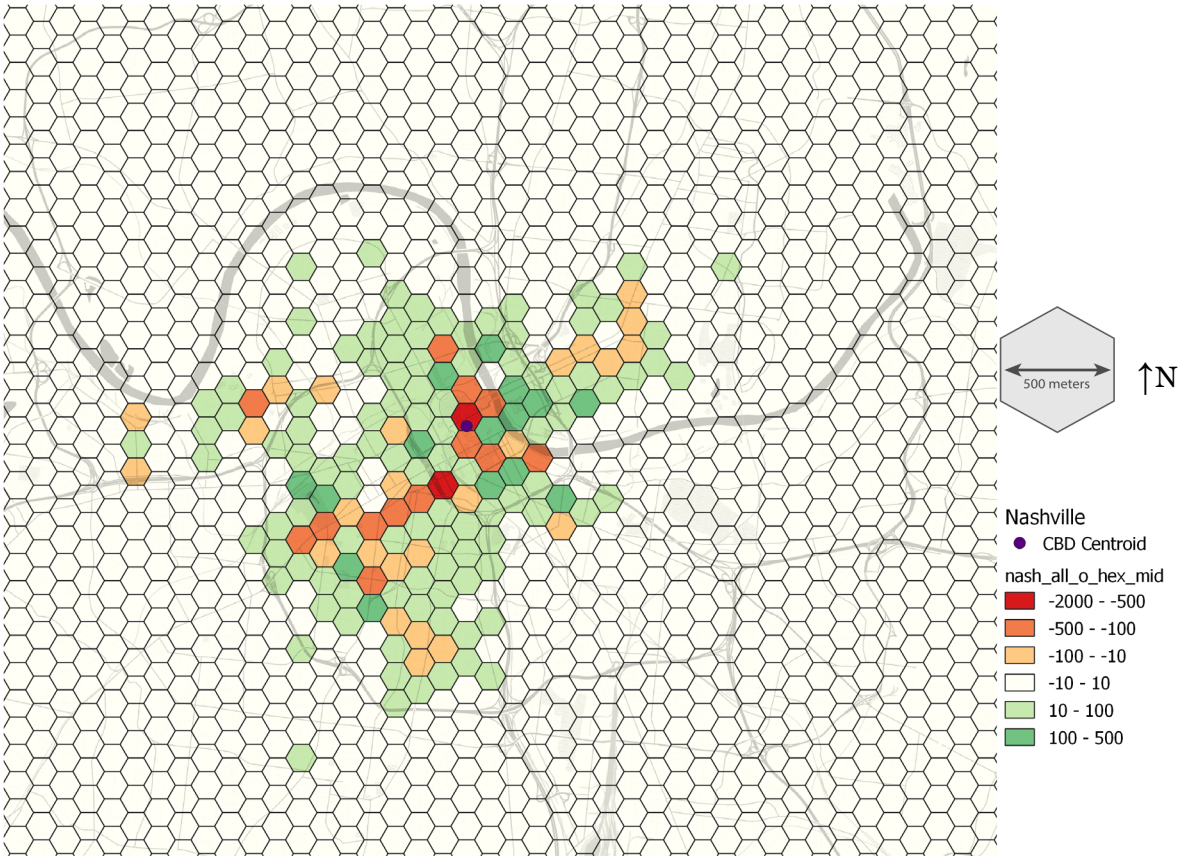

Figure 5.9 – Nashville Midday, Difference between Destination and Origin Counts per Hex

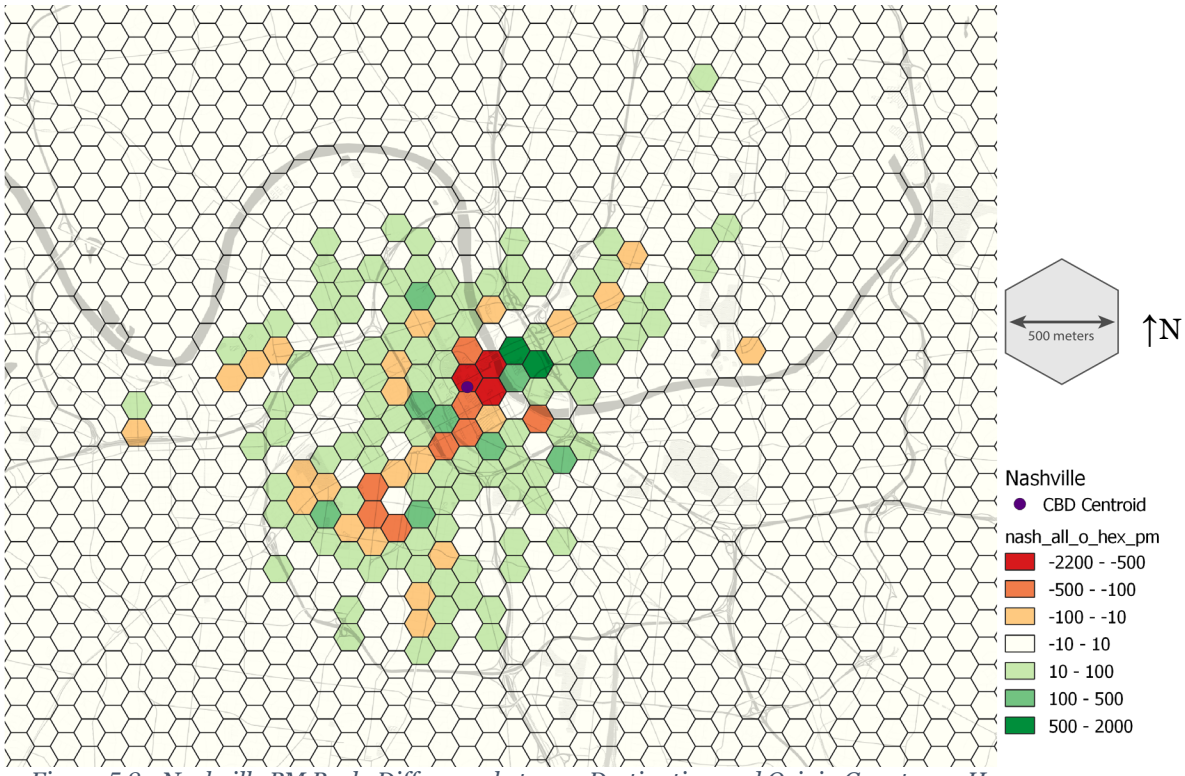

Figure 5.8– Nashville PM Peak, Difference between Destination and Origin Counts per Hex

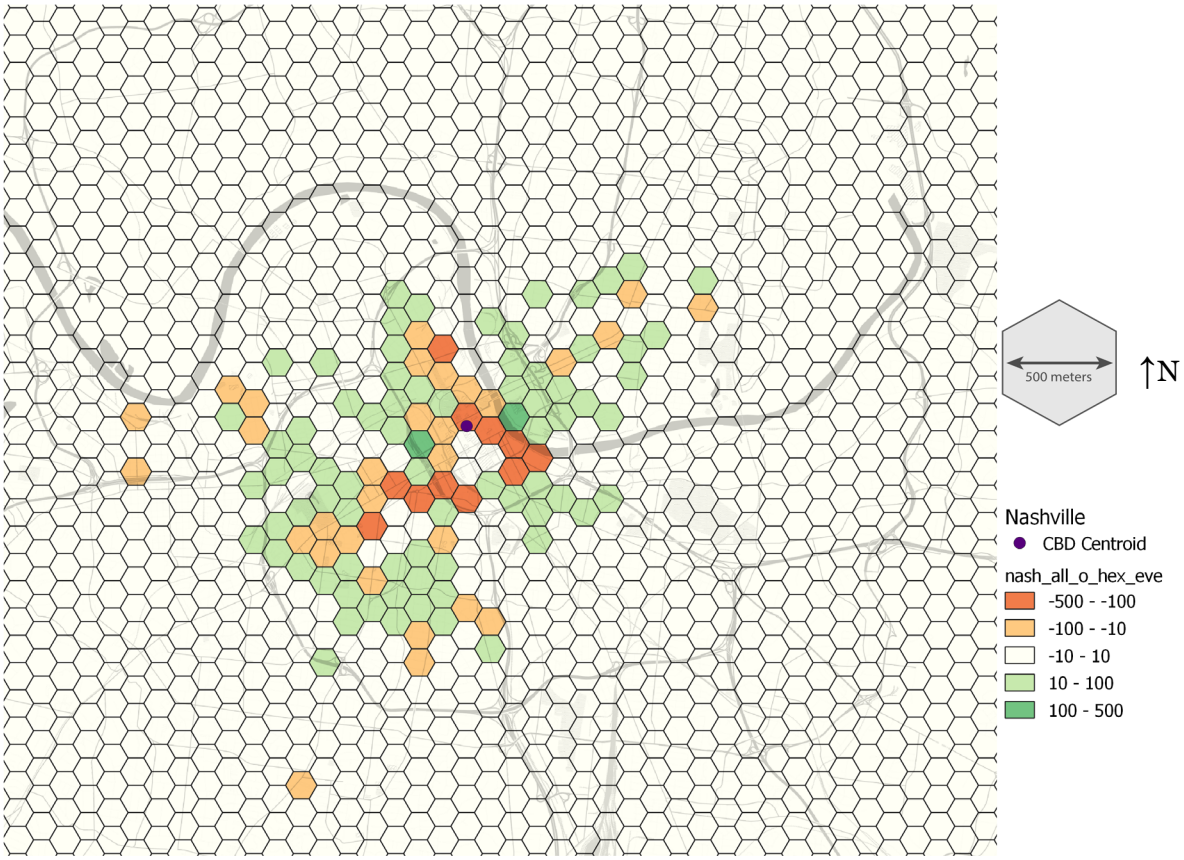

Figure 5.11 – Nashville Evening, Difference between Destinations and Origins per Hex

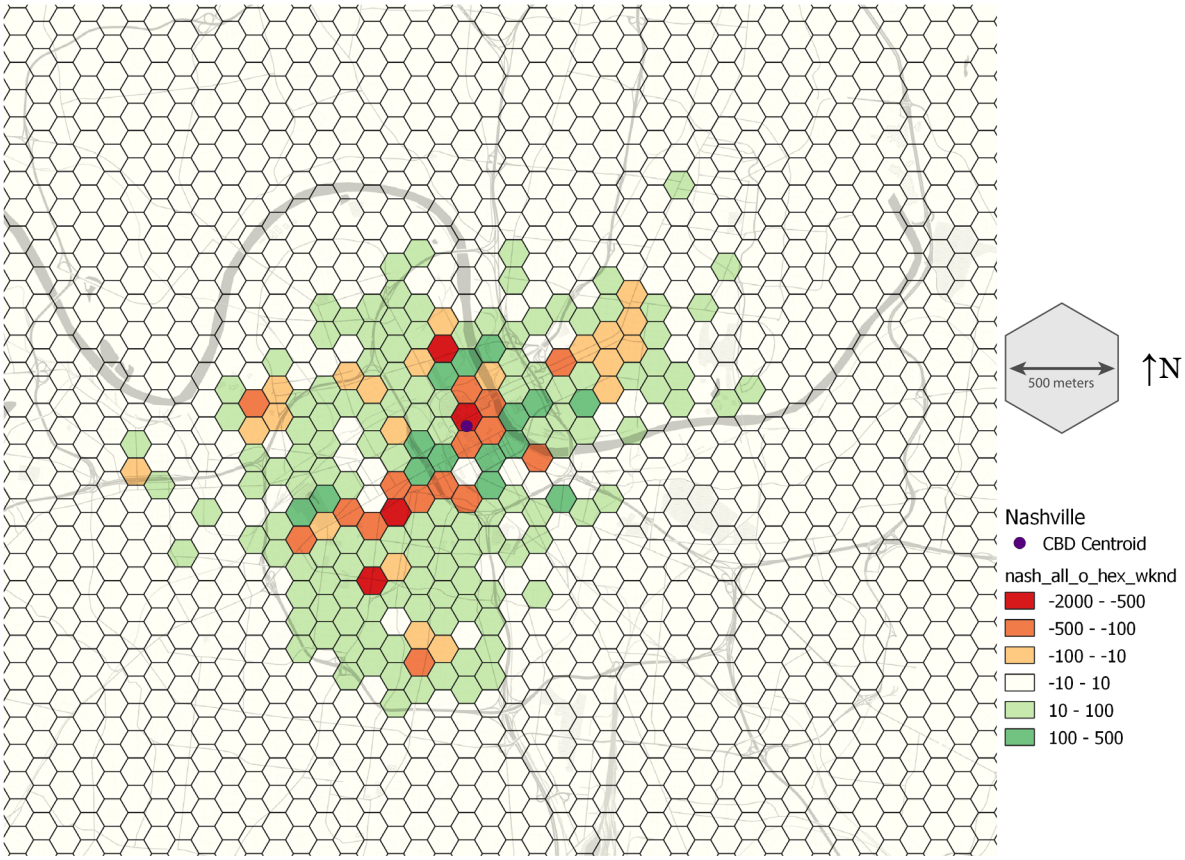

Figure 5.10– Nashville Weekends, Difference between Destination and Origin Counts per Hex

Portland

Portland	*Morning*	*Midday*	*PM Peak*	*Evening*	*Weekend*
	00:00-10:30	*10:31-15:30*	*15:31-18:30*	*18:31-23:59*	*00:00-23:59*
Trips	23,026	56,709	48,028	26,941	75,450

Table 5.2 – Portland Number of Trips by Time Bin

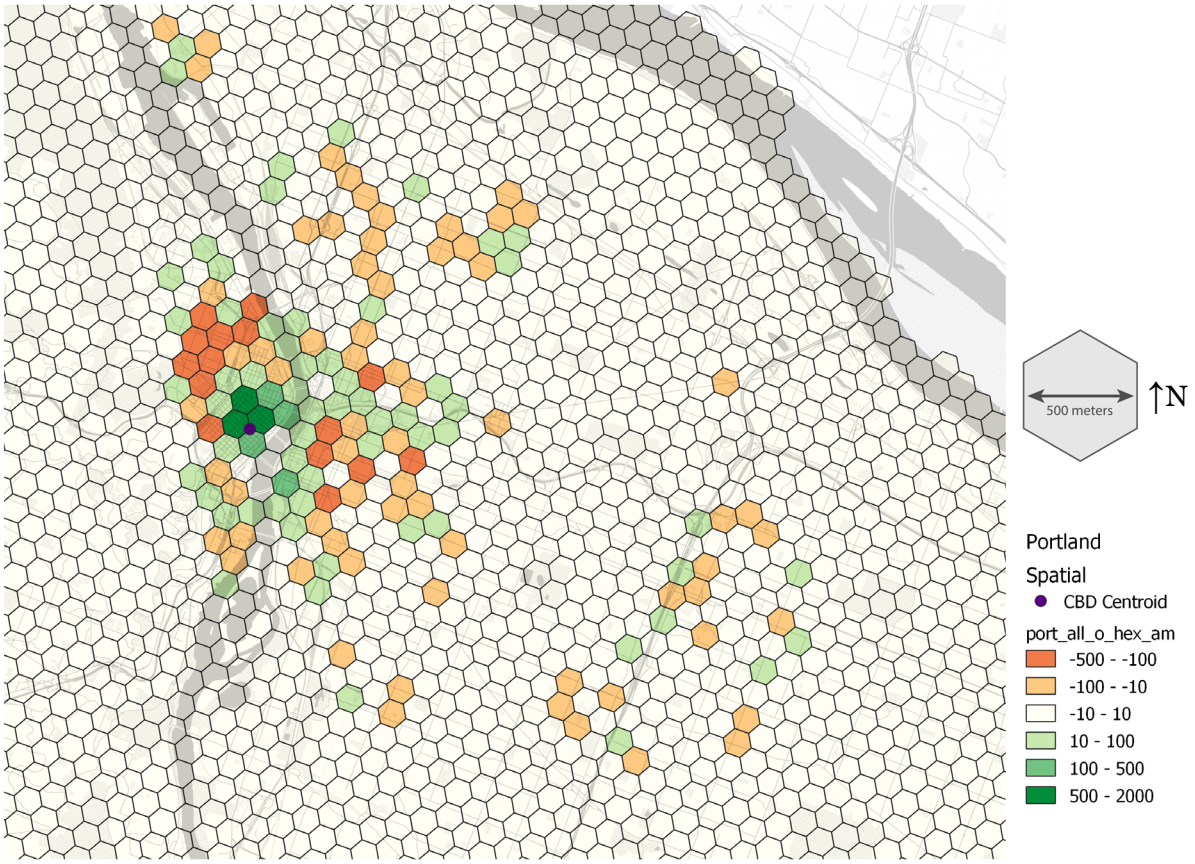

Figure 5.12 – Portland Morning, Difference between Destination and Origin Counts per Hex

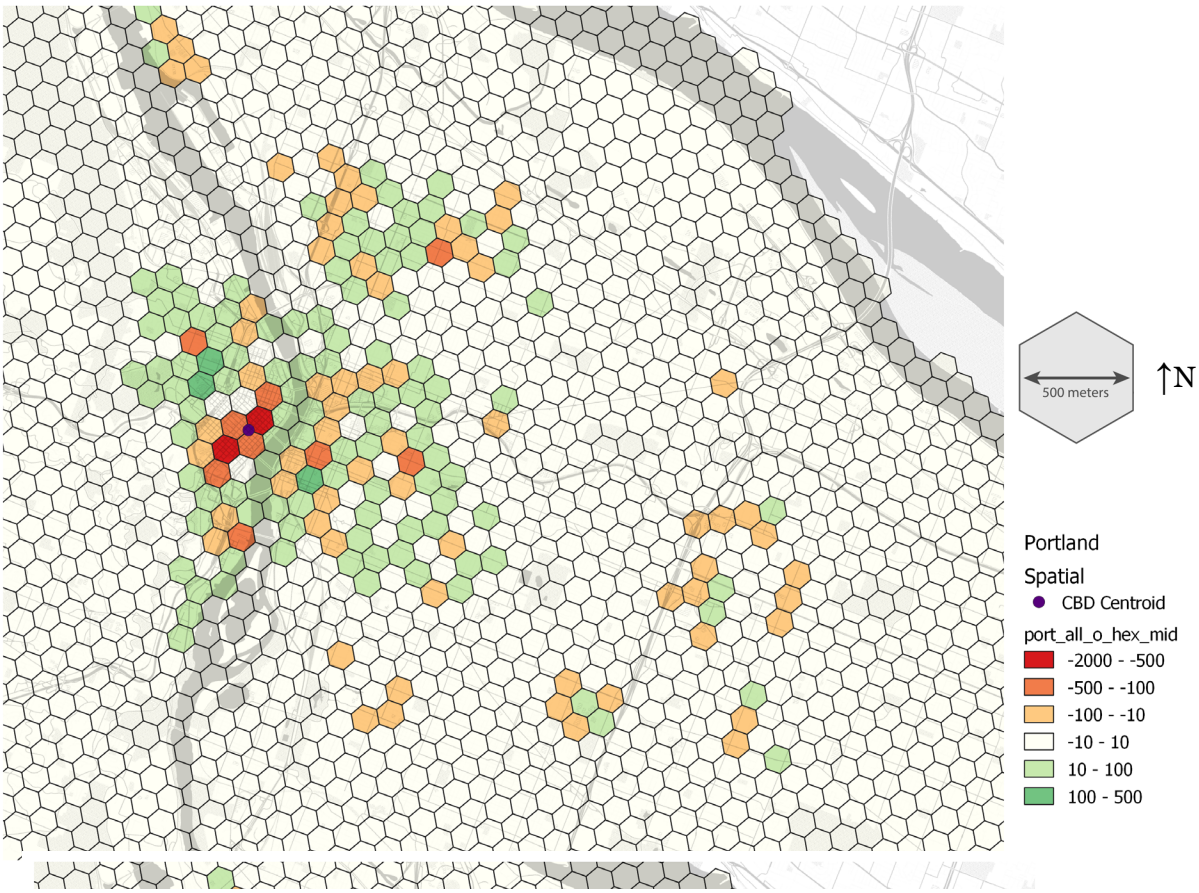

Figure 5.13 – Portland Midday, Difference between Destination and Origin Counts per Hex

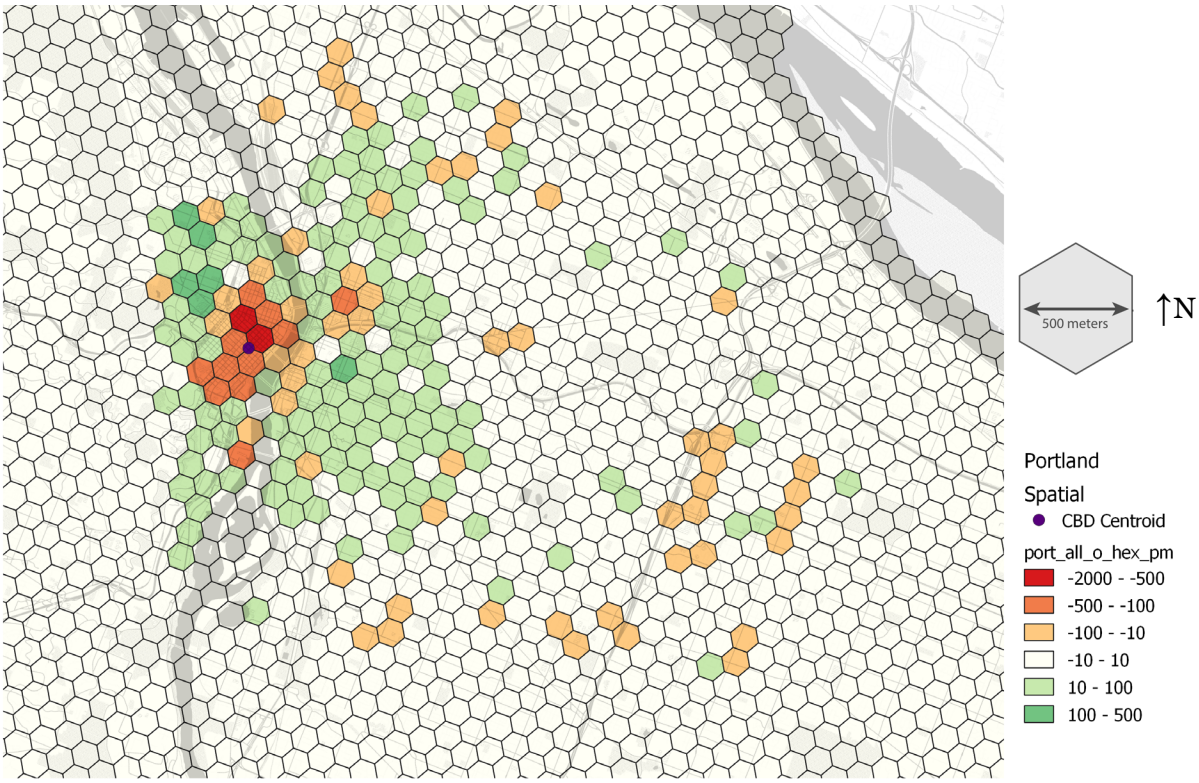

Figure 5.14 – Portland PM Peak, Difference between Destination and Origin Counts per Hex

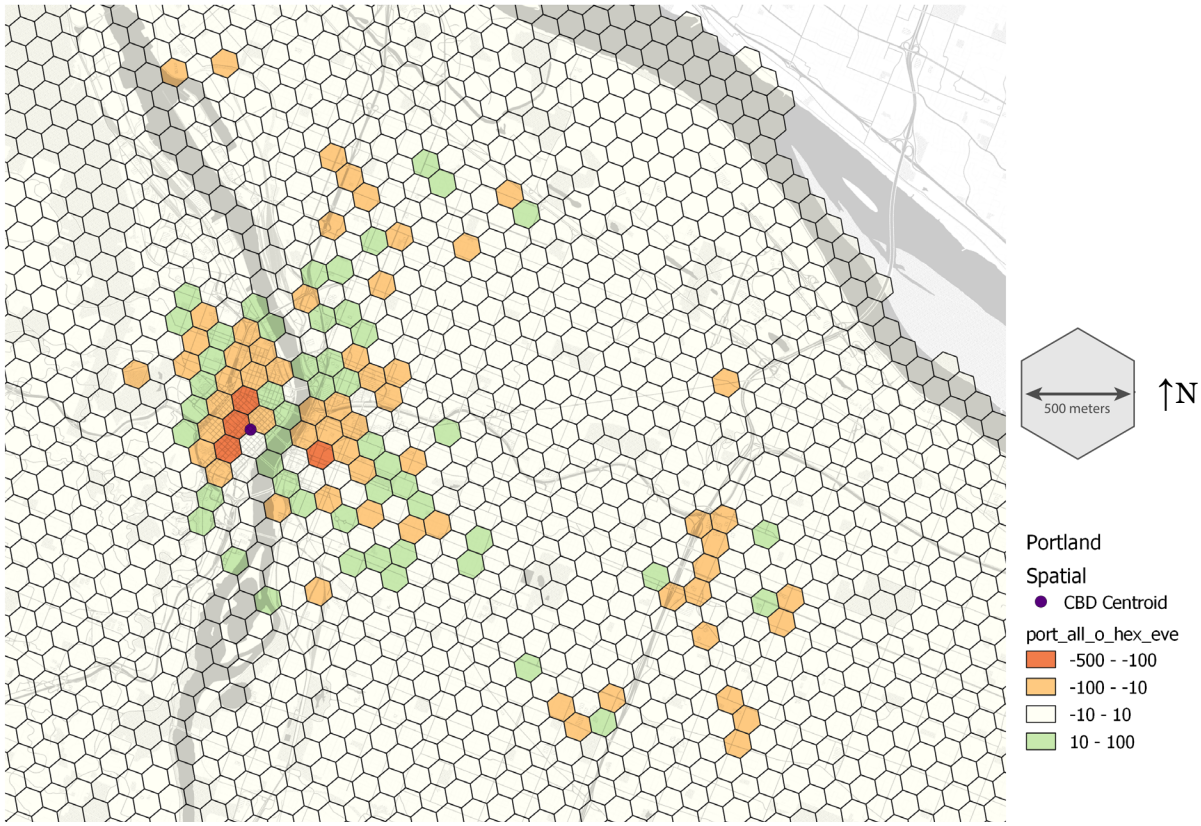

Figure 5.16 – Portland Evening, Difference between Destination and Origin Counts per Hex

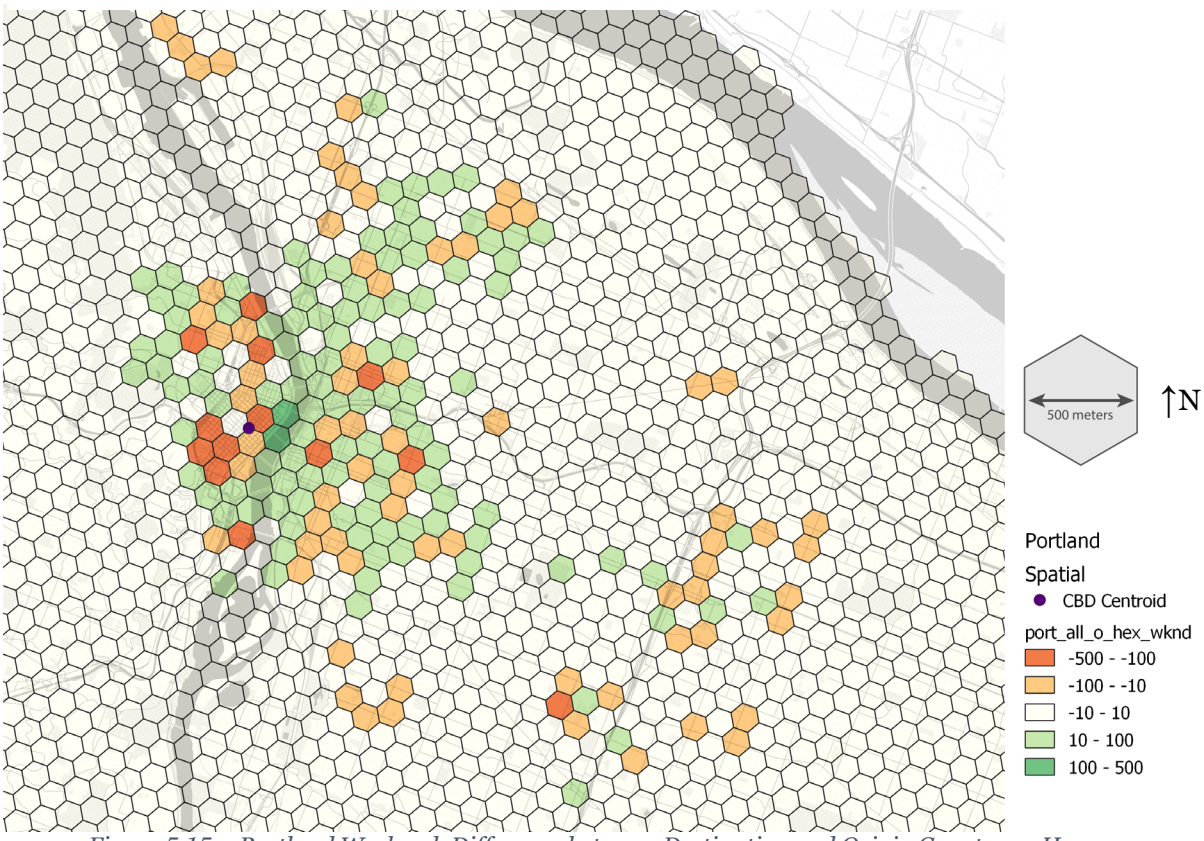

Figure 5.15 – Portland Weekend, Difference between Destination and Origin Counts per Hex

San Diego

San Diego	*Morning 00:00-10:30*	*Midday 10:31-15:30*	*PM Peak 15:31-18:30*	*Evening 18:31-23:59*	*Weekend 00:00-23:59*
Trips	150,697	318,005	228,485	135,163	388,522

Table 5.3 – San Diego Number of Trips by Time Bin

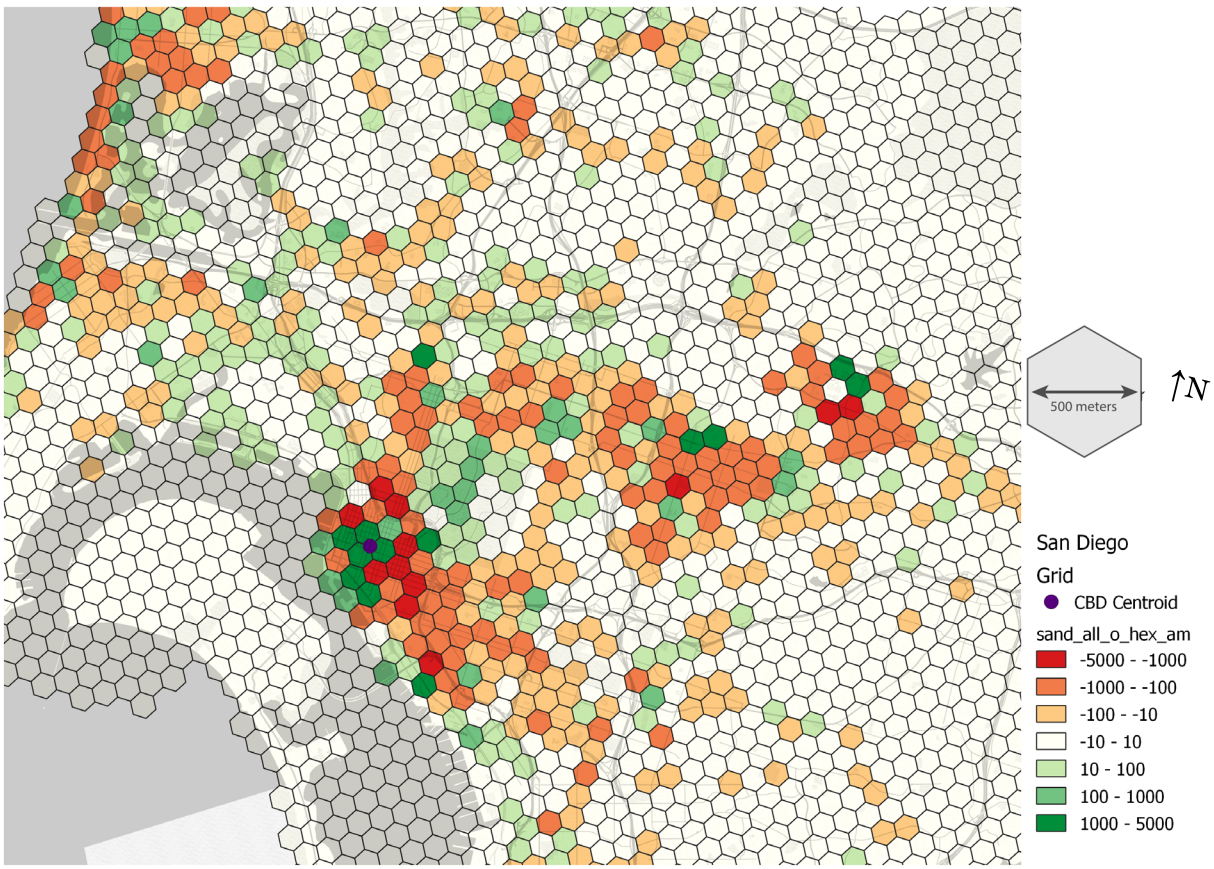

Figure 5.17 – San Diego Morning, Difference between Destination and Origin Counts per Hex

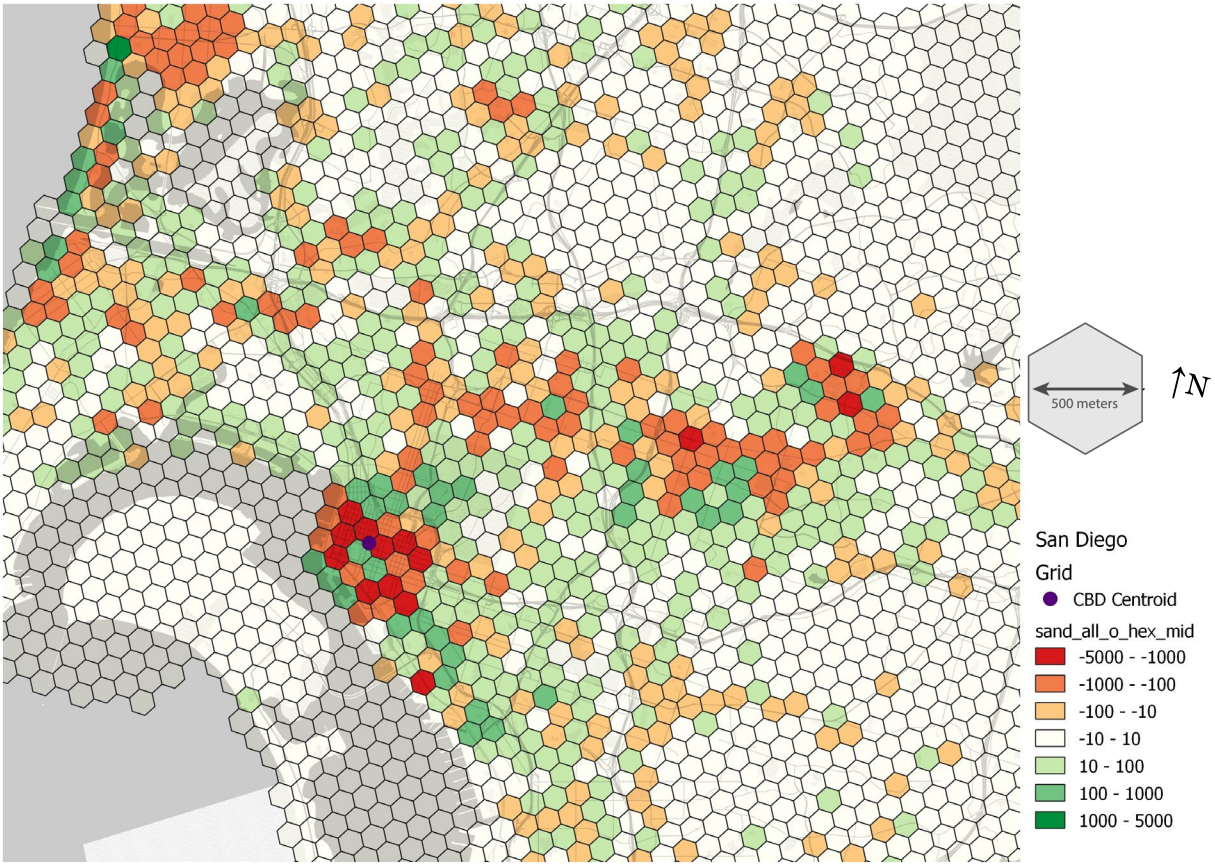

Figure 5.18 – San Diego Midday, Difference between Destination and Origin Counts per Hex

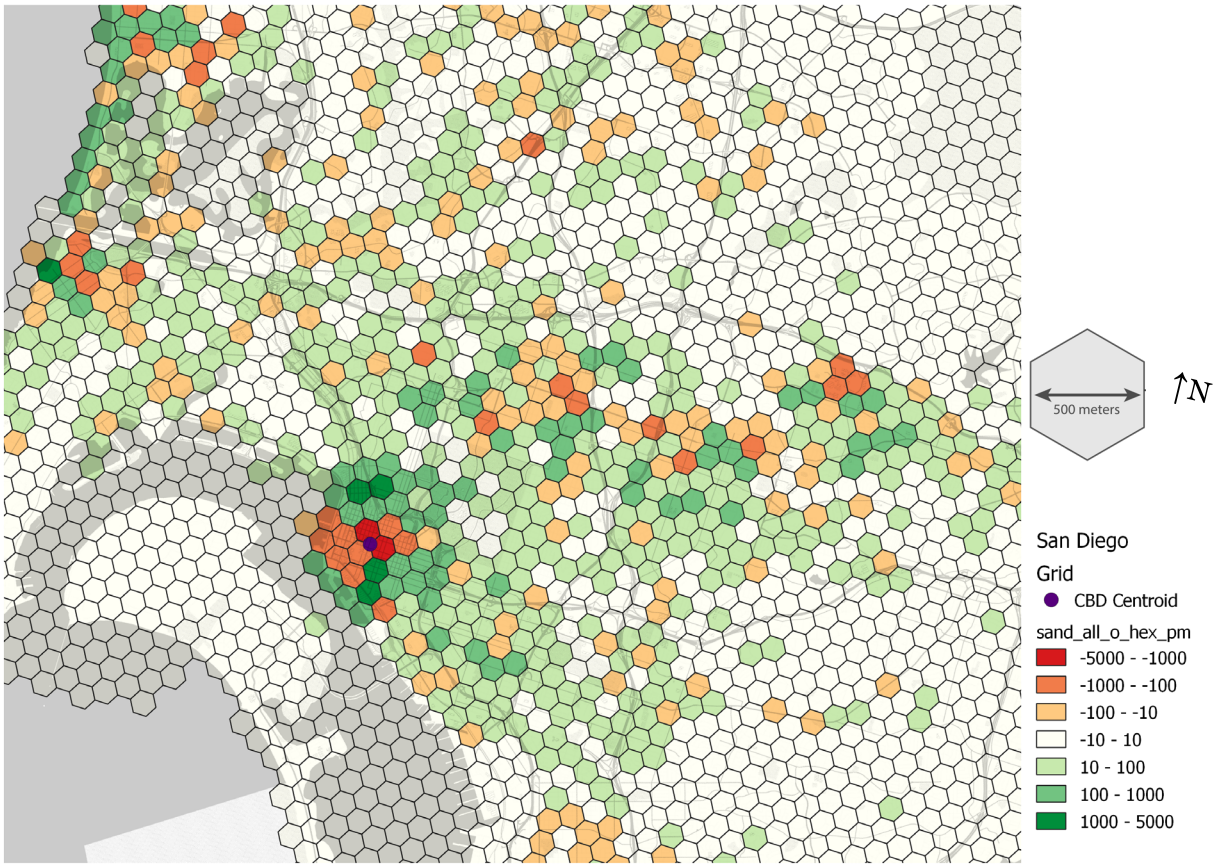

Figure 5.19 – San Diego PM Peak, Difference between Destination and Origin Counts per Hex

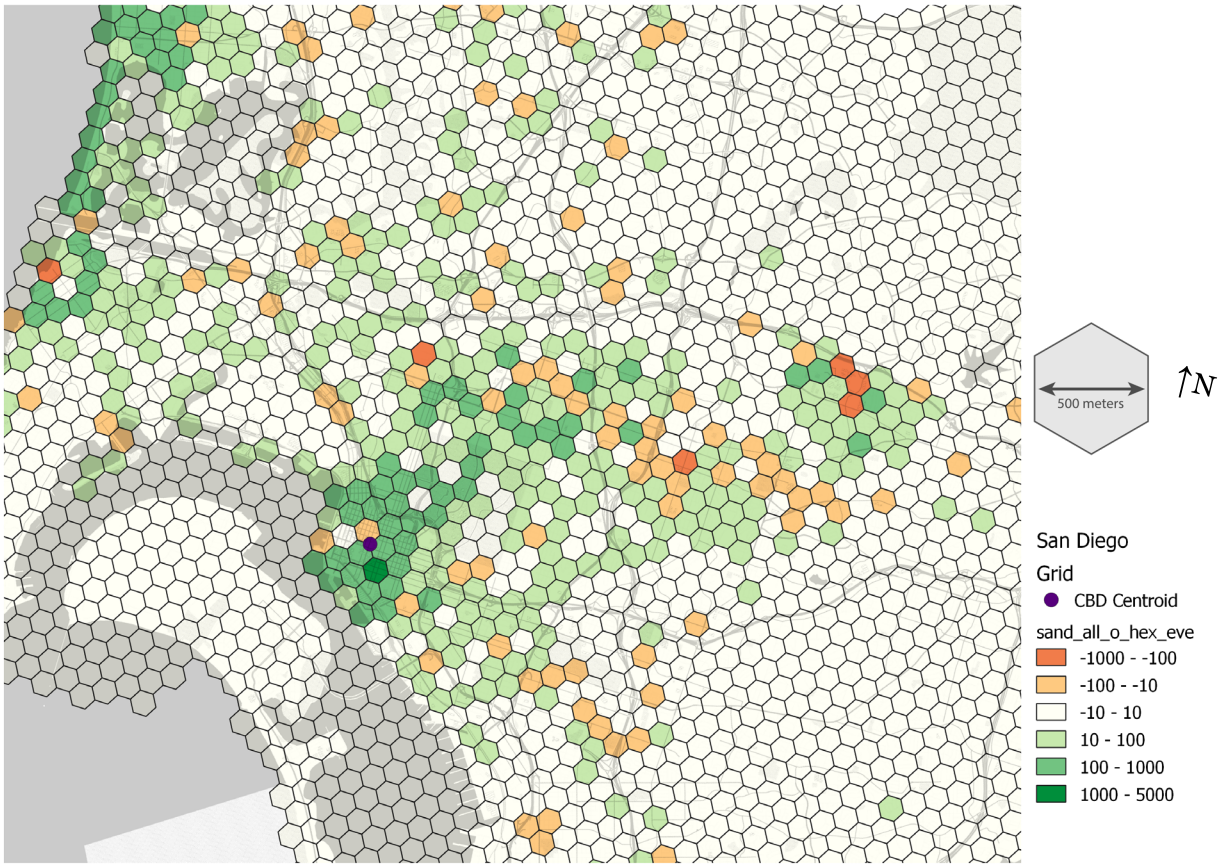

Figure 5.20 – San Diego Evening, Difference between Destination and Origin Counts per Hex

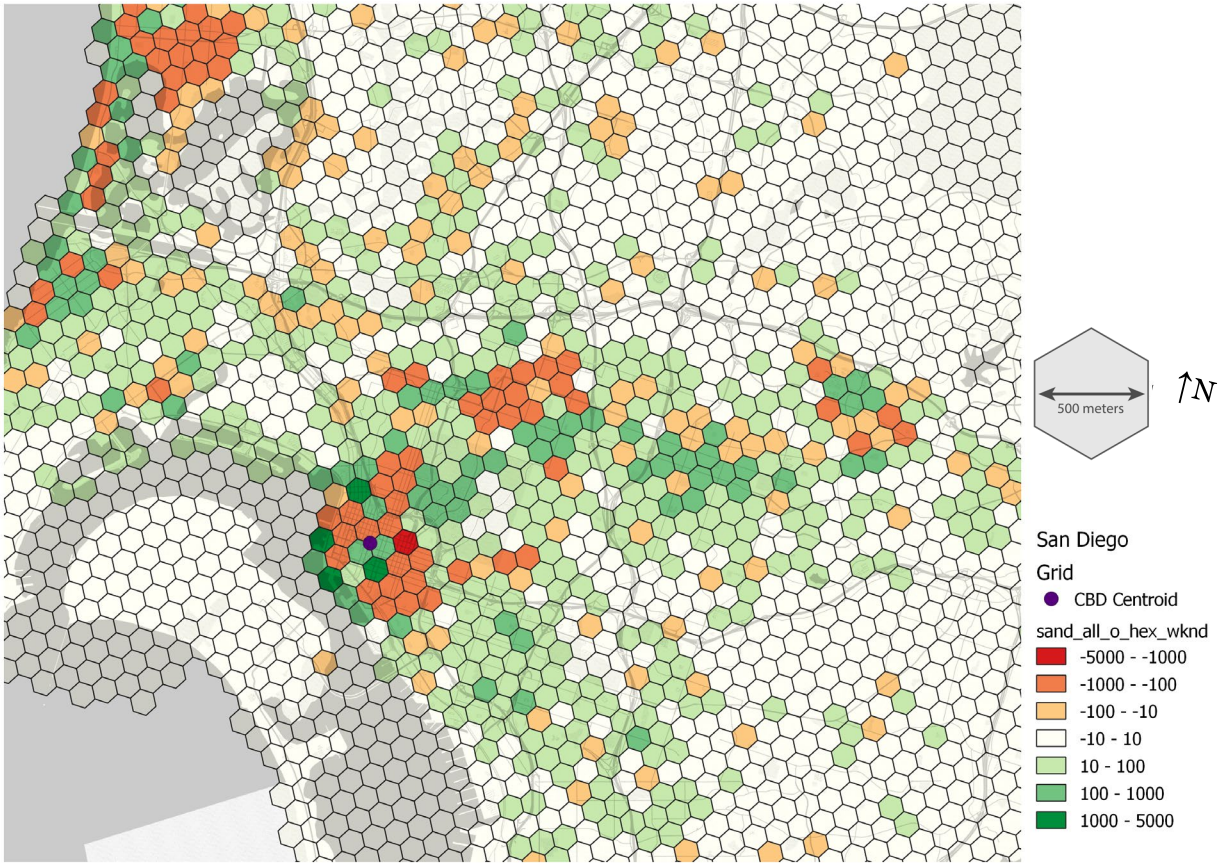

Figure 5.21 – San Diego Weekend, Difference between Destination and Origin Counts per Hex

Spatial Regression Tables

Nashville

Nashville Trip Count

	Dependent variable:					
	o_count_log					
	Overall	Morning	Midday	PM Peak	Evening	Weekend
	(1)	(2)	(3)	(4)	(5)	(6)
Job Density	-0.030***	0.049***	0.019***	0.010	0.008**	-0.017*
	(0.007)	(0.005)	(0.006)		(0.003)	(0.009)
Community of Concern Index	-0.005	-0.027***	-0.008	-0.008*	-0.041***	-0.047***
	(0.007)	(0.005)	(0.007)	(0.005)	(0.006)	(0.006)
Median Household Income	0.008***	-0.008***	0.001	0.002*	0.0001	0.007***
	(0.002)	(0.001)		(0.001)		(0.003)
Population Density	-0.001	0.002	-0.007***	-0.006***	-0.003***	-0.001
	(0.002)		(0.002)	(0.001)	(0.001)	
Street Length	-0.001	0.001	-0.001	-0.0003	-0.001	-0.003**
	(0.001)			(0.0002)	(0.001)	(0.001)
Rebalance Points	1.261***	0.827***	1.015***	0.969***	0.890***	1.104***
	(0.008)	(0.006)	(0.007)	(0.007)	(0.007)	(0.008)
Distance to CBD	-0.185***	0.165***	0.032***	0.0001	0.015**	-0.097***
	(0.014)	(0.007)	(0.011)		(0.007)	(0.016)
Near Transit	0.124***	-0.086	0.007	-0.003	-0.015**	0.018
	(0.009)				(0.007)	(0.011)
Rho	0.3523	0.4686	0.3398	0.3579	0.3913	0.3767
Rho Std. Err.	0.0079	0.0087	0.0083	0.0077	0.0094	0.0086
Akaike Inf. Crit. (LM)	3384	-122	151	-409	-77	1999
Akaike Inf. Crit.	1,709.836	-2,257.851	-1,261.144	-1,998.664	-1,593.379	291.547
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 5.4 – Nashville Count Regression

Nashville Rebalance Points

	Dependent variable:
	rebal_log
	Overall
Job Density	0.059***
	(0.011)
Community of Concern Index	-0.177***
	(0.009)
Household Median Income	0.048***
	(0.003)
Population Density	-0.103***
	(0.004)
Street Length	-0.008***
	(0.001)
Distance to CBD	-0.577***
	(0.020)
Near Transit	1.067***
	(0.012)
Rho	0.8413
Rho Std. Err.	0.0067
Akaike Inf. Crit. (LM)	12700.4
Akaike Inf. Crit.	6,451.212
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 5.5 – Nashville Rebalance Regression

Nashville Mean Cost

	Dependent variable:					
	o_cost_mean_log					
	Overall	Morning	Midday	PM Peak	Evening	Weekend
	(1)	(2)	(3)	(4)	(5)	(6)
Job Density	-0.165***	-0.046**	-0.144***	-0.149***	-0.174***	-0.171***
	(0.034)	(0.021)	(0.025)	(0.025)	(0.023)	(0.028)
Community of Concern Index	0.156***	0.162***	0.148***	0.215***	0.153***	0.075***
	(0.031)	(0.020)	(0.024)	(0.025)	(0.022)	(0.025)
Median Household Income	0.021**	0.019***	0.021**	0.027***	0.027***	0.031***
	(0.010)	(0.007)	(0.008)	(0.009)	(0.008)	(0.008)
Population Density	0.056***	-0.025***	-0.020**	-0.022***	-0.015*	0.022*
	(0.015)	(0.009)	(0.009)	(0.005)	(0.009)	(0.013)
Street Length	0.010	-0.006*	-0.0004	0.004	-0.002	-0.011**
	(0.007)	(0.003)	(0.001)		(0.002)	(0.005)
Rebalance Points	1.032***	0.958***	1.049***	1.099***	1.066***	1.088***
	(0.024)	(0.018)	(0.020)	(0.021)	(0.019)	(0.021)
Distance to CBD	-0.727***	-0.207***	-0.467***	-0.476***	-0.482***	-0.651***
	(0.065)	(0.037)	(0.046)	(0.046)	(0.042)	(0.052)
Near Transit	0.696***	0.197***	0.504***	0.384***	0.308***	0.453***
	(0.044)	(0.027)	(0.035)	(0.035)	(0.032)	(0.037)
Rho	0.4508	0.2561	0.4027	0.403	0.4263	0.4563
Rho Std. Err.	0.0143	0.016	0.0143	0.0142	0.014	0.0135
Akaike Inf. Crit. (LM)	21946	15211	18444	18849	17702	19732
Akaike Inf. Crit.	21,107.380	14,970.060	17,750.770	18,152.120	16,915.000	18,783.670
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 5.6 – Nashville Cost Regression

Nashville Mean Distance

	Dependent variable:					
	o_dist_mean					
	Overall	Morning	Midday	PM Peak	Evening	Weekend
	(1)	(2)	(3)	(4)	(5)	(6)
Job Density	-154.694***	-37.219*	-124.102***	-122.338***	-94.453***	-157.721***
	(33.159)	(21.819)	(22.937)	(26.633)	(21.738)	(26.489)
Community of Concern Index	150.610***	106.778***	80.213***	140.672***	89.245***	101.504***
	(29.807)	(17.260)	(19.881)	(21.044)	(19.077)	(27.552)
Median Household Income	26.275**	16.759**	11.693	20.421***	22.287***	29.237***
	(10.936)	(6.830)	(8.193)	(7.626)	(6.487)	(9.309)
Population Density	18.764	-13.143	-16.291*	-20.562*	-17.288*	3.387
			(9.068)	(10.702)	(10.099)	
Street Length	-8.791	-3.415	-3.973	-1.783	-4.503	-17.322***
			(7.553)	(2.558)	(4.555)	(4.878)
Rebalance Points	345.007***	405.135***	440.411***	436.330***	393.514***	426.568***
	(24.547)	(15.840)	(15.762)	(15.999)	(15.139)	(21.516)
Distance to CBD	-627.519***	-158.803***	-363.383***	-378.400***	-297.258***	-539.455***
	(64.751)	(25.408)	(44.563)	(49.706)	(41.623)	(52.695)
Near Transit	452.087***	117.017***	318.630***	244.807***	241.336***	359.508***
	(43.926)	(13.269)	(28.985)	(28.722)	(26.289)	(38.042)
Rho	0.2187	0.1734	0.2898	0.2796	0.2771	0.2272
Rho Std. Err.	0.0183	0.0188	0.0169	0.0172	0.0174	0.0179
Akaike Inf. Crit. (LM)	110713	104352	104523	104872	103958	108912
Akaike Inf. Crit.	110,579.600	104,272.700	104,247.900	104,621.900	103,721.800	108,760.300
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 5.7 – Nashville Distance Regression

Portland

Portland Trip Count

	Dependent variable:					
	o_count_log					
	Overall	Morning	Midday	PM Peak	Evening	Weekend
	(1)	(2)	(3)	(4)	(5)	(6)
Job Density	0.259***	0.0005	0.134***	0.161***	0.142***	0.149***
	(0.024)	(0.018)	(0.020)	(0.020)	(0.019)	(0.021)
Community of Concern Index	-0.050***	0.015	-0.028**	-0.039***	-0.052***	-0.054***
	(0.014)	(0.010)	(0.011)	(0.011)	(0.011)	(0.012)
Median Household Income	0.026**	-0.0003	0.018*	0.021**	0.034***	0.029***
	(0.012)	(0.009)	(0.010)	(0.009)	(0.009)	(0.010)
Population Density	0.119***	-0.041***	-0.005	0.031***	0.048***	0.056***
	(0.012)	(0.009)	(0.010)	(0.010)	(0.009)	(0.010)
Street Length	0.051***	-0.002	0.011*	0.012**	0.012**	0.025***
	(0.007)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)
Rebalance Points	1.266***	0.837***	0.964***	0.952***	0.896***	1.063***
	(0.014)	(0.011)	(0.012)	(0.011)	(0.011)	(0.012)
Distance to CBD	0.236***	-0.151***	-0.092***	-0.044	0.048*	0.018
	(0.034)	(0.027)	(0.030)	(0.029)	(0.027)	(0.031)
Near Transit	0.301***	-0.111***	0.050**	0.094***	0.095***	0.099***
	(0.024)	(0.018)	(0.020)	(0.020)	(0.019)	(0.021)
Rho	0.5097	0.3149	0.3476	0.4149	0.4643	0.4435
Rho Std. Err.	0.0148	0.0171	0.0165	0.0163	0.0163	0.0158
Akaike Inf. Crit. (LM)	4455	2499	3087	3208	3138	3623
Akaike Inf. Crit.	3,463.693	2,184.427	2,679.548	2,621.443	2,439.006	2,934.799
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 5.8 – Portland Trip Count Regressions

Portland Rebalance Points

	Dependent variable:
	rebal_log
	Overall
Job Density	0.695***
	(0.037)
Community of Concern Index	0.036*
	(0.021)
Household Median Income	0.037**
	(0.018)
Population Density	0.151***
	(0.018)
Street Length	-0.007
	(0.011)
Distance to CBD	-0.363***
	(0.051)
Near Transit	0.440***
	(0.037)
Rho	0.691
Rho Std. Err.	0.0198
Akaike Inf. Crit. (LM)	6238.6
Akaike Inf. Crit.	5,376.653
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 5.9 – Portland Rebalance Regression

Portland Mean Cost

	Dependent variable:					
	o_cost_mean_log					
	Overall	Morning	Midday	PM Peak	Evening	Weekend
	(1)	(2)	(3)	(4)	(5)	(6)
Job Density	0.553***	0.242***	0.423***	0.514***	0.417***	0.537***
	(0.073)	(0.067)	(0.075)	(0.072)	(0.072)	(0.077)
Community of Concern Index	0.121***	-0.040	-0.054	-0.020	-0.026	0.042
	(0.041)	(0.038)	(0.042)	(0.041)	(0.040)	(0.044)
Median Household Income	-0.150***	-0.055*	-0.035	-0.033	0.064*	0.013
	(0.036)	(0.032)	(0.036)	(0.035)	(0.035)	(0.037)
Population Density	0.550***	0.147***	0.281***	0.382***	0.371***	0.474***
	(0.038)	(0.033)	(0.037)	(0.036)	(0.036)	(0.039)
Street Length	0.187***	0.010	0.083***	0.080***	0.066***	0.115***
	(0.022)	(0.020)	(0.022)	(0.021)	(0.021)	(0.023)
Rebalance Points	0.438***	1.062***	0.897***	0.812***	0.845***	0.721***
	(0.034)	(0.034)	(0.036)	(0.034)	(0.034)	(0.036)
Distance to CBD	0.686***	0.292***	0.569***	0.713***	0.679***	0.641***
	(0.101)	(0.093)	(0.104)	(0.099)	(0.099)	(0.107)
Near Transit	1.069***	0.218***	0.781***	0.922***	0.944***	0.812***
	(0.074)	(0.067)	(0.076)	(0.073)	(0.073)	(0.078)
Rho	0.5603	0.3143	0.4775	0.5308	0.5235	0.5022
Rho Std. Err.	0.0236	0.0276	0.0254	0.024	0.0241	0.0251
Akaike Inf. Crit. (LM)	8661	7809	8557	8525	8491	8721
Akaike Inf. Crit.	8,123.375	7,680.132	8,217.800	8,057.278	8,039.457	8,330.218
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 5.10 – Portland Cost Regressions

Portland Mean Distance

	Dependent variable:					
	o_dist_mean					
	Overall	Morning	Midday	PM Peak	Evening	Weekend
	(1)	(2)	(3)	(4)	(5)	(6)
Job Density	256.673***	182.182***	257.200***	355.285***	187.282***	324.138***
	(69.472)	(60.953)	(74.322)	(69.248)	(63.246)	(76.235)
Community of Concern Index	112.805***	-74.874**	-19.244	-19.824	19.907	74.598*
	(39.020)	(33.099)	(99.330)	(37.563)	(56.393)	(42.487)
Median Household Income	-103.598***	-37.793	-18.748	-32.962	-10.977	-0.581
	(33.050)	(29.157)	(51.943)	(38.039)	(39.104)	
Population Density	335.024***	117.992***	214.029***	202.758***	226.804***	297.281***
	(33.798)	(27.274)	(37.498)	(36.423)	(34.258)	(34.257)
Street Length	99.135***	8.144	42.716**	53.630**	18.489	71.308***
	(20.105)		(19.671)	(22.017)	(21.405)	(22.095)
Rebalance Points	183.681***	546.386***	452.296***	352.121***	338.193***	353.717***
	(30.931)	(27.573)	(30.505)	(32.620)	(31.273)	(34.329)
Distance to CBD	435.442***	288.432***	455.528***	475.002***	387.624***	489.078***
	(95.933)	(82.670)	(116.991)	(95.102)	(92.358)	(104.925)
Near Transit	663.014***	158.006***	460.939***	636.396***	682.894***	543.450***
	(68.633)	(52.809)	(68.183)	(72.058)	(66.457)	(73.762)
Rho	0.4332	0.24	0.3532	0.3798	0.3469	0.3585
Rho Std. Err.	0.0257	0.0294	0.0276	0.0273	0.0275	0.0274
Akaike Inf. Crit. (LM)	36739	35886	36457	36841	36571	37026
Akaike Inf. Crit.	36,497.480	35,824.320	36,308.390	36,671.220	36,429.400	36,876.400
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 5.11 – Portland Distance Regressions

San Diego

San Diego Trip Count

	Dependent variable:					
	o_count_log					
	Overall	Morning	Midday	PM Peak	Evening	Weekend
	(1)	(2)	(3)	(4)	(5)	(6)
Job Density	0.029***	-0.034***	0.003**	-0.018***	-0.014***	-0.027***
	(0.011)	(0.006)	(0.001)	(0.005)	(0.005)	(0.006)
Community of Concern Index	0.020	0.095***	0.014	0.031**	0.011	-0.070***
	(0.025)	(0.016)	(0.013)	(0.013)	(0.011)	(0.014)
Median Household Income	0.002	-0.008	-0.001	-0.005	-0.012**	-0.001
	(0.059)	(0.010)	(0.002)	(0.005)	(0.005)	(0.003)
Population Density	0.001	-0.063***	-0.072***	-0.037***	-0.033***	-0.024***
	(0.010)	(0.009)	(0.007)	(0.007)	(0.007)	(0.008)
Street Length	0.014	-0.034***	-0.021***	-0.021***	-0.030***	-0.017***
	(0.009)	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)
Rebalance Points	1.134***	0.876***	0.981***	0.925***	0.817***	0.991***
	(0.010)	(0.009)	(0.009)	(0.008)	(0.008)	(0.009)
Distance to CBD	0.070*	0.227***	0.156***	0.130***	0.175***	0.055**
	(0.036)	(0.021)	(0.018)	(0.017)	(0.017)	(0.022)
Near Transit	0.021	0.051***	0.052***	-0.002	-0.017	-0.091***
	(0.055)	(0.019)	(0.016)		(0.013)	(0.017)
Rho	0.0662	0.2614	0.1414	0.1639	0.2737	0.2214
Rho Std. Err.	0.011	0.0117	0.0099	0.0099	0.0113	0.0101
Akaike Inf. Crit. (LM)	6433	10246	8661	8035	9064	9098
Akaike Inf. Crit.	6,351.210	9,776.813	8,465.124	7,774.821	8,525.667	8,630.499
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 5.12 – San Diego Trip Count Regressions

San Diego Rebalance Points

	Dependent variable:
	rebal_log
	Overall
Job Density	0.252***
	(0.008)
Community of Concern Index	-0.359***
	(0.021)
Household Median Income	0.060***
	(0.008)
Population Density	0.083***
	(0.012)
Street Length	0.139***
	(0.006)
Distance to CBD	-1.143***
	(0.027)
Near Transit	1.919***
	(0.025)
Rho	0.7782
Rho Std. Err.	0.0084
Akaike Inf. Crit. (LM)	18128.9
Akaike Inf. Crit.	13,960.620
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 5.13 – San Diego Rebalance Regression

San Diego Mean Cost

	Dependent variable:					
	o_cost_mean_log					
	Overall	Morning	Midday	PM Peak	Evening	Weekend
	(1)	(2)	(3)	(4)	(5)	(6)
Job Density	0.161***	-0.004	0.040**	0.040**	-0.0004	0.021
	(0.019)		(0.020)	(0.017)		(0.021)
Community of Concern Index	0.213***	0.270***	0.163***	0.228***	0.281***	0.220***
	(0.049)	(0.041)	(0.043)	(0.045)	(0.038)	(0.046)
Median Household Income	0.026	-0.002	0.024	0.019	-0.005	-0.014
	(0.023)	(0.006)		(0.017)		(0.020)
Population Density	0.188***	-0.028	0.028	0.024	0.047***	0.087***
	(0.029)	(0.022)		(0.021)	(0.012)	(0.030)
Street Length	0.255***	0.035***	0.115***	0.108***	0.043***	0.121***
	(0.012)	(0.009)	(0.010)	(0.011)	(0.011)	(0.013)
Rebalance Points	0.850***	1.040***	1.025***	1.059***	1.062***	1.072***
	(0.020)	(0.019)	(0.020)	(0.020)	(0.019)	(0.020)
Distance to CBD	0.354***	0.208***	0.197***	0.293***	0.291***	0.332***
	(0.059)	(0.053)	(0.060)	(0.058)	(0.053)	(0.065)
Near Transit	0.094	0.137***	0.169***	0.111*	0.111***	0.007
	(0.065)	(0.052)	(0.053)	(0.058)	(0.033)	
Rho	0.4329	0.2021	0.2906	0.3159	0.2287	0.3523
Rho Std. Err.	0.015	0.0168	0.016	0.016	0.0163	0.0156
Akaike Inf. Crit. (LM)	22548	20234	21610	21709	20709	21943
Akaike Inf. Crit.	21,837.920	20,096.610	21,312.440	21,355.170	20,528.250	21,490.990
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 5.14 – San Diego Cost Regressions

San Diego Mean Distance

	Dependent variable:					
	o_dist_mean					
	Overall	Morning	Midday	PM Peak	Evening	Weekend
	(1)	(2)	(3)	(4)	(5)	(6)
Job Density	58.311***	-17.630**	-5.381	49.223***	7.623	-4.436
	(16.934)	(7.755)	(4.326)	(17.634)	(39.135)	
Community of Concern Index	203.068***	207.688***	234.087***	179.835***	269.527***	128.761***
	(44.918)	(31.514)	(41.219)	(38.162)	(32.963)	(43.158)
Median Household Income	6.197	-1.615*	6.363***	7.110	-7.638	-16.710
		(0.838)	(2.450)		(29.626)	(11.271)
Population Density	69.158***	-17.330*	4.845	3.493	-33.354*	56.340**
	(26.672)	(8.961)		(4.219)	(18.721)	(24.676)
Street Length	131.481***	19.678***	64.034***	49.735***	29.493***	76.295***
	(12.063)	(7.047)	(9.453)	(9.853)	(8.380)	(11.861)
Rebalance Points	275.922***	415.538***	394.353***	387.232***	399.067***	409.753***
	(14.342)	(12.466)	(15.260)	(14.270)	(18.452)	(14.854)
Distance to CBD	-17.467	20.217***	32.853	-86.861	114.197	48.703
		(7.467)	(24.303)	(56.147)	(76.574)	(31.249)
Near Transit	48.731***	114.815***	157.554***	-2.529	80.167	6.228
	(13.286)	(33.601)	(48.052)	(3.000)	(72.510)	(7.230)
Rho	0.251	0.1933	0.1593	0.2201	0.1709	0.1903
Rho Std. Err.	0.0189	0.0192	0.0195	0.0194	0.0207	0.0196
Akaike Inf. Crit. (LM)	94968	91009	93891	93516	90837	95225
Akaike Inf. Crit.	94,806.520	90,916.270	93,830.050	93,398.110	90,766.570	95,138.310
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 5.15 – San Diego Distance Regressions

Weather

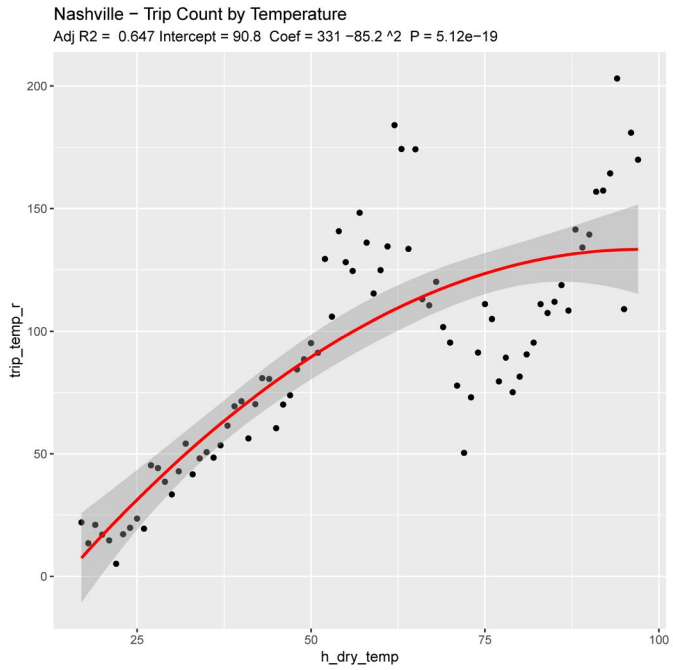

Figure 5.22 – Nashville Count by Temperature

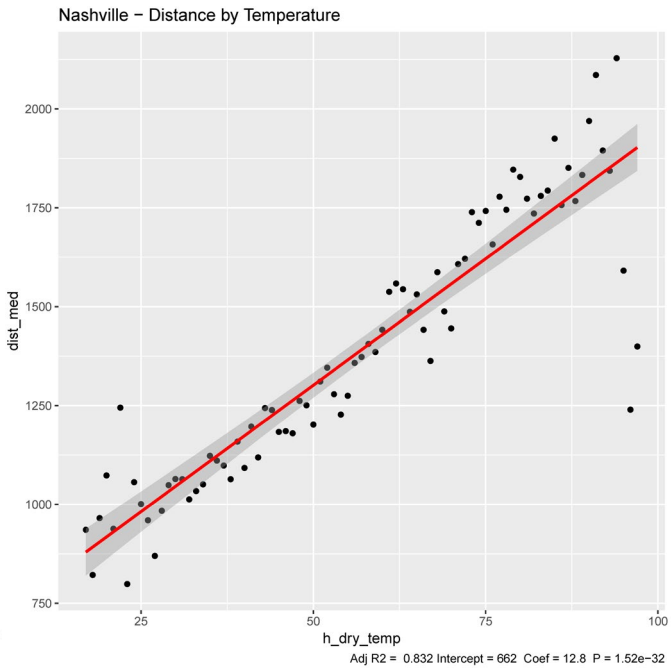

Figure 5.23 – Nashville Distance by Temperature

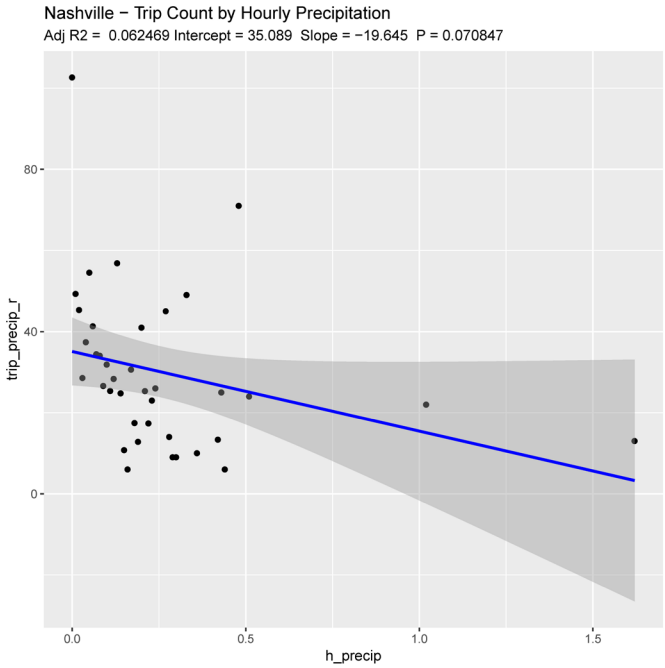

Figure 5.24 – Nashville Count by Precipitation

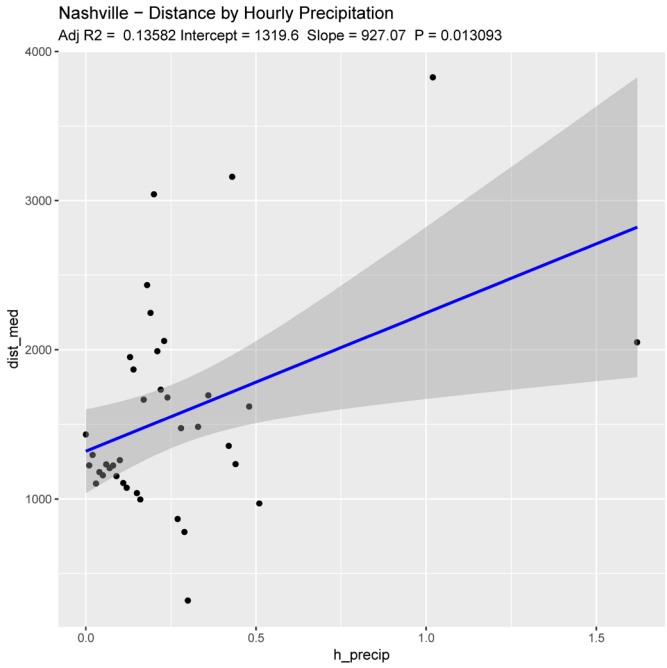

Figure 5.25 – Nashville Distance by Precipitation

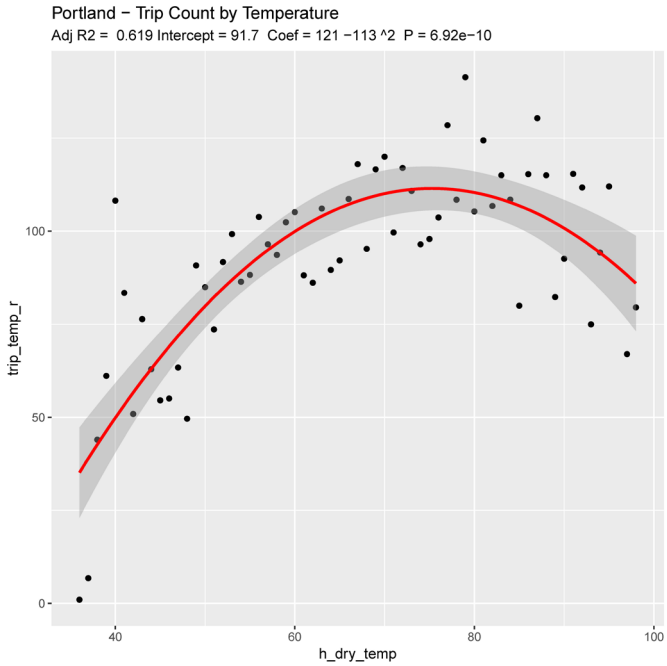

Figure 5.29 – Portland Count by Temperature

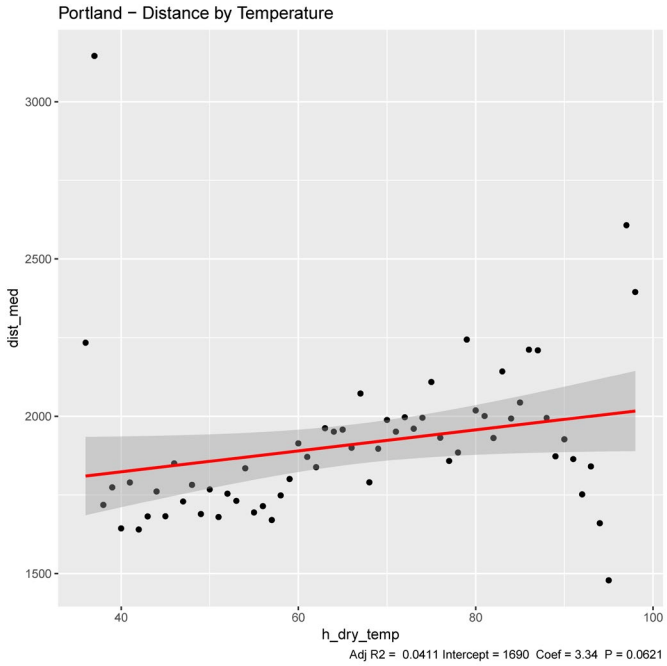

Figure 5.28 – Portland Distance by Temperature

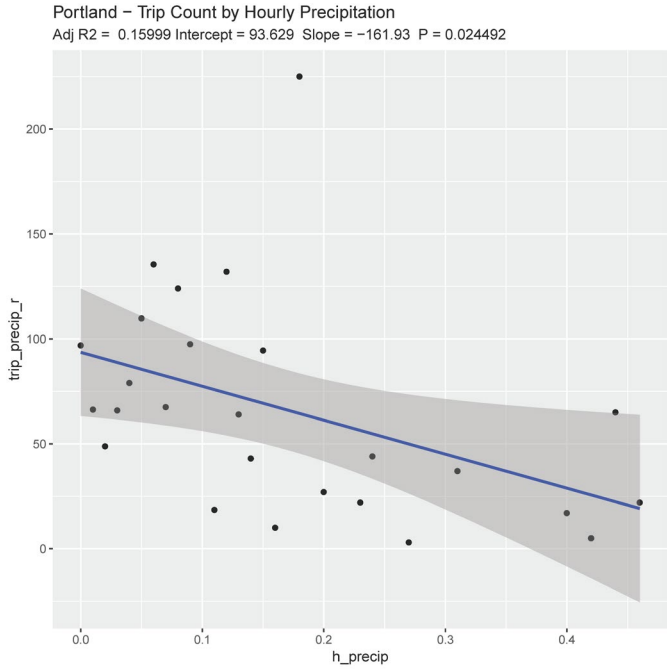

Figure 5.27 – Portland Count by Precipitation

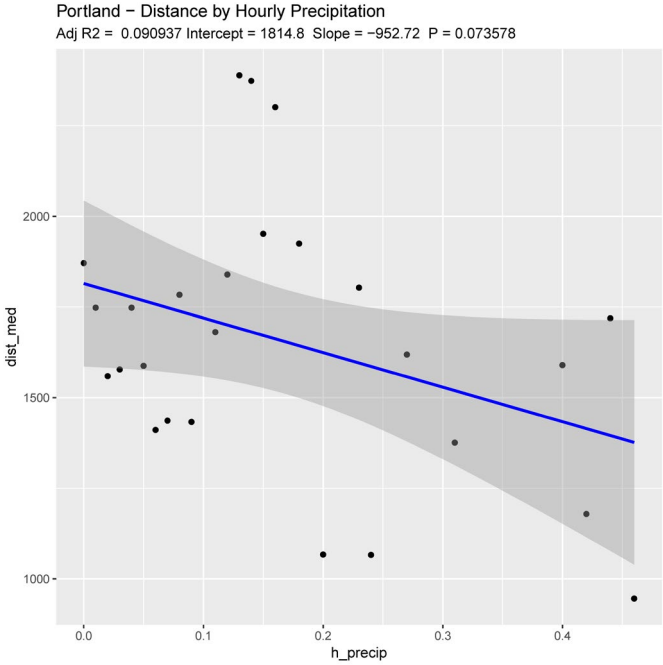

Figure 5.26 – Portland Distance by Precipitation

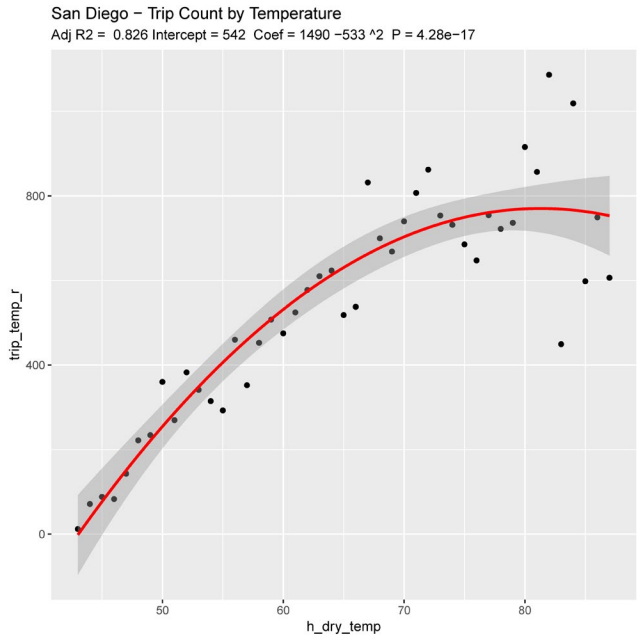

Figure 5.30 – San Diego Count by Temperature

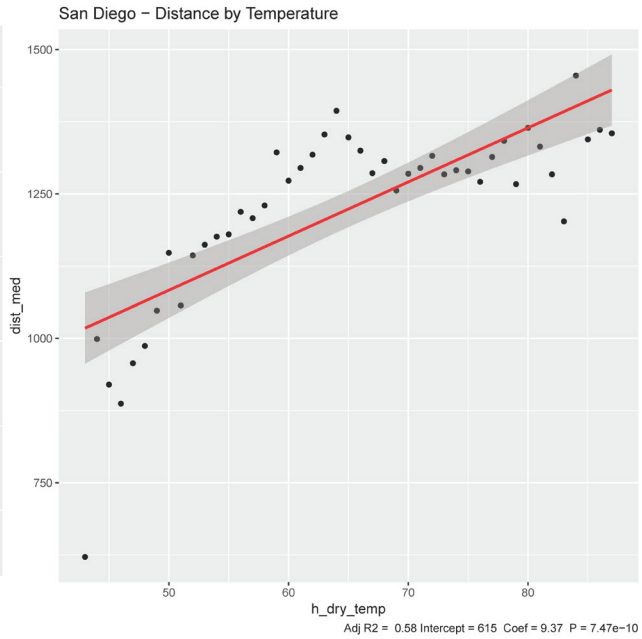

Figure 5.33 – San Diego Distance by Temperature

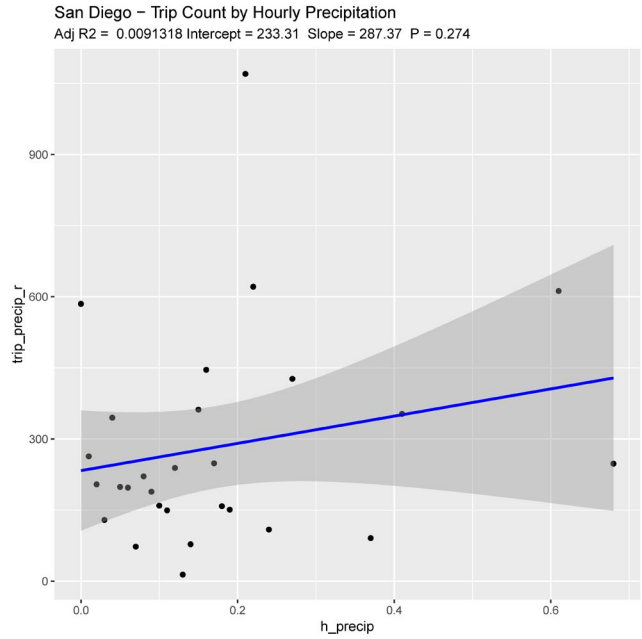

Figure 5.32 – San Diego Count by Precipitation

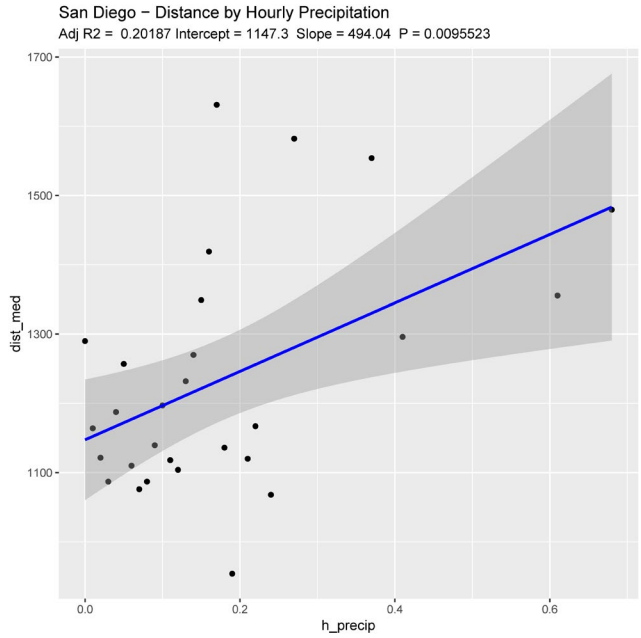

Figure 5.31 – San Diego Distance by Precipitation

Open Trip Planner

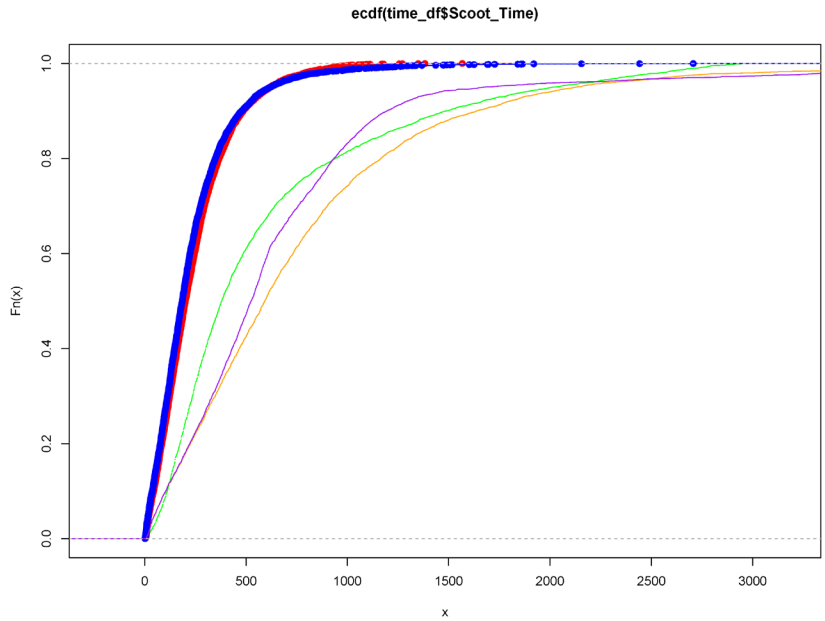

Figure 5.34 – San Diego Travel Time Cumulative Distribution Function – 10,000 Trip Subset

Scooter	Car	Walk	Bike	Transit

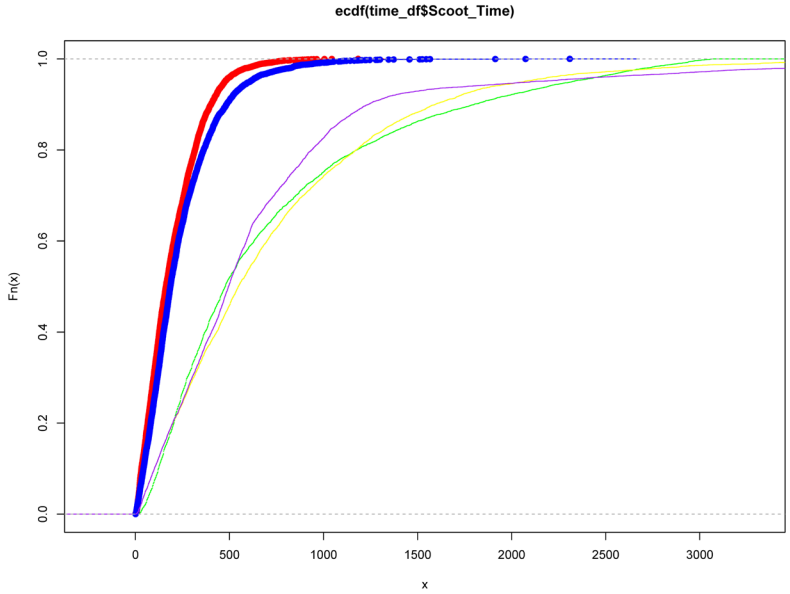

Figure 5.35 – Nashville Travel Time Cumulative Distribution Function – 10,000 Trip Subset

Chapter 6 – Findings

Descriptive Analysis

Histograms

The histograms plotting the distribution of rides in Nashville (Figure 5.2) show a general increase in the number of trips over the course of the day, with the most trips occurring around 6 P.M on weekdays. This time distribution follows the overall pattern described in other reports on e-scooter use, particularly the NACTO report (NACTO, 2019). We see the total number of rides increasing on Friday, with significantly higher ridership on Saturday and Sunday. One possible explanation of this relative surge in use in the strength of the Nashville tourism industry. A common theme in local news coverage of e-scooters has been the assertion that they are primarily used by tourists and for fun (Plazas, 2019). We are unable to address that question directly, but the usage numbers over the course of the study period indicate significantly heavier use on weekends. Additionally, the lack of a defined AM peak may indicate that existing mobility spikes, such as the morning commute to work, may not strongly influence scooter use counts. However, we do see a higher proportion of rides earlier in the day on weekdays as compared with weekends in Nashville. While there is no clear peak, this early morning usage could be indicative of increased work ridership.

The histogram for Portland (Figure 5.1) looks markedly different than the Nashville distributions. On weekdays, we can see three distinct peaks, roughly corresponding with morning work commute, lunch time, and the evening commute or departure from work. The size of the peaks increases as the day goes on. However, each of these three peaks is roughly consistent across days, with a departure from the pattern on Friday. As in Nashville, Friday usage was the highest for a weekday, and the definition of the peaks appears to be less, potentially indicating that usage on Friday is a hybrid of usage on a weekday and a weekend day. As in Nashville, we see higher usage on weekends than on any individual weekday, but the differences appear to be slight. However, the shape of the histogram on a weekend day in Portland lacks peaks or definition, with the most rides occurring in the midafternoon. This pattern is shared with Nashville.

San Diego presents another set of histograms with a divergent pattern across days. Most weekdays appear to exhibit a Midday and PM peak at roughly the same hours as in Portland. However, evidence of a weekday AM peak in scooter usage from this set of histograms is scant.

Weekend ridership in San Diego displays a drastically different trend than in either Nashville or Portland. There appears to be a significant decrease in scooter activity between a peak at 12:30 and a second peak around 5:30 PM. More exploration of the specifics of use patterns on weekends in San Diego is needed to put forward a potential explanation for this phenomenon. However, it is unique among the cities, and an important feature to note.

Figure 5.4 is a histogram of all trips in each city, binned by trip length. We see that in Nashville and San Diego, the largest bin contains trips between 500m and 750m in length. The story in Portland is similar, with the largest bin containing trips between 750m and 1Km. Furthermore, despite the cutpoint of the x-axis on these graphs at 10 kilometers, we see evidence of a very long-tailed distribution for trip length. However, the bulk of trips are under 5 kilometers in all three cities. This common trip length has significant implications for how this mode may fill gaps and address common travel needs (Vale, 2015), though conclusions cannot be drawn from this set of histograms.

Figure 5.5 indicates the percentage of trips that are less than 100m in reported length. This is a significant number of trips across all days in all cities, reaching over 10% for Saturdays in Nashville. There a number of possible explanations for scooter trips that register these very short distances. A possibly convincing explanation is that as a new mode, a number of people try this mode, and quickly (within 100m) decide to end the ride. If this phenomenon explained a portion of the below 100m trips, it would track with the increase in this category on weekends. The percentages displayed here are of the total trips taken on that day in the relevant city, so generally higher use on weekends does not explain this effect. However, I float the possibility that there may be a connection between trying this new device and the flexibility of a weekend. We cannot determine how many of these short rides are taken by new users, as we have no user data for each ride.

Another possible explanation is equipment failures or deficiencies that do not become evident until the ride has begun, such as brake failures or no throttle power. This is a plausible contributing factor, but as we have not information from the Status Changes endpoint to any trip data for this analysis, we can only speculate. Other possible contributing factors may include GPS errors or a lack of infrastructure in which the rider feels comfortable operating a scooter.

Time Bin Maps

The three sets of five maps at the beginning of Chapter 5 display net change in scooters through trip activity. In each city, we see a clear pattern of movement towards the Central Business District during the AM block, with net gains for most hexes close to the CBD centroid.

We see the reverse in the PM block, with most hexes near the CBD centroid losing scooters, and areas further away from the CBD showing net gains. This shared pattern is an indication that scooters move in the way one might expect for transportation taking people to and from work locations at normal work hours. The AM and PM time blocks include only trips on weekdays. During the other two weekday blocks, Midday and Evening, we find less predictable patterns of movement. During the Midday hours, we see a general movement of scooter trips that begin near the CBD, with more destinations than origins in hexes just outside the CBD. However, the magnitude of this effect appears to be less pronounced than during the PM hours, as seen in Nashville (Figure 5.2).

The Evening maps present mixed stories. In Nashville (Figure 5.4), the overall trend of dispersion from the CBD is present but weakened, with no hex falling into either of the categories at the tail of the distribution. When we consider that these maps are showing tens of thousands of trips (or hundreds of thousands, for San Diego), the reality that the origins and destinations are often within 500 net difference from even is astonishing. Many hexes, even near CBD zones, are within 10 trips of even in/outflows during these time blocks. These calculations do not include rebalancing, and only measure trip activity. The near-equilibrium states in many active hexes lends support to a general principle underlying dockless mobility, where user flows create natural asset management patterns (Pal & Zhang, 2017). In Nashville and Portland, we see trends away from from the CBD during the evening, but San Diego shows a strong net gain of scooters near the CBD during this time block.

The weekend trip maps (Figures 5.10, 5.15, 5.20) present another mixed story between cities. In all three cities, we see general movement away from parts of the CBD and movement into adjacent hexes, with a magnitude of mixing not necessarily visible during other time blocks. In Portland and San Diego, one can observe net gains in hexes that border water, while nearby CBD zone hexes experience net loss. During weekday AM and PM blocks, these hexes show net changes with the same directionality, so the divergence is of note on the weekend. The factors contributing to this divergence bear further investigation, but may indicate the presence of attractor attributes that have different gravities on weekends or weekdays (Ma et al., 2019).

Regression Models

The regression models are presented as a 3x3x6+3 set. For each of three cities, the modeled dependent variables are the number of origin points per hex, the mean cost of a trip originating in a hex, the mean distance of a trip originating in a hex, and the number of rebalance points (as described in Chapter 4) in a hex. For the three variables excluding

rebalance points, each model was run on the full set, then five more times on each time bin subset. The rebalance regressions were run only on the full set for each city.

We are able to compare within cities using the six iterations of each model, divided by time of day. We are also able to compare between cities on a directional and significance basis, though we cannot compare effect sizes between cities.

With the exception of the Communities of Concern Index, all variables in the models are log-scaled, including the dependent variable. This means that the magnitude of the coefficient represents the change in the dependent variable given a 1% change in the value of the independent variable.

The Rho values provided below each model indicate the degree of improvement over a non-spatial model, with larger values representing more improvement. In all models, the combination of the Rho value with the Rho Standard Error indicate significant improvement and justify the use of a spatial lagged model.

The Akaike Information Criteria are displayed for both the Linear Model (non-lagged) version and the lagged version of the model. While the magnitude of these statistics is not individually revealing, a smaller (or more negative) AIC for the lagged version of the model again indicates an improvement with the incorporation of spatial attributes.

Density

Our null hypothesis was based on the count of scooter trips exactly mirroring density, so we will start by analyzing the impact of different types of density on scooter activity. Some of the top line results would appear to be self-evident, but the inferences that we can draw from the numerical outputs from these models are multi-faceted. In each iteration of the spatial models, and across cities, Job Density emerged as an important factor in predicting where trips were likely to occur. The association was positive, indicating that greater job numbers in a hex in a strong and consistent predictor of ridership. All hexes are the same size, so the raw number of jobs in a hex and the density of jobs are equivalent measures.

An important lurking variable in all of these analysis is the presence of tourists, visitors, and other people whose presence in a hex would not be typically captured in demographic data. With the available national datasets, the spatial overlap between high core job density and the presence of many people not present for work is difficult to disentangle.

Another interesting feature of Job Density is the relationship with Mean Trip Distance. In Nashville, we see that rides starting in areas with high job density are shorter, with a persistently significant negative coefficient across all time periods. One possibility is that

destinations for riders beginning in these places are closer, simply as a function of overall increased density. This is a possibility, and aligns with general conceptions of urban attractiveness economic, with central locations appearing more attractive because of their proximity to a greater number of destinations. Without making a judgement on what types of destinations are found attractive by scooter users, this theory might provide a good partial explanation for the inverse effect of job density on scooter trip length.

However, we see the opposite relationship in Portland, with a consistent positive relationship between trip distance and job density. These coefficients mean that at all times of day, trips that start in areas of the city with more jobs will tend to be longer, on average, than those that start in less job-dense areas. The consistency of this effect would necessitate an alternate explanatory hypothesis. We must also consider San Diego. In San Diego, the directionality of the job density coefficient is only significant for the AM and PM Peak periods, and in opposite directions. In this case, we have a plausible possible explanation if scooters are used heavily for commuting during these periods, and most commute trips in San Diego are to a dense urban core. During the morning, trips starting in job-dense areas are shorter, possibly because these riders have shorter distances to travel to reach the CBD zone. In the PM Peak, this reverses, and trips that start in job-dense areas have further to travel to reach the non-work destinations. These interpretations could be tested by extensions of this work covered in Chapter 7.

An additional possible interpretation of the distance predictor is related to the finding that more trips start in job dense areas, and the urban design that characterizes those neighborhoods. In all three focus cities, the densest hexes exist on a regular street grid. According to survey results (PBOT, 2019b), many people who try scooters have never used a bike to navigate a city. If they begin a trip with a destination in mind (an assumption that I will address later), navigating to that destination within a regular grid is likely to be a more efficient process due to relative directional ease, even while experiencing the city in a novel way. Conversely, areas with low job density take a variety of urban and street forms, and selecting an efficient route may pose more of a navigational challenge, even if the intended destinations are comparable distances from the origin point. This is conjecture, as we have not analyzed revealed routing information compared with idealized routes returned by Open Trip Planner.

We see the same directionality and significance of Job Density on Mean Trip Cost in each city as was found for Mean Trip Distance. This makes sense, as Cost and Distance are both tightly related to the duration of a trip. The cost of the trip is calculated as $1.00 + $0.15*minutes. We have analyzed both separately because they are representative of different

aspects of use, and both have directly interpretable units (meters and cents). Additionally, analyzing these related but separately calculated variables separately allows us to sidestep some of the long-tail distribution peculiarities of the Distance values.

Population density emerged as a factor with mixed associations between cities. This is not a surprising result, and the presence of this association represents the core null hypothesis for the spatial regression analysis. If no factors influenced how many scooter trips begin in an area besides the number of individuals present to ride them, we would expect the distribution of scooter trips to exactly match population density. While there is a positive effect, we do not see a particularly tight correlation. In San Diego, the association between population density and trip generation is significant and negative in each time block, directly falsifying the null hypothesis.

Marginalization

The Communities of Concern Index (CoCI) appears as a significant associative variable in all three cities, across different sets of regressions. In Nashville and Portland, CoCI was significant in the negative direction during some time bins, and not a significant associative variable during others. This suggests that in these cities, even after accounting for the other factors in the regression, fewer trips start in communities higher on this index of marginalization. The time bins during which this variable was significant differed between these two cities, but the largest effect size in both Portland and Nashville was seen during the Weekend time bin. This significant, negative association was shared in San Diego as well. A possible explanation here could find a difference between "directed" trips during weekdays, and "leisure" trips on weekends. Additionally, if a larger share of trips on weekends are taken by tourists, those trips may be less likely to start in high CoCI hexes.

In both San Diego and Nashville, hexes higher on this index were associated with the origin location of more expensive and longer trips. This finding is in line with the results of previous studies on marginalization and dockless mobility (Qian & Jaller, 2019). There are a number of possible explanations, but without further analysis of the trip purposes (discussed in Chapter 7), we are limited in the conclusions we can draw. However, if we are to assume that individuals are "consuming" scooter use to a degree that attempts to address their mobility needs, higher rates of consumption in marginalized communities could be a very promising finding for this new mobility paradigm. Significant further exploration is required to support this claim on a broad scale, but there are a number of positive indications in this regard. This includes a recent ridership profile survey conducted by a scooter company showing that its riders are significantly more diverse both ethnically and economically than the bike commuter

population (Wachunas, 2019). The Overall model results from Portland are in line with these conclusions as well. However, the finding of higher scooter consumption in marginalized communities did not appear to be significant during most time bin regressions in Portland. The Portland user survey (PBOT, 2019b) also lends credence to this claim.

The following exploration operates under the assumption that trips originating in a hex are primarily taken by people who live in that hex. The potential pitfalls with this approach are expanded upon under the Limitations heading.

Geographic Factors

An additional major difference between the cities included in the analysis is the strength of the public transportation network. Nashville has only bus lines, and the Nashville public transit system is generally regarded as "slow inefficient and outdated" (Stockton, 2018). A recent and contentions ballot initiative in Nashville would have invested many billions of dollars into public transportation improvements in the region, but was narrowly defeated after major political spending by the Koch network. The debate and narrative surrounding this episode clearly show that Nashvillians view their public transit network as inadequate. The implications for scooters as a viable mode in this context are wide-ranging.

Portland and San Diego both have well established light rail and bus rapid transit systems. While these systems suffer from many of the same problems to most North American public transit agencies, including deferred maintenance, decreasing ridership, and uncertain financial futures that prevent long term investment (Zureiqat, 2010).

The density of street network within a hex did not show a consistent pattern across cities or time bins. As such, we make no claims about the generalizable impact of this factor on scooter activity.

Rebalancing Regressions

The set of possible systemic underlying patterns leading to lower ride counts in high CoC hexes is equally broad, and likely work in tandem with some of the demographic factors. One straightforward explanation for ride differences are differences in scooter availability, either because fewer scooters are placed within these hexes under the direction of Gust or fewer rides end in these hexes. By including Rebalance Points in our spatial lag models, we have separated these possible explanations, and can examine the validity of both. Additionally, we decided to run a regression for each city looking at Rebalancing Points as the dependent variable, to understand what factors are associated with this crucial behavior.

We see that for every spatial lag model that examines the number of rides that start in a hex, Rebalance Points emerges as highly significant. This is a logical result, but does confirm that increased scooter availability is almost always the most important factor in scooter consumption.

The Rebalance Points regressions indicate that in all cities, job density is a significant positively associated variable with rebalance locations. Conversely, the distance to the CBD is a significant negatively associated variable in all three cities. This result is interpreted as rebalancing activity generally happens to move scooters closer to the city center, rather than dispersing them because they end up concentrated. In all cities, the Near Transit indicator was positive and significant. While the largest values of this variable occur at the CBD, the peaks are distributed across the cities along major corridors. The strength of this variable, which in all cities had a greater magnitude than distance to CBD, indicates that rebalancing occurs not just towards the center, but towards more active streets likely to have bus stops.

Finally, we examine the impact of the Communities of Concern Index on rebalancing activity. In Nashville and San Diego, hexes higher on this index were less likely to see rebalancing activity, both with significant coefficients of moderate magnitude. However, in Portland we see the opposite, with Communities of Concern actually being weakly positively significant in predicting where rebalancing is more likely to occur.

One potential reason for this difference is the active stance taken by Portland to mandate the placement of scooters in East Portland, which roughly aligns with the CoC index used in this study (Theen, 2019). This is a major policy difference between Nashville, which started with very few regulations on scooters (Garrison, n.d.) and San Diego (Pell, March 5, & 2019, n.d.), which started with none at all. The structure of the Portland pilot, with permits granted to operators who proved they could meet certain standards, allowed the city to exert far more control over the deployment of scooters.

There other potential explanations for the lack of predictive power for the CoC Index coefficient in Portland. A larger percentage of Portland is marked as a community of concern, possibly meaning that the range of variable scale is diminished. Another possibility is a greater overlap between tourist centers and areas marked as Communities of Concern, or areas with high job densities and communities of concern.

We chose to focus on rebalancing as a dependent variable because of its critical role as a policy lever.

Weather

The weather analyses presented are initial explorations into the role of weather as a predictor of scooter activity. These linear models were run separately from the spatial lag regressions, as the external data was attached to scooter trips on a date and time basis, rather than on a geographic basis. As described in Chapter 4, every trip was tied to the reported weather in the city by the hour of the trip origin. Temperature and precipitation were the independent variables, while trip count and median trip distance were the dependent variables. Once these are conducted in each city, we have a 2x2x3 set of analyses.

Temperature shows a consistently significant positive relationship with both the number of trips that occur, and with the distance people travel. we see a very tight linear relationship between temperature and median distance in Nashville. This linear relationship suddenly and dramatically degrades and falls off around 95 degrees. We assert that people ride longer distances as the weather gets warmer, up to the point where the heat becomes extremely uncomfortable. We see a similar linear relationship in Portland, with the drop-off point around 90 degrees. Regional differences in temperature tolerance and acclimation might explain the difference of this inflection point between cities. San Diego shows a slightly more complex distribution of median distances, with grouped residuals indicating the presence of other effects not captured by the linear model.

The potential implications of this relationship are wide-ranging, and based on the exploratory analysis we cannot directly support any underlying phenomenon. However, this relationship lends significant support to the significance of the hedonic expectation of this mode (Madigan et al., 2016). Simply put, people appear to enjoy riding scooters more when the weather is warmer, so they choose to ride them over longer distances. We must also be aware of how this desire may play out among two major trip type categories: work commute trips and leisure trips. Leisure trips, especially trips that have no pre-determined destination or major consequences for variation in arrival time, may be most easily extended in warm weather. The story is quite different with work commute trips, which connect an origin/destination pair at a fixed distance, and late arrival holds significant consequences. In this case, potential contribution to this linear relationship may not be that people choose to extend existing trips. As temperature increases, people might rather consider scooters a more competitive option for trips of greater length. As a result, a larger portion of longer, pre-existing travel behavior might be shifting onto scooters at warmer temperatures. However, the analyses so far conducted cannot support this claim, and it remains a question for future research.

As temperature relates to distance, the scatter plot indicated the possible presence of a non-linear relationship, so this model was run as a polynomial with a squared term. This model showed significantly improved fit over a simple linear regression, and is presented here. The dependent variable in the count models is the number of trips during the temperature category, divided by the number of hours in which the matching temperature was recorded *and* a scooter trip appears in the database. Details of this statistic are in Chapter 4.

As with median distance, we generally see an increase in the number of trips that occur as the temperature increases. In Portland and San Diego, we see a leveling-off and possible decline in the trip generation rate based on temperature above 80 degrees. In Nashville, the scatter plot indicates the presence of a factor beyond the temperature in terms of predicting the "trip temperature rate" statistic.

While we have not conducted analyses to confirm this possible underlying explanation, the scatter plot is likely mediated by the regular peaks in scooter activity that coincide with regular temperature ranges. The count of trips in any given hour is driven in large part by time of day, as is the temperature. Cooler temperatures typically occur early in the morning and late at night, with the highest temperatures falling in the early afternoon. This shared lurking variable reduces the conclusions that can be drawn from the models presented.

Another effect of the temperature distribution within a city is the smaller number of hours that feed into categories at the tails of the independent variable. This results in reduced representativeness for the summary statistic used in the model fit. While this is a drawback for temperature, the impacts of this methodological effect are particularly pronounced on the Precipitation analyses.

In both Nashville and San Diego, we see the presence of outliers in the precipitation. As an illustration, we can dive into the outliers in Nashville. The two outlier precipitation points represent the data from 18:00 – 19:00 on Wednesday, May 16th, 2018. This covered two weather reports, one indicating 1.02 inches of rain in the previous hour, and the next indicating 1.62 inches of rain in the previous hour. These two outlier points represent only 36 trips. The method used of running a linear model rather than a weighted model that accounts for these disparities is evident given the distribution of precipitation rates during analysis hours. As a result, we draw no conclusions from the San Diego or Nashville precipitation models.

The distribution of precipitation rate across hours is substantially more even in Portland, a city renowned for its rainy climate. We can begin to draw inferences from these models, even as we acknowledge the drawbacks of the method overall. The models indicate that as precipitation increases, both the number of trips that occur and the length of those trips is likely

to decrease. The Portland count model indicates that for every additional 1/10 of an inch of rain, 16 fewer trips will occur. The Portland distance model suggests that for every additional 1/10 of an inch of rain, the median trip length will decrease by 95 meters. These effects, while statistically significant, are relatively modest in magnitude.

Further analysis is needed to substantiate any predicted effects from these exploratory models. However, if additional study yields results that bolster the preliminary findings from Portland, this would constitute a relatively unintuitive relationship, with inclement weather not having a relatively minimal impact on the scooter trip rate or distance.

Many people cite protection from the elements as a reason they prefer using a private automobile (Ramezani, Pizzo, & Deakin, 2018). As a result, one would expect a more precipitous decline in scooter use during rain. A steeper negative model slope might indicate that the viability of scooters as a robust and flexible mobility option is highly weather dependent. However, we do not see a strong effect. There may be a gap between the perceived intensity of precipitation as a disincentive and the revealed intensity of precipitation as a disincentive. As we see from literature on e-bike mode shift, personal experience is the most important factor in creating long-term habit change (T. Jones, Harms, & Heinen, 2016; Popovich et al., 2014), and these changes are working against powerful forces of path dependency and ingrained behavior (Rosenfield, Attanucci, Zhao, & Zhao, 2019). With increased scooter experience, this gap may narrow. Significant numbers of people do unquestioningly commute or travel by bicycle in the rain, so there is certainly hope that scooters may find their way into wide adoption for all weather commuting.

Mode Comparison

At this stage, the comparison of travel between modes is preliminary, but initial results indicate the potential for exciting findings in future iterations of this research. A discussion of how this framework will be iterated and expanded upon can be found in Chapter 7. All results discussed below should be considered preliminary and illustrative. Specific values are mentioned as a direct interpretation of the graphs presented in Chapter 6, but the results of a more comprehensive analysis may differ significantly.

Two Cumulative Distribution Function (CDF) graphs are presented, one for Nashville and one for San Diego. A cumulative distribution function plot has the variable of interest on the x-axis, with the y-axis representing the percentile value of each data point, scaled from 0 to 1. These plots examine travel time as the variable of interest. To construct each line, each value is

plotted on the x-axis. The trips are ranked by duration from shortest to longest, and the y coordinate is determined by the percentile rank of the value within the point set. This construction technique provides a curve with consistently positive slope, though the instantaneous slope varies along the function. When multiple CDFs are displayed on the same plot, one compares performance overall by comparing the area under the curve, with larger regions indicating better overall performance. One can also examine comparative performance at a specific value of the variable of interest by comparing the y-values. In a vertical comparison, the y-value represents the percentage of trips on the mode that have a shorter duration than the x-value. By comparing across a y-value, you find x-values that indicate the duration at which the same percentage of trips is shorter than the x-value. An important note on these comparisons is that because each set of records (divided by mode) is ranked and ordered separately, vertical comparisons are not comparing the same O/D pair. Rather, vertical comparisons show only the percentage of trips below a certain duration on each mode.

The data used in this analysis is drawn from the 10,000 trip randomized subset from each city. As referenced, Open Trip Planner was configured using the parameters described in Chapter 4. Time values for scooter trips (in green) are based on actual trips, cleaned using the standard outlier procedure described in Chapter 4. Time values for all other modes were calculated using the exact O/D location for the scooter trip. The trip start time and date were also replicated from the scooter trip, as this has importance for transit routing. We must note that the calculated time performance of the non-scooter modes reflects idealized flow conditions: no congestion, no stoplights, and on-time transit performance. We would expect CDF plots of revealed data on these modes to be generally lower overall. This concern is addressed in Chapter 7.

In both cities, we see that Car and Bicycle performance is significantly better than the other analyzed modes, and show very similar performance to each other. Hypothetical Car travel time is generally shorter than Bike time in Nashville, with San Diego showing the opposite. Due to the significant separation between the pair of Car and Bike and the Transit, Walk, and revealed Scooter, I focus this exploration on a comparison of the last three modes.

In Nashville, results indicate that the shortest 20% of scooter trips (approximately 3 minutes of recorded scooter time) could be completed faster than the shortest 20% of public transit trips. Scooter trips between approximately three minutes and nine minutes in length represent approximately 35% of the 10,000 trip sample. This same percentile slice on public transit would not have been able to perform as quickly. Scooter trips between nine minutes and 40 minutes are approximately 20% of the sample, starting at the 55th percentile. This range of

trips is completed faster on transit. Trips longer than 40 minutes represent approximately 4% of the sample.

Within Nashville, approximately 80% of scooter trips were completed in under 20 minutes. The same is true of walking trips. appear to be faster than walking until the length of the scooter trip reaches approximately 20 minutes. Above 20 minutes, the percentage of walking trips that are completed rises faster than scooter trips are completed. Approximately 80% of the sample O/D pairs were connected faster on scooter than with a hypothetical walk. For the 20% that remain, a few theories can be offered. We know that route choice in a new mode can be difficult (Papinski & Scott, 2011). It is possible that scooter riders chose an extremely circuitous route due to unfamiliarity. They could also have traveled on a circuitous route due to obstacles such as unsafe infrastructure, a factor that Open Trip Planner does not consider for Walk routing. Finally, these trips could be serving a partial transportation function as well as a partial recreation function, in which efficient travel between a pair of coordinates is not the sole motivating goal.

In San Diego, only the shortest 10% of scooter trips (under two minutes in duration) were slower than the shortest 10% of routed public transit trips. Scooter trips between two minutes and fifteen minutes in duration represent approximately 70% of the scooter trips in the San Diego 10,000 trip subset. The 10th to 80th percentile of transit trips took more time to complete than the parallel category of scooter trips. The 80th to 95th percentile of scooter trips are comprised of trips between fifteen minutes and half an hour in duration. The same percentile range of public transit trips were completed faster. At the 95th percentile of scooter trips by duration, Walk, Transit, and revealed scooter trips converge.

From the CDF plot of San Diego, we can say that 70% of scooter trips linked O/D pairs faster than 70% of public transit trips. However, we cannot say that each 70% segment is comprised of the same O/D pairs. For instance, a scooter trip that began and ended in downtown might have an O/D pair that could be quickly linked by transit, but the rider chose to use the scooter to travel to a friends apartment, pick something up, and then return to work. Her revealed trip would be ranked by the duration of actual time, while the matching O/D transit trip would be ranked significantly better. However, with such a large gap between the CDF curves, we can understand that there are significant and meaningful differences between the modes examined.

Chapter 7 – A Mode Shift Analysis Framework

Theory

Mode Choice

Much of transportation modeling builds on the theory of planned behavior. Transportation models often assume that individuals have reasonably good information about their options, and as a result of this information, make decisions about how to proceed by carefully weighing costs and benefits of each option. This logic underlies the Discrete Choice analysis, built from a series of Utility Functions (Ben-Akiva et al., 1985).

To operationalize this framework into a large scale transportation model, this discrete choice function is nested within a four step model. In the first step, trips are generated based on demographic information about the geographic zone. The second step is trip distribution, in which the purposes generated for each trip are matched with a destination that will allow fulfillment of that purpose. The third step is the mode choice analysis, using a logit function as described above. The fourth step is the assignment of a route upon which each traveler will proceed.

The four-step travel model is used across the world as a reliable predictor of travel behavior, and the flexibility provided in the trip generation definition step allows for effects testing of new developments or transportation interventions. However, the reality of many four-step models is somewhat different from this idealized version. In the United States, each Metropolitan Planning Organization is required to keep a model of regional transportation changes. These models are adjusted over time as environmental and demographic factors shift, but the underlying data and structure are rarely adjusted. In a major metropolitan region, the MPO model has expanded through slow accretion of new information. To maintain backwards compatibility, the nuances accounted for with new information are limited. The amount of data driving the model is immense, and the division between deprecated and still active information is sometimes unclear. Due to the complexity and recursive nature of the model structure, programming any new adjustment can take multiple days, and verification that the new parameters are an accurate representation of the intended scenario is near impossible. During a conversation with MPO staff, it was revealed that while the results of the model are intended to be deterministic, they appear stochastic, even to those who are employed to manage and understand the model. A single run of the model typically takes a week and costs approximately $175,000 (Meeting Notes, 2018).

In this context, there has been a significant appetite for more flexible, scenario-based modeling tools. While the magnitude of data incorporated in the MPO model may have been a distinctly positive quantity during the heyday of cybernetics, there is a now a widespread recognition that many datasets, both old and new, are woefully unrepresentative (Favaretto et al., 2019). This is a particularly pernicious problem with so-called "black box" models, where the input information is either confidential or the model execution is not public record.

A scenario-modeling tool allows for relatively straightforward adjustment of input parameters, and prioritizes flexibility and speed over the exact return value. Instead of providing a (theoretically) precise integer value of trips like a four-step model would, a scenario model might provide a range with confidence intervals, and be optimized for sensitivity analysis.

Methodology

Scooter Mode

As described in Chapter 1, there are a number of questions about the future of mode shift following the widespread introduction of shared electric micro-mobility. Beginning in Chapter 4, I described the background and some potential uses of the Open Trip Planner tool. In a traditional four-step model, the most computationally intensive steps are the utility calculations and the route assignment. In an age of cheap computational power, one can remove this as an obstacle to better demand modeling by modifying this open source routing engine to suit the needs of the project. As a critical component of this proposed modeling framework is its flexibility, I will highlight easily adjusted parameters with (*).

By leveraging the recently released *otpr* and *opentripplannerr* R packages, one can incorporate this powerful routing engine and all the configuration options directly within the *sf* R environment. This coding and data analysis environment is geared towards large spatial datasets and vectorized operations. The modifications described below will be partially implemented in R, and partially through modification of JSON configuration files that control the build of the routing graph. References to a hex geography refer to the same 500 meter hexagonal grid that was used in the rest of this research. However, the size of the hex cells(*) can be adjusted if initial results indicate that finer geographic fidelity is required, especially within city centers. Additionally, non-hex network buffers(*) around public transportation stops(*) and hubs need to be highlighted and separately categorized.

In this research framework, one should modify the existing bicycle mode characteristics to better reflect the ways in which an electric scooter moves through a city. This allows not only insight into how trips taken on scooters might perform on other travel modes, but the ability to simulate stochastically-generated trips on all existing modes plus scooters. The default

configurations of the bicycle routing engine set a consistent speed of 11 miles per hour. One could adjust that on a trip-level basis, with simulated scooter trips traveling at the speed of the average speed of trips departing from the same hex(*), mediated by time of day(*). The speed will vary based on segment, in the same way that car speeds vary with provided speed limits. A guide for these inferences is (Arellano et al., 2019). An important and useful feature of the bike routing as built into Open Trip Planner is the incorporation of elevation change as a decision factor in route choice. Hill tolerance is a critical difference between bicycle operation and an electrified micro-mobility option such as e-scooters or e-bikes. Open Trip Planner provides the option to modify this option in a triangle tradeoff framework between Slope Tolerance(*), Safety Need(*), and Time(*). Safety need is a measure of the minimum safety level required for a potential user to consider it in a network segment, with longer segments being weighted more in the overall safety consideration of a route. A short segment on a very dangerous street, such as a 45 mile per hour divided highway, may have equal weight to a longer segment on less dangerous a 35 mile per hour four lane road. The third factor in the triangle is Time, which prioritizes more direct routes at the expense of Slope and Safety. Based on analysis of existing scooter trips along the directness criteria described at the beginning of this chapter, Time is not an important characteristic in the decisions of scooter riders on routing. Slope Tolerance is increased, as the powered device makes hills a much smaller routing disincentive. One is left with a heavily weighted triangle on Safety, which aligns well with news reporting and survey responses of scooter riders on what types of infrastructure they prefer (NACTO, 2019; PBOT, 2019b).

Importantly, Open Trip Planner also allows for the use of this modified bicycle as an access mode to or from transit. This is a critical level of nuance that significantly expands the range of trips for which a scooter might feasibly be used, based on the distribution of trip lengths in the sample. Open Trip Planner permits taking a bicycle on transit as a default, this setting should be adjusted for the scooter mode(*). The maximum distance permitted as an "access" leg(*) is adjustable as well, an important consideration given the range differences between bicycles and electric micro-mobility (Fyhri et al., 2017).

Another modification of the scooter mode is incorporation of the scooter search process that typically precedes a ride. One might stochastically generate a routing through-point within a five minute(*) walking buffer of the trip origin. The trip will then simulate walking to the through point, then routing via scooter to the destination. The distance of the walking buffer should be related to the density of scooter deployment, which might vary by the factors described in the regression analysis in Chapter 6. The price calculation will include two separate components: the cost of the scooter ride, and the value of time(*) incurred by the scooter search.

Open Trip Planner incorporates modeling of station-based bike share, which allows a degree of sensitivity analysis in the location of the stochastic through-point. One could simulate a docked scooter sharing service by making modifications to the bikeshare mode(*) instead of the bicycle mode.

Finally, the other analyses described in this research provide some ability to model differences in mode choice in relation to weather. This is sometimes included as an explanatory variable in a utility function within a logit model, and with a mode that is so exposed to the elements, we must pay close attention to the scaling of this variable(*). One could provide separate explanatory variables relating to temperature(*), precipitation(*), humidity(*), and others. As an extension, one could explore refactoring a composite index of weather/scooter congruence.

Existing Modes

As described earlier, the routing conditions in Open Trip Planner return idealized itineraries. Congestion is not considered in the calculation of travel time or route choice. Parking search and stop lights are not factored in as well. To bring the car mode in line with reality, a set of scale factors will be used to address these deficiencies.

Congestion data for each segment can be constructed for a city by running a traveling salesman set of trips through the Google Maps API at different times of day over a period of days. This should capture an approximation of the differences in travel time during each period, while the traveling salesman route configuration is an attempt to get congestion information for as many segments as possible. An alternate workflow would be to scrape a series of screenshot images of the display of predicted traffic information in a city for every possible time segment. These images could then be associated with map files, so the depicted streets line up with a shapefile of the street network. Raster inference would then be used to tie the raster color to the centerline shapefile, with a new column for each time of day. This has the potential to create a massive dataset. Both of these congestion options risk replicating the “black box” mentality of older demand modeling, and both have significant external dependency.

To address parking concerns, a stochastic through-point would be added within a variable network distance(*) from the destination. The buffer size would vary by time of day as it relates to parking utilization, as well as by trip type. Work commute trips are significantly more likely to involve parking at a dedicated, reserved spot with no spot search costs incurred (Hamre & Buehler, 2014). Once the route arrived by car at the stochastic through point, it would proceed on foot to the listed destination point.

The cost of parking would be a scale factor based on time of day and a normal distribution of stay durations(*) for each trip type. There would again be a small time-value cost(*) incurred based on the walk distance.

The expense of operating a car will be generally estimated, using a conservative estimate of approximately $4.00/hour(*).

To simulate TNC and ride-hail trips, one would instead generate a stochastic start location, and modify the provided origin to act as a through point. To model shared trips, one would first generate a direct route linestring, then calculate a stochastic through-point within a five minute driving distance of the route centroid. If a shared trip is over 15 minutes in duration, two stochastic through points will be generated. These represent the route detours and time penalties incurred through the activities of non-modeled riders.

To estimate cost for TNC rides, I intend to use representative API calls to the TaxiFareFinder website. With a geographically randomized set of O/D pairs and times of day, one could extrapolate a geographically and time weighted regression of cost factors, comprised of the origin location, the time of day, and the destination location. Each should provide a component of the estimated fare. Once a well-fitting model is constructed, this will be used to estimate a TNC fare with confidence intervals(*). Local surcharges or tolls could be added as needed(*). The fare for a shared ride will be calculated from the fare for a standard ride, multiplied by a fractional coefficient(*) representing the sharing discount.

Transit routing will be subject to the same congestion coefficients(*) as car travel if the leg is on a bus. To simulate construction of bus priority lanes, these coefficients(*) could be modified or scraped on targeted segments. Transit fare will be calculated using the local fare scheme, with the addition of a wait penalty(*).

Origin / Destination Matrix

The geographic inputs for this model will be drawn from three sources: revealed behavior MDS records, the Longitudinal Employment Household Dynamics Survey, and a grid with random offset. This methodology is based on (Imbens, 1992). Most transportation modeling exercises rely on the construction of Transportation Analysis Zones, or TAZs. This is a time-consuming and data-intensive process, though certainly provides a degree of value to the final model (Zhao & Zhao, n.d.). However, the relation of TAZs to trip generation (Chmielewski, 2019) is also intensive and errors at this stage can cascade through the model.

An argument can be made that relying heavily on representing the world as it currently exists does a disservice to the overall utility of the model, especially if it is intended for scenario planning rather than evaluating specific infrastructure projects. By not defining the current state

of trip origins and distributions, and instead allowing generative and attractive factors to be modified as parameters, one substantially increases the exploratory power of the model while reducing the impact of data inconsistencies.

Finally, for the micro-mobility questions, the route assignment step is not a focus of this work. Records should be kept of the routes used to generate the explanatory variables within the utility functions that feed the multinomial logit model for each O/D pair.

Trip Classification

In addition to time based classifications, one could tag each O/D pair with a series of dummy variables that can be used to analyze segments at a later date. These include weather dummy variables and a categorical variable relating to a matrix of trip types, inferred from basic population and employment density statistics. Additionally, each trip will be tagged with the hex ID of the origin and destination points, in order to facilitate comparison and linkage with the regression results and descriptive statistics described elsewhere in this thesis.

Expected Results and Use

The proposed framework should allow for a number of result types. A simple predicted mode split for all tested trips will be the most straightforward output. This will vary based on the parameters defined during model configuration. Depending on the intended audience, the configuration of these parameters could be set through an RShiny web application. This would allow easy sharing of the tool, and eliminate the need for coding knowledge to test different scenarios.

Other potential results from this framework include highlighted neighborhoods where mode shift might be most substantial, identified times of day during which micromobility could play the largest role, or potential locations for transfer hubs between first mile/last mile micromobility and long-range public transportation.

Chapter 8 – Discussion

Transport Marginalization

There are a few potential competing explanations for the trip generation differences. In general, we can attribute this pattern to either individual preferences and characteristics, or to environmental and systemic factors. Along the individual dimension, there are a number of possible explanations, most of which have been previously put forward in popular reporting on scooters or extrapolated from studies on other travel modes. The spatial regressions reveal connections and correlations, and allow for prediction of the dependent variable, in this case scooter trip features, when the underlying variable is known. With an index as a predictive variable, the inferences drawn from the coefficient can be wider ranging. For this study, components of the Communities of Concern Index were tested both independently and together in the same regression, and the Index proved to be a stronger and more consistent overall predictor than the individual component factors. While this strengthens the model, it means that there are many competing rationales to explain the underlying effect. It is likely that many of the possible reasons outlined below contribute the results seen in the spatial model.

The most basic possibility is that people in these hexes are simply less interested in using shared electric scooters. This is unlikely based on the results of the Populus survey (Clewlow, 2019) measuring stated interest and the Lime ridership survey (Wachunas, 2019) assessing the ethnicities and household incomes of scooter users in a variety of cities. Another possible implication is that fewer people in these hexes have access to the technology required to easily unlock and use a scooter. The requirements are a smartphone with GPS and a data plan, as well as a credit card, and while there is broad market penetration of these products in the U.S. generally, evidence of a Digital Divide persists (Pierce, 2018). Some reports have indicated that even if someone has these two technological components, there is a hesitancy to provide credit card information to a new and unknown app (Groth, 2019). The potential negative effects of credit theft would be larger for an individual with fewer resources, and this could be considered a prudent decision.

Pride is an important factor that influences mode choice and transportation decisions (Moody & Zhao, 2019). Hoffman explores the shame associated with using a bicycle to travel if that choice is bounded by income or legal restrictions (Hoffmann, 2016). While our study did not address this question, the concept of an electric scooter be so divergent from a bicycle that

the non-car shame is insignificant, or even inverts to be expressed as modal pride. Reports conducted by scooter companies hint at this possibility, though there has yet to be robust research on the topic. If a significant divergence was found in attitudes towards bicycles and bicycle users as compared with scooters and scooter users, this could be a powerful explanatory variable in mode shift analyses. As Hoffman implies, we might expect to see the magnitude and even the direction of this effect shift with socio-economic status. If "scooter pride" is found to be more common among low-income individuals, this would point to another equity advantage of this mode over existing micro-mobility.

An important potential explanation for lower ridership in high CoCI hexes is the perceived or actual cost of using these scooters (Phillips, 2014). While some scooter companies brand heavily based on the cost structure, this is not true for all providers. Additionally, the cost structure, which is inherently time-based, creates the potential for price uncertainty with potential users. Especially if a potential rider does not use a bicycle to navigate (and based on mode split numbers for the cities in question, they likely do not), attempting to estimate how many minutes it may take them to reach their destination may be a difficult task. If a user makes an estimate with an expected margin of error of plus or minus five minutes, the expected fare for the trip has a range of $1.50 between the considered possibilities. When these scenarios are viewed in the context of humans' well-documented impairments in time estimation, pricing estimation, and travel time estimation in unfamiliar conditions, the possible price for a ride becomes even less transparent (Kahneman, 2003). There is some evidence that financial risk tolerances vary by income level (C. Wang, 2019), with higher income individuals who have a larger cushion to absorb financial setbacks being comfortable with riskier behavior. The financial risk inherent in signing up for an unknown app and a scooter ride with unknown cost differs by income and wealth available. The number of rides with costs that balloon is quite small, with a $10 ride falling in the 95% percentile of ride costs in Nashville.

The questions of payment are a good example of the *procedural equity* aspects of this new technology. The ability of individuals to navigate the meta-experience before and after the use of the scooter is just as critical to the experience and potential for mode shift as the *product equity* outcomes like travel efficiency.

Many people do not have reliable access to card-based funding, and even some who do prefer to operate primarily in cash (Groth, 2019). Cash payments on scooters would be impractical, but there are many ways to facilitate cash-grounded transactions. Some scooter companies follow the model of the transit card. Scooter companies could sell refillable cards through a network of existing bodegas and convenience stores. These cash-funded pay as you go

cards, which interface with the scooter through RFID, would bring scooters onto a more level playing field with existing transit options.

Mobile data access, mentioned at the beginning of this thesis as a prerequisite and foundational technology for making shared electric scooters a reality, must also be accessed by users. While smartphone market penetration rates have consistently increased and these devices are widely used across income levels, unfettered data access still remains a limitation (Groth, 2019). Even if the price of the ride is comprehensible to potential riders, the possible data implications of using a scooter for 40 minutes may be significant, or may at least appear to be so. To a potential customer, the legibility of this factor is not necessarily obvious, and as with many of the other topics discussed, the relative impact of a data overage on one's finances varies across incomes.

One issue that was raised in a conversation with a mobility equity advocacy group was a concern about racial profiling. Similar to the pervasive and widely acknowledged practice of over-policing of Black drivers who are operating fully in the law, there are concerns that policing of scooter use would take on a racial dimension. This is especially a concern when the regulatory environment in which people operate scooters is unclear (Anderson-Hall, 2019). Despite instructions present in the apps that allow people to unlock scooters, confusion persists about where to ride scooters: sidewalk or road. Most states mandate the use of e-scooters on the road, though Colorado famously mandates that they only operate on sidewalks(Arellano et al., 2019). However, especially for riders not used to a mode other than walking or driving, feeling safe is of paramount importance, and they may operate in whichever environment feels most safe at the time, banking on the leniency of the governing structure to permit what is technically an infraction. If this assumption of good faith in the police system is not present, as is the case in many communities of color that have experienced decades of systemic bias and aggressive policing, people may be less willing to operate scooters with regulatory uncertainty, simply because the possibility of a police interaction as a result is too great.

Many scooter companies provide fare equity programs, which give a discount to riders who qualify as low-income. This qualification can either be through proof of receipt for federal or state benefits, as with many bike share systems, or through direct income verification. The type of discount varies between providers, from a reduced per-minute rate to a reduction or elimination of the unlock fee (HR&A, 2018). However, despite the presence of these low-income programs, use of these programs remains limited. During their four month pilot in 2018, the City of Portland reported only 43 registered low-income users (PBOT, 2019a). This may be indicative of procedural hurdles required to access the low-income programs. As in other

transportation contexts, the procedural hurdles may appear to pose too much of a challenge for residents to consider enrolling, despite the benefits (Phillips, 2014).

A critical dimension of equity that has gone relatively unexamined in this thesis is the effect of gender on scooter use. As described in Chapter 1, women and men use bicycles at different rates and for different activities (Garrard et al., 2008). The data used in this analysis provided no insights as to the gender of the users, and we are not able to make any claims on whether scooters are more or less gender equitable than bikes as a result of this research. However, surveys conducted by scooter companies and cities indicate that ridership is indeed more balanced across genders than bicycling surveys conducted by the American Association of Bicyclists (NACTO, 2019; PBOT, 2019b; Wachunas, 2019). This is encouraging, and policies to construct the protected multi-modal street designs that are preferred at greater rates by women should be prioritized. Scooters and other forms of micro-mobility continue to benefit from this expanded network, as do bicycles.

Shared electric scooters, as described in this paper, have a relatively high prerequisite for physical agility, when compared with a car or public transit. As we see with ride-hailing, companies have not yet figured out how to effectively serve disabled potential clients (Birtchnell et al., 2018; Cass, Shove, & Urry, 2005).

Related to the previous point about physical agility, while we were not able to draw meaningful conclusions about age from the data provided, news reporting and surveys of riders show a population significantly younger than the population as a whole (Hardt & Bogenberger, 2019). This is not surprising, but does speak to the ways in which this new mode may not effectively serve as many residents as other modalities (Plazas, 2019).

Equity in Urban Design

There are significant urban design implications created by the rise of scooters. One trend that has been observed in other is the strong preference for bike lanes and protected bicycle infrastructure for scooter riders (Dill & McNeil, 2013; PBOT, 2019b). Many studies have observed differences in risk tolerances in cyclists, especially between genders (Garrard et al., 2008; Krull, 2018). A big result of these findings has been an increased push in the past decade for cities to construct safer bicycle infrastructures. Critically, studies show that the construction of each additional piece of this protected, or low-stress bicycle network, has an effect greater than the sum of its parts (Marqués, Hernández-Herrador, & Calvo-Salazar, 2014; Marqués, Hernández-Herrador, Calvo-Salazar, Herrera-Sánchez, & López-Peña, 2015). While new protected bicycle infrastructure has been hailed in many quarters of the bicycling cycling

advocacy community, there has been opposition as well (Braun, 2019; Lee et al., 2017). In many cases, cities prioritize protected bicycle projects in areas undergoing significant real estate investment (Flamm & Rivasplata, n.d.). The construction of bike lanes has become, in many cities, synonymous with the onset of gentrification and displacement of long term-low income residents (Garrett & Taylor, 2012; Krishnamurthy, n.d.).

The question of bike lane displacement was covered in a recent book by Dr. Melody Hoffman, entitled "Bike Lanes are White Lanes"(Hoffmann, 2016). This book chronicles a history of the cycling advocacy movement among communities of color, and highlights major gaps between the ambient narrative of what a bicyclist looks like, and the true diversity of the population of bicyclists. One major issue that she draws attention to is the privilege that is a prerequisite to be considered a "bike commuter", in that for most "bike commuters", there is an inclement weather does not wholly prevent them from getting to work. The popular press image of a bike commuter is typically white and male, and usually makes an active choice to use a bike to work. This implies the presence of another viable commute option, whether it is a personal car or public transportation. The less common narrative of who bikes to access society are people who either cannot afford car a car or who are prevented from holding a license. The implications of car unaffordability, which were discussed earlier, are clear. The licensing guidelines for motor vehicle licenses in many states restrict licenses to people on parole or people, and license suspensions are not uncommon as even without the presence of moving violations (Groth, 2019). Racially biased enforcement systems pervade regulatory frameworks in the United States. The disproportionate percentage of police enforcement actions that target Black individuals is linked to a carceral system that preys on communities of color. One of the many ramifications of these systems is vast sectors of people being legally shut out of private car operation (Cass et al., 2005; Le Vine, Lee-Gosselin, Sivakumar, & Polak, 2013). Biking from need rather than choice , similar to the concept of being "car free" rather than "transit dependent" may describe people with similar patterns of behavior, who might value similar infrastructure investments and transportation policies (Hoffmann, 2016). However, the gap in perceived legitimacy and political power is real.

The regression results above reveal that, in spite of assumed reduced access to safe infrastructure designed to accommodate non-car modes (Braun, 2019; Langston, 2019), we see higher consumption of scooters in marginalized communities. This speaks partially to failure of existing transportation options to meet the needs of residents of these communities (Golub, Marcantonio, & Sanchez, 2013). While our study does not directly examine the dispersion of multi-modal infrastructure, the likely growth in micro-mobility provides cities with a broader

imperative to explore additional micro-mobility infrastructure (Clewlow, Regina R, 2018; Nuñez, Bisconsini, & Rodrigues da Silva, 2018).

An additional consideration of urban design is the pavement quality and tolearances present in cities and new micro-mobility infrastructure. Many cities are now aligning their bicycle and micro mobility lane installation plans to follow existing road repaving schedules, or modifying their repaving schedules to meet the needs of micro-mobility infrastructure. This makes a fair amount of sense from a municipal budgeting and resource allocation standpoint, and often allows for the installation of more radical lane treatments than might be otherwise possible absent repaving. However, as cities make these decisions, it is important to reflect on who the new road surface is intended to serve. This alignment requires a connection between the Capital Projects arm and the Repair arms of the City. Even in small cities, the decision-makers for these bureaucratic divisions might not be within the same department. If transportation planning has been slow to adapt to meet the needs of non-car users, infrastructure construction departments have been even slower (Geels, 2012). The field of pavement studies with relation to bicycle comfort has become most active in only the past two years (Fyhri et al., 2017; Nuñez et al., 2018). However, despite the similarities to bicycles in terms of infrastructure needs in many regards, the smaller tires, smaller suspension profile, and standing posture of electric scooters may necessitate a different level of attention in designed appropriate road conditions (McHugh, 2019).

Figure 8.1 – A visual comparison of the tire size between an electric scooter and a standard bicycle

As currently constructed, the scooter's tolerance to absorb physical shocks is substantially more limited than a bicycle or a car. This is partially due to the physics and mass differential between the vehicles. A physical jolt delivers significantly more energy to the rider on a smaller vehicle than on a larger one, where the kinetic energy is dissipated across more mass. Improving the grade of asphalt used on micromobility construction projects is a way for cities to literally pave the way towards a more micro-mobility friendly future. The issue of pavement quality is particularly challenging for cities that experience long winters, as they must deal with harsher freeze/thaw cycles and more potholes as a result.

Environmental Impacts

While the research conducted during this project does not directly address the environmental and climactic effects of scooter use, this is an indispensable part of the analytical frame. A drive to understand how to reduce the intense environmental degradation that is partially a result of our national car culture is a motivating factor behind this research(Shu & Bazerman, n.d.). Cities are in urgent need of radical change along environmental and equity axes, and too often these are pitted against each other (Gross, Neil, 2018).

A reduction in consumption is required to reduce global resource throughput (Daly, 1996). However, there is an inherent fairness concern with this claim: those who had previously benefited from more extravagant consumption continue to enjoy the downstream advantages while simultaneously imploring those who were shut out of this former definition of Affluence to consume less (Chertow, 2000; Kysar, n.d.). In the micro-mobility space, we appear to have an opportunity to bring the fairness concern and the consumption concern into alignment. As described in Chapter 1, smaller vehicles require less energy to propel a human passenger at the same rate. Electric motors have dramatically increased in efficiency advantage over internal combustion motors (Chang, Wu, Lai, & Lai, 2016). Not only is the conversion from stored energy to kinetic energy more efficient, but the process of obtaining fuel is drastically different. The production and distribution of petroleum is a tremendously environmentally damaging process in and of itself (Williams, Haley, & Jones, 2015). While significant amounts of electricity continue to be derived from polluting sources (Lawrence Livermore National Laboratory, 2019), the technology currently exists to produce near boundless electricity in a carbon-neutral way (Nivola & Crandall, 1995; Plummer, 2018). Deployment of these energy generation systems is ongoing, but there is no reason to wait for one section of the global energy puzzle to be in assembled before working on a related section (Dennis, Colburn, & Lazar, 2016; Quay, 2010).

Not only are scooters more efficient, but they occupy significantly less physical space. The use of space or density of activity is one of the strongest predictors of environmental impact (C. Jones & Kammen, 2014). As discussed in Chapter 1, a tremendous share of public space within cities is designated for use by expensive, private automobiles. Reallocation of this space towards environmentally-friendly modes will begin to transform behavior, with compound effects as people continue to adopt micromobility (McFadden, 2007).

Transportation Networks

Much of the discussion thus far has been about electric scooters as a new mode within the transportation system in a competitive stance with existing options. While our analysis focused primarily on scooter data based on data availability, we must understand scooters as a small piece of the broader mobility system. Even the most dedicated scooter advocate does not envision a future where every trip is made on these devices (Clewlow, 2019). Rather, these devices have the potential to replace or complement portions of the service profiles of a number of different modes (Replogle & Fulton, 2014). The construction of new transportation infrastructure famously induces demand (Salvucci, n.d.; Young, 2015). While this is axiomatically known among transportation planners with respect to automobile infrastructure, the same appears to be true of bicycle facilities (Schmidt & Meyer, 2009; Xu, Wang, Wang, & Liu, 2019). The critical point in this discussion of demand is that while some trips may be reallocated from other modes or routes, a significant portion of this new use did not exist before the new policy. Scooters perform well in the Bunten and Rolheiser framework, hinting that continued adoption may be likely (Bunten & Rolheiser, 2019).

This research has primarily examined the direct impacts of electric shared scooters. The ways in which shared scooters interact with existing mobility networks will likely vary dramatically over time and across geographies. However, the supplemental and complementary effects of this new mode are evident in the survey responses from Portland (PBOT, 2019b), where respondents indicated how they would have made their most recent scooter trip if a scooter had not been available. 7.5% of respondents stated that they would not have made the trip at all, indicating the presence of unmet travel demand. Nearly 40% of respondents indicated they would have made the trip in a privately owned vehicle, a composite figure which includes transportation network company trips. 10% of respondents indicated they would have taken public transportation, and the largest individual category of respondents, 37%, indicated they would walk. This survey was conducted during and following the four-month pilot period of

scooter deployment in Portland, a context which diverges from full deployment in two relevant ways.

The first is the temporally-bounded nature of the pilot. While a four month period is not insignificant, there is evidence from other technology adoption studies that many people who may eventually adopt a technology would not have shifted over such a short period (El Zarwi et al., 2017; Zarwi, Vij, & Walker, 2017).

The Portland scooter pilot in 2018 limited providers to a total of 2,500 vehicles. It is unclear whether this cap directly limited the overall provision of scooter services. However, if shared electric scooters are to be treated like a piece of the transportation commons, in the same way that road access or public transit service are framed as generally available (caveats to this claim are covered in Chapter 1), a cap may increase scarcity, and change the spatial distribution of the good. Rather than increasing demand with a lowered supply, one might actually find a reduction in demand with scarce supply. If transportation is a unique good, a basic component of full societal participation (Cass et al., 2005), people will tend to select options from the transportation market that are seen as reliable. Consistent availability and predictability is a major factor that influences perceptions of reliability (Gehrke & Welch, 2019; Ma et al., 2019).

The reliable placement of docked bikeshare may be one factor that contributes to a use pattern that mirrors other forms of transportation, with a significant AM and PM commute peak. The presence of this peaked pattern for annual members only indicates that pricing consistency is another factor that may contribute to perceptions of reliability (Rosenfield et al., 2019).

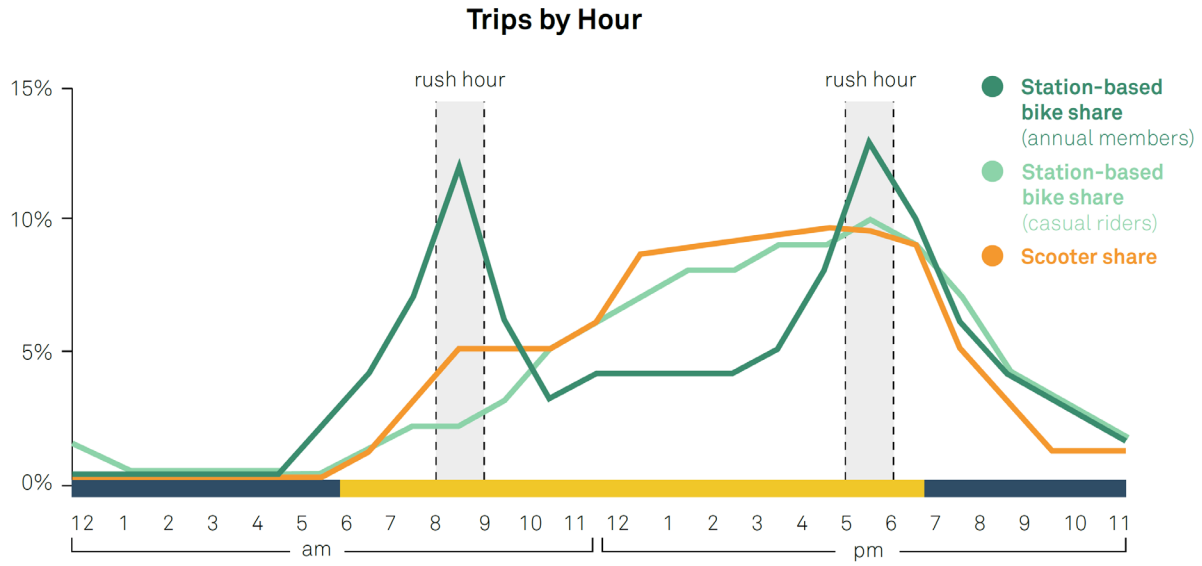

Figure 8.2 – Micro-mobility trips by hour, national composite data. From (NACTO, 2019)

Beyond the comparisons of scooters to other existing modes, we must recognize the role that scooters appear to playing in the broader mobility system. It is clear that electric shared scooters can and have begun to enhance the options and mobility of people across cities. In certain dimensions, including gender equity, racial identification, and wealth, they appear to be doing better than traditional bike share or private bicycles. The comparative nature of this research must be placed in context within a mobility system that is more than the sum of its parts.

In particular, the capacity of shared micro-mobility to function as a first mile / last mile connection is intriguing to policy-makers (Shaheen & Chan, 2016). No clear conclusions can be drawn from the research at hand with regards to first mile function, but the mode shift framework in Chapter 7 would provide significant insight into that question. Survey results from Portland indicate that while some people access scooters from transit or use scooters to access transit, this segment currently comprises about 6% of scooter trips (PBOT, 2019b). With long-term regular deployment, this would likely rise as more people incorporated them into routinized travel.

The policy considerations of these findings and this discussion comprise Chapter 9.

Limitations

This study provides interesting results but remains firmly exploratory. Data from electric shared scooters is new, and robust analysis methods are still being built up. There were a number of limitations on this analysis, which I will describe in the following section.

While MDS is a comprehensive data type, as of now it does not include any information about the user. This is obscured for privacy reasons, and makes sense for a public-facing data product. However, though Gust has User ID keys to link trips together by user, a decision was made to not provide an anonymized version of that key for research purposes. While this is a reasonable decision, it limited the types of traditional transportation analysis that could be undertaken. Most transportation analyses bundle bunches of transportation activity in an attempt to detect patterns at the individual level that aggregate to wide trends. However, we were limited to a spatial analysis, and could not engage directly with these modeling tools.

While the scooter trips in this analysis represent the complete universe of scooter activity on Gust during the study period (within the three cities), there were other companies operating in the same cities in 2018. Customer surveys indicated low amounts of brand loyalty, and while

there were some differences in the physical vehicle between brands, respondents typically selected those that appeared to be closest to their current location. In Portland, the city recorded 700,369 trips over the four month period (PBOT, 2019a). The trips we analyzed comprised only a portion of the total trips. While we would expect our findings to be fairly representative across scooter providers in the same city, there has been growth in other forms of dockless electric mobility that might have different characteristics. These are outside the scope of this analysis.

The spatial multivariate regression includes demographic factors as an input variable, with the assumption being that at least a large contingent of those who use scooters departing from a hex either live or work there. In fact, we have no guarantee that this is the case. However, this is an established method in transportation planning, from the four-step method onwards.

An additional limitation is the number of trips that were less than 100 meters long. While we provided those rates broken down by city and day of the week, many of these very short trips were not filtered with the rest of the outliers and may have thrown off the analysis if we intended to look only at trips that served as transportation. However, we felt that the inclusion of these trips was important, as each interaction with a scooter, even if the user does not end up using it for long, is an important piece of information, and we considered it as a trip.

Besides being disaggregated by time of day and week, the passage of time did not factor into our spatial analysis. This creates a potential confound, especially in areas where scooters arrived significantly later than others due to a growing fleet size. We collapsed these effects over many months. Future iterations of this work must employ time series analysis to understand how these effects change over time.

A limitation on the analysis of cost is important to note. As described in Chapter 3, MDS contains both a Standard Cost and Actual Cost field. However, in every data record received from Gust, these two fields contained identical values. This will lead to serious potential errors when calibrating an eventual mode shift model. There are two possible ways for users to pay below Standard Cost, one of which is quite prevalent. The first one is that users may be part of a low-income program, meaning they typically receive 50% off each ride. The second, much more common possibility is the use of a promotional code that provides either a free ride, or some amount of credit allowing for a free ride up a certain length. While very few individuals have signed up for the scooter fare equity programs, huge numbers of people have tried scooters using a promo code, which is not captured in the data. Specifically, if a trip is compared in an eventual multinomial logit model and with a listed cost of $4.30 and an actual cost of $0, this will throw of the calibration and the validity of the recommendations.

The case studies examined in this research were selected in part because Gust expected the city to be comfortable sharing ride-level data with an outside researcher. While the cities are relatively diverse in terms of size, demographics, and geography, caution must always be exercised when extrapolating findings from one geographic context to another.

Chapter 9 – Final Thoughts

In most states, cities are given control over streets within their boundaries, with the exception of state routes and interstates. This control is not unlimited, and is bound by commerce clause considerations and must comply with any state-level legislation or regulation. While cities had long been cautious about creating policy that may run contrary to the wishes of the state, the rise of ride-hailing served as a wakeup call in this regard. Many transportation departments have become proactive in regulating transportation, especially with relation to new or emerging modes. As ride-hailing proliferated, many cities began to levy a per-trip surcharge in an attempt to bring in additional revenue. Studies of ride-hailing have generally found that it increases overall Vehicle Miles Traveled (Chen et al., 2015; Shaheen, 2018). With the goal of reducing the environmental impacts of transportation, policy makers typically seek to reduce VMT (Nivola & Crandall, 1995). Given the impacts of ride-hailing, a per-trip surcharge acts as a Pigouvian tax to reduce the attractiveness of a behavior with negative externalities.

It is significantly easier to levy a tax on something new than to increase the tax rate for an existing tax. Both of these approaches are generally politically easier than implementing a tax on something that was formerly un-taxed (Gössling & Cohen, 2014; Millard-Ball, 2009). Scooter companies were new entrants into the urban transportation space, and their appeal as taxable entities was compounded by the prevalent coverage of stock valuations (Trefis, 2018). A strong argument was made that scooters relied entirely on the provision of public space and amenities for their business model, and it made sense for the city to recoup some of the value created using assets paid for with public funding. Notably, some scooter companies publicly agreed with the reasoning and called for a tax. However, they argued that if a tax was to be levied for use of public infrastructure, it should cover all users of that space, or at least all commercial users of public road space. Many have argued that pricing of road space is one of the most effective policy levers available to change the transportation paradigm (King, Manville, & Shoup, 2007; TfL, n.d.). This pricing can take many forms, from parking regulation to congestion pricing to

tolls. However, scooter companies argued that the tax should be scaled according to the extent of the direct negative externalities, including infrastructure wear and tear, and emissions.

Micro-mobility advocates were not successful in convincing any city to implement the comprehensive road pricing tax. However, many cities have now implemented a per-device fee that providers must pay to register each new device, or a per-trip surcharge, or both. While this has brought in revenue for cities, some have argued that, given the direction of the societal externalities of scooter use may be positive, this tax ends up discouraging behavior aligned with municipal transportation, sustainability, and equity goals. This debate should be informed by additional analyses of the externalities of scooter use, and must also consider non-transportation externalities such as sidewalk occupancy.

There has recently been a surge of bills introduced at the state level that would wholly or partially block cities from designing and implementing local rules regarding scooters. These preemption bills have been considered in 26 different states (NACTO, 2019). The creation of regulatory consistency for e-scooters is generally a good goal, especially given the current patchwork state of legality (Anderson-Hall, 2019). However, some of these bills limit the ability of cities to levy additional taxes. Another municipal power that may be limited under state bills is the ability of cities to impose permitting requirements that mandate certain types of activity, including targeted rebalancing.

From the outset of their pilot program, Portland mandated that each company was required to maintain at least 100 scooters in East Portland, a majority-minority community that has lower access to city transportation. East Portland emerged in our analysis as a regional Community of Concern. The rebalance regression models showed a small positive associative coefficient between the CoCI and rebalance locations in Portland (p. 88). This stands in sharp contrast to the significant negative effects of moderate magnitude that described the relationship between CoCI and rebalance locations in Nashville (p. 84) and San Diego (p. 92). We then note that the number of rebalance points in a hex was consistently the strongest positive predictor of the total count of trips originating in a hex. A scooter must be present for a scooter ride to start, but this simple explanation for the effect does not minimize the clear powerful mediating relationship between rebalancing and trip generation. Finally, we contextualize what sorts of trips occur by noting that in all three cities, hexes higher on the CoCI were associated with trips that were overall longer and more expensive. We assume people are using scooters in ways that serve their needs, and that these longer trips are somewhat reflective of greater unmet transportation needs (Boarnet, Giuliano, Hou, & Shin, 2017; Geurs & van Wee, 2004; Qian & Jaller, 2019). Therefore, I recommend that to address issues of transportation

inequity and access, cities should pursue policies that prioritize active rebalancing of scooters to historically marginalized neighborhoods.

This recommendation of rebalance requirements is in line with policies being pursued in a number of cities, including the other two cities in this study. However, the Nashville policy is significantly less exacting than the Portland policy, and appears to an encouragement rather than a requirement (Elrod & Allen, 2018). Mandated compliance is still under debate. San Diego has recently approved updated rules regarding scooter usage, which will take effect in June 2019 (Pell et al., n.d.). Oakland, a city not included in this analysis, uses Communities of Concern as its rubric for rebalancing. This policy has proved to be generally popular with residents, as well as effective for maintaining an appropriate distribution of scooters throughout the city. This policy is currently at risk of nullification through state preemption.

As cities and possibly states continue to explore potentially effective regulatory frameworks, I urge them to consider two major factors. First, implementation of a tax on micro-mobility may be politically popular, may yield short term revenues, and may be appropriate given the degree to which unregulated dockless mobility has overtaken public space in the past (Rahim Taleqani et al., 2019). However, the microeconomic ramifications of this tax on the uptake of what appears to be a positive contributor to the overall mobility system must be part of this revenue calculation. Second, rebalancing requirements are a highly effective policy lever to create capacity for broader access to this mode. Most providers already offer low-income programs, but enrollment may be arduous, especially if separate enrollment is required for each scooter provider. Rebalancing is another, complementary path to providing increased access in a way that places the effort required on the provider rather than the low-income resident.

If electric, shared micro-mobility can truly improve the capacity of marginalized communities to meet their mobility needs, prioritizing mode shift becomes not just an environmental and transportation goal for cities, but a social justice goal as well. This new mobility model has spread with unparalleled speed, and has the potential to radically reshape the ways in which people are able to access public space and the urban landscape. New technology is not politically neutral, and the associations and systemic frames that the shared electric scooter propagates are directly related to who is able to ride them and why they are used. Surveys reveal widespread interest across demographic boundaries; it is up to mobility providers and cities to ensure that micro-mobility reaches its full potential as a key component in a sustainable, flexible, and equitable transportation system.

Chapter 10 - Sources

A data standard for Mobility as a Service Providers who work within in the City of Los Angeles : CityOfLosAngeles/mobility-data-specification [Python]. (2019). Retrieved from https://github.com/CityOfLosAngeles/mobility-data-specification (Original work published 2018)

Anderson-Hall, K. (2019). 19-05267 Governing Micro-Mobility: A Nationwide Assessment of Electric Scooter Regulations. 18.

Arellano, J. F., Eng, P., & Fang, K. (2019). SUNDAY DRIVERS OR FAST AND FURIOUS: SPEED AND RIDER BEHAVIOUR OF ELECTRIC SCOOTER SHARE USERS IN DOWNTOWN SAN JOSE, CA. 9.

Ashford, F. N. A., & Caldart, C. C. (2008). 3 Economics and the Environment. 62.

Bartlett, A. A. (n.d.). REFLECTIONS ON SUSTAINABILITY, POPULATION GROWTH, AND THE ENVIRONMENT - REVISITED. 34.

Ben-Akiva, M. E., Lerman, S. R., & Lerman, S. R. (1985). Discrete choice analysis: theory and application to travel demand (Vol. 9). MIT press.

Bertrand, M., & Mullainathan, S. (2003). Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination (Working Paper No. 9873). https://doi.org/10.3386/w9873

Birtchnell, T., Harada, T., & Waitt, G. (2018). On the verge of change: Maverick innovation with mobility scooters. Environmental Innovation and Societal Transitions, 27, 118–128. https://doi.org/10.1016/j.eist.2017.11.002

Bishop, J. D. K., Doucette, R. T., Robinson, D., Mills, B., & McCulloch, M. D. (2011). Investigating the technical, economic and environmental performance of electric vehicles in the real-world: A case study using electric scooters. Journal of Power Sources, 196(23), 10094–10104. https://doi.org/10.1016/j.jpowsour.2011.08.021

Boarnet, M. G., Giuliano, G., Hou, Y., & Shin, E. J. (2017). First/last mile transit access as an equity planning issue. Transportation Research Part A: Policy and Practice, 103, 296–310. https://doi.org/10.1016/j.tra.2017.06.011

Boserup, E. (1981). Population and technological change: a study of long-term trends. Chicago: University of Chicago Press.

Braun, L. M. (2019). 19-05013 Widening the gap? Area-level associations between bicycle commuting, sociodemographic advantage, and bike lane access in 22 U.S. cities. 15.

Bunten, D., & Rolheiser. (2019, ongoing). A new classification of mobility.

Burns, L., Jordan, W., & Scarborough, B. (2013, January). Transforming Personal Mobility. The Earth Institute.

Cairns, S., Behrendt, F., Raffo, D., Beaumont, C., & Kiefer, C. (2017). Electrically-assisted bikes: Potential impacts on travel behaviour. Transportation Research Part A: Policy and Practice, 103, 327–342. https://doi.org/10.1016/j.tra.2017.03.007

Canepa, B. (2007). Bursting the Bubble: Determining the Transit-Oriented Development's Walkable Limits. Transportation Research Record, 1992(1), 28–34. https://doi.org/10.3141/1992-04

Cass, N., Shove, E., & Urry, J. (2005). Social Exclusion, Mobility and Access. The Sociological Review, 53(3), 539–555. https://doi.org/10.1111/j.1467-954X.2005.00565.x

Chang, C.-C., Wu, F.-L., Lai, W.-H., & Lai, M.-P. (2016). A cost-benefit analysis of the carbon footprint with hydrogen scooters and electric scooters. International Journal of

Hydrogen Energy, 41(30), 13299–13307. https://doi.org/10.1016/j.ijhydene.2016.05.168
Chen, L., Mislove, A., & Wilson, C. (2015). Peeking Beneath the Hood of Uber. Proceedings of the 2015 ACM Conference on Internet Measurement Conference - IMC '15, 495–508. https://doi.org/10.1145/2815675.2815681
Chertow, M. R. (2000). The IPAT Equation and Its Variants. Journal of Industrial Ecology, 4(4), 13–29. https://doi.org/10.1162/10881980052541927
Chmielewski, J. (2019). Transport Demand Model Management System. IOP Conference Series: Materials Science and Engineering, 471, 102068. https://doi.org/10.1088/1757-899X/471/10/102068
Clark, C. (1951). Urban Population Densities. Journal of the Royal Statistical Society. Series A (General), 114(4), 490–496. https://doi.org/10.2307/2981088
Clewlow, R. R. (2019). 19-03991 The Micro-Mobility Revolution: The Introduction and Adoption of Electric Scooters in the United States. 13.
Clewlow, R. R., & Mishra, G. S. (2017). Disruptive transportation: The adoption, utilization, and impacts of ride-hailing in the United States. University of California, Davis, Institute of Transportation Studies, Davis, CA, Research Report UCD-ITS-RR-17-07.
Clewlow, Regina R. (2018, July). The micro-mobility revolution: The introduction and adoption of electric scooters in the United States. Populus.
Conway, M., Salon, D., & King, D. (2018). Trends in Taxi Use and the Advent of Ridehailing, 1995–2017: Evidence from the US National Household Travel Survey. Urban Science, 2(3), 79. https://doi.org/10.3390/urbansci2030079
Coughlin, J. (1996). Tragedy of the Concrete Commons.
Coughlin, J. F. (1997). Can a Cabinet Official Innovate?: Entrepreneurial Politics and National Transportation Policy. Public Works Management & Policy, 1(3), 209–223. https://doi.org/10.1177/1087724X9700100302
Cronon, W. (1992). Nature-s metropolis: Chicago and the Great West (3. print). New York: Norton.
Daly, H. E. (1996). Beyond Growth: The Economics of Sustainable Development. Beacon Press.
Dennis, K., Colburn, K., & Lazar, J. (2016). Environmentally beneficial electrification: The dawn of 'emissions efficiency.' The Electricity Journal, 29(6), 52–58. https://doi.org/10.1016/j.tej.2016.07.007
Desai, U. (2002). Environmental Politics and Policy in Industrialized Countries. MIT Press.
di Prampero, P. E. (1986). The Energy Cost of Human Locomotion on Land and in Water.
Dill, J., & McNeil, N. (2013). Four Types of Cyclists?: Examination of Typology for Better Understanding of Bicycling Behavior and Potential. Transportation Research Record: Journal of the Transportation Research Board, 2387(1), 129–138. https://doi.org/10.3141/2387-15
Dill, J., & Rose, G. (2012). Electric Bikes and Transportation Policy: Insights from Early Adopters. Transportation Research Record: Journal of the Transportation Research Board, 2314(1), 1–6. https://doi.org/10.3141/2314-01
Dwyer, J., & Hu, W. (n.d.). Driving a Car in Manhattan Could Cost $11.52 Under Congestion Plan. 42.
El Zarwi, F., Vij, A., & Walker, J. L. (2017). A discrete choice framework for modeling and forecasting the adoption and diffusion of new transportation services. Transportation Research Part C: Emerging Technologies, 79, 207–223. https://doi.org/10.1016/j.trc.2017.03.004
Elrod, J., & Allen, B. (2018, August). Nashville | Second Substitute Bill BL2018-1202 (as amended). Retrieved May 23, 2019, from https://www.nashville.gov/Metro-Clerk/Legislative/Ordinances/Details/7d2cf076-b12c-4645-a118-b530577c5ee8/2015-2019/BL2018-1202.aspx

Ermagun, A., Lindsey, G., & Hadden Loh, T. (2018). Bicycle, pedestrian, and mixed-mode trail traffic: A performance assessment of demand models. Landscape and Urban Planning, 177, 92–102. https://doi.org/10.1016/j.landurbplan.2018.05.006
Favaretto, M., De Clercq, E., & Elger, B. S. (2019). Big Data and discrimination: perils, promises and solutions. A systematic review. Journal of Big Data, 6(1), 12. https://doi.org/10.1186/s40537-019-0177-4
Ferrara, M., Liberto, C., Nigro, M., Trojani, M., & Valenti, G. (2019). Multimodal choice model for e-mobility scenarios. Transportation Research Procedia, 37, 409–416. https://doi.org/10.1016/j.trpro.2018.12.210
Fishman, E., & Cherry, C. (2016). E-bikes in the Mainstream: Reviewing a Decade of Research. Transport Reviews, 36(1), 72–91. https://doi.org/10.1080/01441647.2015.1069907
Flamm, B., & Rivasplata, C. (n.d.). Perceptions of Bicycle-Friendly Policy Impacts on Accessibility to Transit Services: The First and Last Mile Bridge, MTI Report 12-10. 114.
Fuller, D., Gauvin, L., Kestens, Y., Daniel, M., Fournier, M., Morency, P., & Drouin, L. (2011). Use of a New Public Bicycle Share Program in Montreal, Canada. American Journal of Preventive Medicine, 41(1), 80–83. https://doi.org/10.1016/j.amepre.2011.03.002
Fyhri, A., & Fearnley, N. (2015). Effects of e-bikes on bicycle use and mode share. Transportation Research Part D: Transport and Environment, 36, 45–52. https://doi.org/10.1016/j.trd.2015.02.005
Fyhri, A., Heinen, E., Fearnley, N., & Sundfør, H. B. (2017). A push to cycling—exploring the e-bike's role in overcoming barriers to bicycle use with a survey and an intervention study. International Journal of Sustainable Transportation, 11(9), 681–695. https://doi.org/10.1080/15568318.2017.1302526
Garrard, J., Rose, G., & Lo, S. K. (2008). Promoting transportation cycling for women: The role of bicycle infrastructure. Preventive Medicine, 46(1), 55–59. https://doi.org/10.1016/j.ypmed.2007.07.010
Garrett, M., & Taylor, B. (2012). Reconsidering Social Equity in Public Transit. Berkeley Planning Journal, 13(1). https://doi.org/10.5070/BP313113028
Garrison, J. (n.d.). Bird scooters: Nashville seizes hundreds of electric scooters; Company to... https://www.tennessean.com/story/news/2018/06/07/bird-scooters-impo... 3.
Ge, Y., Knittel, C. R., MacKenzie, D., & Zoepf, S. (2016). Racial and Gender Discrimination in Transportation Network Companies (Working Paper No. 22776). https://doi.org/10.3386/w22776
Geels, F. W. (2012). A socio-technical analysis of low-carbon transitions: introducing the multi-level perspective into transport studies. Journal of Transport Geography, 24, 471–482. https://doi.org/10.1016/j.jtrangeo.2012.01.021
Gehrke, S. R., & Welch, T. F. (2019). A bikeshare station area typology to forecast the station-level ridership of system expansion. Journal of Transport and Land Use, 12(1). https://doi.org/10.5198/jtlu.2019.1395
Geurs, K. T., & van Wee, B. (2004). Accessibility evaluation of land-use and transport strategies: review and research directions. Journal of Transport Geography, 12(2), 127–140. https://doi.org/10.1016/j.jtrangeo.2003.10.005
Golub, A., Marcantonio, R. A., & Sanchez, T. W. (2013). Race, Space, and Struggles for Mobility: Transportation Impacts on African Americans in Oakland and the East Bay. Urban Geography, 34(5), 699–728. https://doi.org/10.1080/02723638.2013.778598
Gössling, S., & Cohen, S. (2014). Why sustainable transport policies will fail: EU climate policy in the light of transport taboos. Journal of Transport Geography, 39, 197–207. https://doi.org/10.1016/j.jtrangeo.2014.07.010
Graham, M. R., Kutzbach, M. J., & McKenzie, B. (2014). Design Comparison of LODES and ACS Commuting Data Products (No. 14–38). Retrieved from Center for Economic Studies, U.S. Census Bureau website: https://ideas.repec.org/p/cen/wpaper/14-38.html

Grengs, J., Levine, J., Qing Shen, & Qingyun Shen. (2010). Intermetropolitan Comparison of Transportation Accessibility: Sorting Out Mobility and Proximity in San Francisco and Washington, D.C. Journal of Planning Education and Research, 29(4), 427–443. https://doi.org/10.1177/0739456X10363278

Gross, Neil. (2018). Opinion | Is Environmentalism Just for Rich People? - The New York Times. The New York Times, 4.

Groth, S. (2019). Multimodal divide: Reproduction of transport poverty in smart mobility trends. Transportation Research Part A: Policy and Practice, 125, 56–71. https://doi.org/10.1016/j.tra.2019.04.018

Hall, J. V., & Krueger, A. B. (2016). An Analysis of the Labor Market for Uber's Driver-Partners in the United States (Working Paper No. 22843). https://doi.org/10.3386/w22843

Hallenbeck, S. (2012). User Agency, Technical Communication, and the 19th-Century Woman Bicyclist. Technical Communication Quarterly, 21(4), 290–306. https://doi.org/10.1080/10572252.2012.686846

Hallenbeck, S. (2016). Claiming the bicycle: women, rhetoric, and technology in nineteenth-century America. Carbondale: Southern Illinois University Press.

Hamre, A., & Buehler, R. (2014). Commuter Mode Choice and Free Car Parking, Public Transportation Benefits, Showers/Lockers, and Bike Parking at Work: Evidence from the Washington, DC Region. Journal of Public Transportation, 17(2), 67–91. https://doi.org/10.5038/2375-0901.17.2.4

Handy, S. L., & Clifton, K. J. (n.d.). Evaluating Neighborhood Accessibility: Possibilities and Practicalities. 12.

Hankey, S., Lindsey, G., Wang, X., Borah, J., Hoff, K., Utecht, B., & Xu, Z. (2012). Estimating use of non-motorized infrastructure: Models of bicycle and pedestrian traffic in Minneapolis, MN. Landscape and Urban Planning, 107(3), 307–316. https://doi.org/10.1016/j.landurbplan.2012.06.005

Hardt, C., & Bogenberger, K. (2019). Usage of e-Scooters in Urban Environments. Transportation Research Procedia, 37, 155–162. https://doi.org/10.1016/j.trpro.2018.12.178

Hekkert, M. P., Suurs, R. A. A., Negro, S. O., Kuhlmann, S., & Smits, R. E. H. M. (2007). Functions of innovation systems: A new approach for analysing technological change. Technological Forecasting and Social Change, 74(4), 413–432. https://doi.org/10.1016/j.techfore.2006.03.002

Hoffmann, M. L. (2016). Bike Lanes Are White Lanes: Bicycle Advocacy and Urban Planning. U of Nebraska Press.

HR&A. (2018, December). NYC E-Scooter Benefits Study.

Imbens, G. W. (1992). An Efficient Method of Moments Estimator for Discrete Choice Models With Choice-Based Sampling. Econometrica, 60(5), 1187. https://doi.org/10.2307/2951544

Intergovernmental Panel on Climate Change. (2018). Global warming of 1.5°C. Retrieved from http://www.ipcc.ch/report/sr15/

Jacobsen, R. B., & Eighmy, J. L. (1980). A Mathematical Theory of Horse Adoption on The North American Plains. Plains Anthropologist, 25(90), 333–342. https://doi.org/10.1080/2052546.1980.11909002

Jones, C., & Kammen, D. M. (2014). Spatial Distribution of U.S. Household Carbon Footprints Reveals Suburbanization Undermines Greenhouse Gas Benefits of Urban Population Density. Environmental Science & Technology, 48(2), 895–902. https://doi.org/10.1021/es4034364

Jones, T., Harms, L., & Heinen, E. (2016). Motives, perceptions and experiences of electric bicycle owners and implications for health, wellbeing and mobility. Journal of Transport Geography, 53, 41–49. https://doi.org/10.1016/j.jtrangeo.2016.04.006

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. American Economic Review, 93(5), 1449–1475.
Khan, S., & Kushler, M. (2013). Plug-In Electric Vehicles: Challenges and Opportunities. 46.
King, D., Manville, M., & Shoup, D. (2007). The political calculus of congestion pricing. Transport Policy, 14(2), 111–123. https://doi.org/10.1016/j.tranpol.2006.11.002
Krishnamurthy, V. (n.d.). Driven to Congestion: How the Planning, Engineering and Politics of Transportation Established, Preserves and Perpetuates the Automobile City. 167.
Kroesen, M. (2017). To what extent do e-bikes substitute travel by other modes? Evidence from the Netherlands. Transportation Research Part D: Transport and Environment, 53, 377–387. https://doi.org/10.1016/j.trd.2017.04.036
Krull, L. (2018). The land use and built environment factors impacting where women are using bikeshare in Boston.
Kysar, D. A. (n.d.). Climate Change, Cultural Transformation, and Comprehensive Rationality. 37.
Langston, S. (2019, May 21). Family of man killed on scooter wants them banned from Nashville. Retrieved May 22, 2019, from WKRN website: https://www.wkrn.com/news/local-news/family-of-man-killed-on-scooter-wants-them-banned-from-nashville/2015226802
Larsen, J., Student, M., El-Geneidy, A., & Yasmin, F. (n.d.). Beyond the Quarter Mile: Examining Travel Distances by Walking and Cycling, Montréal, Canada. 24.
Lawrence Livermore National Laboratory. (2019, March). United States Energy Usage Flowchart.
Le Vine, S., Lee-Gosselin, M., Sivakumar, A., & Polak, J. (2013). A new concept of accessibility to personal activities: development of theory and application to an empirical study of mobility resource holdings. Journal of Transport Geography, 31, 1–10. https://doi.org/10.1016/j.jtrangeo.2013.04.013
Lee, R. J., Sener, I. N., & Jones, S. N. (2017). Understanding the role of equity in active transportation planning in the United States. Transport Reviews, 37(2), 211–226. https://doi.org/10.1080/01441647.2016.1239660
Lewin-Koh, N. (n.d.). Hexagon Binning: an Overview. 12.
Lindsey, G., Maraj, M., & Kuan, S. (2001). Access, Equity, and Urban Greenways: An Exploratory Investigation. The Professional Geographer, 53(3), 332–346. https://doi.org/10.1111/0033-0124.00288
Litman, T. (2019). Implications for Transport Planning. 39.
Lowry, M. B., Furth, P., & Hadden-Loh, T. (2016). Prioritizing new bicycle facilities to improve low-stress network connectivity. Transportation Research Part A: Policy and Practice, 86, 124–140. https://doi.org/10.1016/j.tra.2016.02.003
Low-Stress Bicycling and Network Connectivity. (2017, November 8). Retrieved May 21, 2019, from Mineta Transportation Institute website: http://transweb.sjsu.edu/research/low-stress-bicycling-and-network-connectivity
Ma, X., Ji, Y., & Yuan, Y. (2019). 19-04504 Effects of Docked Bike-Sharing Fleets, Urban Land Uses and Socio-Demographic Characteristics on Dockless Bike-Sharing User Demand. 9.
Madigan, R., Louw, T., Dziennus, M., Graindorge, T., Ortega, E., Graindorge, M., & Merat, N. (2016). Acceptance of Automated Road Transport Systems (ARTS): An Adaptation of the UTAUT Model. Transportation Research Procedia, 14, 2217–2226. https://doi.org/10.1016/j.trpro.2016.05.237
Madison, J. (1787). The union as a safeguard against domestic faction and insurrection.
Malo, S. (2013, June 6). The Story of the Black Band-Aid. Retrieved May 20, 2019, from The Atlantic website: https://www.theatlantic.com/health/archive/2013/06/the-story-of-the-black-band-aid/276542/

Mansky, J. (n.d.). The Motorized Scooter Boom That Hit a Century Before Dockless Scooters. Retrieved April 30, 2019, from Smithsonian website: https://www.smithsonianmag.com/history/motorized-scooter-boom-hit-century-dockless-scooters-180971989/

Marqués, R., & Hernández-Herrador, V. (2017). On the effect of networks of cycle-tracks on the risk of cycling. The case of Seville. Accident Analysis & Prevention, 102, 181–190. https://doi.org/10.1016/j.aap.2017.03.004

Marqués, R., Hernández-Herrador, V., & Calvo-Salazar, M. (2014, September 23). Sevilla: a successful experience of bicycle promotion in a Mediterranean context. 769–781. https://doi.org/10.2495/SC140651

Marqués, R., Hernández-Herrador, V., Calvo-Salazar, M., & García-Cebrián, J. A. (2015). How infrastructure can promote cycling in cities: Lessons from Seville. Research in Transportation Economics, 53, 31–44. https://doi.org/10.1016/j.retrec.2015.10.017

Marqués, R., Hernández-Herrador, V., Calvo-Salazar, M., Herrera-Sánchez, J., & López-Peña, M. (2015, June 2). When cycle paths are not enough: Seville's bicycle-PT project. 79–91. https://doi.org/10.2495/UT150071

McFadden, D. (2007). The behavioral science of transportation. Transport Policy, 14(4), 269–274. https://doi.org/10.1016/j.tranpol.2007.07.001

McHugh, M. (2019, May 15). Scooter Season Is Here Again—but Will It Be Less of a Menace This Time? Retrieved May 23, 2019, from The Ringer website: https://www.theringer.com/tech/2019/5/15/18624332/e-scooters-transportation-portland-bird-lime-spin-micro-mobility-market

Millard-Ball, A. (2009). Cap-and-Trade: Five Implications for Transportation Planners. Transportation Research Record: Journal of the Transportation Research Board, 2119(1), 20–26. https://doi.org/10.3141/2119-03

Miskeen, M., Alhodairi, A., & Rahmat, R. (2013). Modeling a Multinomial Logit Model of Intercity Travel Mode Choice Behavior for All Trips in Libya. International Journal of Civil and Environmental Engineering, 7(9), 10.

Moody, J., Middleton, S., & Zhao, J. (2019). Rider-to-rider discriminatory attitudes and ridesharing behavior. Transportation Research Part F: Traffic Psychology and Behaviour, 62, 258–273. https://doi.org/10.1016/j.trf.2019.01.003

Moody, J., & Zhao, J. (2019). Car pride and its bidirectional relations with car ownership: Case studies in New York City and Houston. Transportation Research Part A: Policy and Practice, 124, 334–353. https://doi.org/10.1016/j.tra.2019.04.005

Moya-Gómez, B., & García-Palomares, J. C. (2016). The impacts of congestion on automobile accessibility. What happens in large European cities? ArXiv:1610.06829 [Cs]. Retrieved from http://arxiv.org/abs/1610.06829

NACTO. (2019). Shared Micro-Mobility in 2018.

Nivola, P., & Crandall, R. (1995). The Extra Mile: Rethinking Energy Policy for Automotive Transportation.

Nuñez, J. Y. M., Bisconsini, D. R., & Rodrigues da Silva, A. N. (2018). Combining environmental quality assessment of bicycle infrastructures with vertical acceleration measurements. Transportation Research Part A: Policy and Practice. https://doi.org/10.1016/j.tra.2018.10.032

Ostrom, E. (1990). Ostrom, E. 1990. Governing the Commons: The Evolution of Institutions for Collective Action. Cambridge University Press.

Pal, A., & Zhang, Y. (2017). Free-floating bike sharing: Solving real-life large-scale static rebalancing problems. Transportation Research Part C: Emerging Technologies, 80, 92–116. https://doi.org/10.1016/j.trc.2017.03.016

Papinski, D., & Scott, D. M. (2011). A GIS-based toolkit for route choice analysis. Journal of Transport Geography, 19(3), 434–442. https://doi.org/10.1016/j.jtrangeo.2010.09.009

PBOT. (2019a). 2018 E-Scooter Findings Report | The City of Portland, Oregon. Retrieved May 23, 2019, from https://www.portlandoregon.gov/transportation/78431
PBOT, P. B. of T. (2019b). 2018 E-Scooter Pilot User Survey Results.
Pearce, D. W., Atkinson, G., & Mourato, S. (2006). Cost-benefit analysis and the environment: recent developments. Paris: Organisation for Economic Co-operation and Development.
Pell, S., March 5, & 2019. (n.d.). Dockless bikes and scooters in San Diego get new rules in June. Retrieved May 20, 2019, from https://www.sandiegoreader.com/news/2019/mar/05/stringers-dockless-bikes-scooters-san-diego-rule/
Phillips, D. C. (2014). Getting to work: Experimental evidence on job search and transportation costs. Labour Economics, 29, 72–82. https://doi.org/10.1016/j.labeco.2014.07.005
Pierce, J. (2018). Digital Divide. In The International Encyclopedia of Media Literacy (pp. 1–8). https://doi.org/10.1002/9781118978238.ieml0052
Plazas, D. (2019). Our problem is not electric scooters like Lime, Bird, Lyft or Uber. It is Nashville's lagging infrastructure. | Plazas. Retrieved May 22, 2019, from The Tennessean website: https://www.tennessean.com/story/opinion/columnists/david-plazas/2019/02/22/scooters-lyft-uber-lime-bird-nashville-safe-roads/2927502002/
Plummer, B. (2018, November). Clean Energy is Surging.
Popovich, N., Gordon, E., Shao, Z., Xing, Y., Wang, Y., & Handy, S. (2014). Experiences of electric bicycle users in the Sacramento, California area. Travel Behaviour and Society, 1(2), 37–44. https://doi.org/10.1016/j.tbs.2013.10.006
Qian, X., & Jaller, M. (2019). STATION-LEVEL ANALYSES FOR BIKESHARE ACTIVITIES IN DISADVANTAGED COMMUNITIES. 5.
Quay, R. (2010). Anticipatory Governance: A Tool for Climate Change Adaptation. Journal of the American Planning Association, 76(4), 496–511. https://doi.org/10.1080/01944363.2010.508428
Rahim Taleqani, A., Hough, J., & Nygard, K. E. (2019). Public Opinion on Dockless Bike Sharing: A Machine Learning Approach. Transportation Research Record: Journal of the Transportation Research Board, 036119811983898. https://doi.org/10.1177/0361198119838982
Ramezani, S., Pizzo, B., & Deakin, E. (2018). Determinants of sustainable mode choice in different socio-cultural contexts: A comparison of Rome and San Francisco. International Journal of Sustainable Transportation, 12(9), 648–664. https://doi.org/10.1080/15568318.2017.1423137
Replogle, M., & Fulton, L. (2014, November). A Global High Shift Scenario. ITDP.
Rodrigue, J. (2017). The Geography of Transit Systems.
Rosenfield, A., Attanucci, J., Zhao, J., & Zhao, J. (2019). Evaluating Travel Demand Management Interventions Using a Randomized Controlled Trial at MIT. Transportation.
Sagoff, M. (2007). The Economy of the Earth: Philosophy, Law, and the Environment (2nd ed.). https://doi.org/10.1017/CBO9780511817472
Salvucci, F. P. (n.d.). • Industries • Customers • Infrastructure • Incrementalism • R&D • Competition. 16.
Schmidt, N., & Meyer, M. D. (2009). Incorporating Climate Change Considerations into Transportation Planning. Transportation Research Record: Journal of the Transportation Research Board, 2119(1), 66–73. https://doi.org/10.3141/2119-09
Services for working with MDS Provider data, built as runnable Docker containers.: CityofSantaMonica/mds-provider-services [Python]. (2019). Retrieved from https://github.com/CityofSantaMonica/mds-provider-services (Original work published 2018)

Shaheen, S. (2018). Shared Mobility: The Potential of Ridehailing and Pooling. In D. Sperling (Ed.), Three Revolutions: Steering Automated, Shared, and Electric Vehicles to a Better Future (pp. 55–76). https://doi.org/10.5822/978-1-61091-906-7_3

Shaheen, S., & Chan, N. (2016, Winter). Mobility and the Sharing Economy: Potential to Facilitate the First- and Last-Mile Public Transit Connections [Text]. https://doi.org/info:doi/10.2148/benv.42.4.573

Sheller, M., & Urry, J. (2006). The New Mobilities Paradigm. Environment and Planning A, 38(2), 207–226. https://doi.org/10.1068/a37268

Shu, L. L., & Bazerman, M. H. (n.d.). Cognitive Barriers to Environmental Action: Problems and Solutions. 27.

Soja, E. W. (2011). Seeking Spatial Justice. Seeking Spatial Justice, 37(111), 5.

Staeheli, L. A., & Mitchell, D. (2008). The people's property? power, politics, and the public. New York : London: Routledge, Taylor & Francis Group.

Stiegler, B. (2010). For a new critique of political economy. Polity.

Stockton, N. (2018, May). Why Nashville Voters Rejected Public Transit | WIRED. Retrieved May 22, 2019, from https://www.wired.com/story/nashville-transit-referendum-vote-plan/

Strimas-Mackey, M. (n.d.). Fishnets and Honeycomb: Square vs. Hexagonal Spatial Grids. Retrieved May 23, 2019, from http://strimas.com//spatial/hexagonal-grids/

TfL. (n.d.). Central London Congestion Charging - Impacts monitoring Fifth Annual Report, July 2007. 279.

Theen, A. (2019, March 25). E-scooters may return April 26, Portland says, with new regulations geared to curb sidewalk riding. Retrieved May 22, 2019, from oregonlive.com website: https://www.oregonlive.com/commuting/2019/03/e-scooters-may-return-april-26-portland-says-with-potentially-6-times-as-many-devices-allowed.html

Tietenberg, T. H. (1990). ECONOMIC INSTRUMENTS FOR ENVIRONMENTAL REGULATION. Oxford Review of Economic Policy, 6(1), 17–33. https://doi.org/10.1093/oxrep/6.1.17

Trefis. (2018). What Drives Value At Electric Scooter Sharing Startups Like Bird? 1.

United Nations, D. of E. and S. A. (2015). 2014 revision of the World Urbanization Prospects [Report]. Retrieved from (ST/ESA/SER.A/366) website: 2014-revision-world-urbanization-prospects.html

United Nations, I. S.-P. P. (2019, June 5). Global Asssessment Report on Biodiversity and Ecosystem Services. United Nations.

Ursaki, J., & Aultman-Hall, L. (2016). Quantifying the equity of bikeshare access in US cities. 95th Annual Meeting of the Transportation Research Board, Washington, DC.

Vale, D. S. (2015). Transit-oriented development, integration of land use and transport, and pedestrian accessibility: Combining node-place model with pedestrian shed ratio to evaluate and classify station areas in Lisbon. Journal of Transport Geography, 45, 70–80. https://doi.org/10.1016/j.jtrangeo.2015.04.009

Wachunas, J. (2019, March). New Lime Report Highlights Diversity Of Electric Scooter Riders. Retrieved May 22, 2019, from https://www.li.me/second-street/new-lime-report-highlights-diversity-electric-scooter-riders

Wang, C. (2019). The Effect of Demographic Characteristics on an Individual's Financial Risk Tolerance (SSRN Scholarly Paper No. ID 3257904). Retrieved from Social Science Research Network website: https://papers.ssrn.com/abstract=3257904

Wang, J., & Lindsey, Greg. (2019). The Effects of Neighborhood Socio-Demographic Characteristics on Bike Share Member Patterns of Use. TRB 2019 Annual Meeting.

Warner, S. B. (1978). Streetcar suburbs: the process of growth in Boston ; 1870 - 1900 (2. ed). Cambridge, Mass. London: Harvard Univ. Press.

Wasserman, D., Rixey, A., Zhou, X. (Elynor), Levitt, D., & Benjamin, M. (2019). Evaluating OpenStreetMap's Performance Potential for Level of Traffic Stress Analysis. Transportation Research Record: Journal of the Transportation Research Board, 036119811983677. https://doi.org/10.1177/0361198119836772

Waters, R. L., Lunsford, B. R., Perry, J., & Byrd, R. (1988). Energy-speed relationship of walking: Standard tables. Journal of Orthopaedic Research, 6(2), 215–222. https://doi.org/10.1002/jor.1100060208

Weiss, D. J., Nelson, A., Gibson, H. S., Temperley, W., Peedell, S., Lieber, A., … Gething, P. W. (2018). A global map of travel time to cities to assess inequalities in accessibility in 2015. Nature, 553(7688), 333–336. https://doi.org/10.1038/nature25181

Wild, C. (n.d.). The suffragette who rode a scooter in 1916. Retrieved April 30, 2019, from Mashable website: https://mashable.com/2015/06/15/1916-suffragette-scooter/

Williams, J. H., Haley, B., & Jones, R. (2015). Policy Implications of deep decarbonization in the United States.

Xu, C., Wang, Y., Wang, C., & Liu, P. (2019). Investigation of Contributing Factors to Travel Demand of Free-floating Bike Sharing: A Geographically Weighted Regression Approach. 8.

Young, J. (2015). Infrastructure: Mass Transit in 19th- and 20th-Century Urban America. Oxford Research Encyclopedia of American History. https://doi.org/10.1093/acrefore/9780199329175.013.28

Zarwi, F. E., Vij, A., & Walker, J. (2017). Modeling and Forecasting the Evolution of Preferences over Time: A Hidden Markov Model of Travel Behavior. ArXiv:1707.09133 [Stat]. Retrieved from http://arxiv.org/abs/1707.09133

Zhao, H., & Zhao, Y. (n.d.). TRAFFIC ANALYSIS ZONES - HOW DO WE MOVE FORWARD? 18.

Zureiqat, H. (2010). Baumol's Cost Disease in Public Transit: Historical Evidence and Future Implications. 17.

.1038/nature25181

Wild, C. (n.d.). The suffragette who rode a scooter in 1916. Retrieved April 30, 2019, from Mashable website: https://mashable.com/2015/06/15/1916-suffragette-scooter/

Zureiqat, H. (n.d.). Baumol's Cost Disease in Public Transit: Historical Evidence and Future Implications. 17.