EVALUATING ACCESSIBILITY OF FEMALE CAREGIVERS IN PHILADELPHIA

FROM A SAFETY PERSPECTIVE

Camille Madeline Boggan

A THESIS

in

City and Regional Planning

Presented to the Faculties of the University of Pennsylvania in
Partial Fulfillment of the Requirements of the Degree of

MASTER OF CITY PLANNING

2021

____________________
Advisor
Megan S. Ryerson
UPS Chair of Transportation
Associate Professor

Planning Thesis Studio Instructor
Francesca Russello Ammon
Associate Professor
Acknowledgements

When I began my time at the Weitzman School of Design, I had hoped to end the next step in my academic and professional career by writing a master’s thesis with a feminist perspective of transportation planning methods. My experiences watching mothers struggle on and off of buses with strollers, children, and shopping bags—before and after I moved to Philadelphia—have continued to inspire the way I approach planning and designing transportation systems. The completion of this paper would not have been possible without the generous support of the faculty and staff in the City and Regional Planning Department, my classmates and colleagues, and my family. I give my utmost thanks to Dr. Francesca Russello Ammon and Dr. Akira Drake Rodriguez for their assistance, guidance, and patience throughout the development of this research. To Josh Davidson and Carrie Long, who have served as second advisers, mentors, abundant sources of planning knowledge, and friends. To Dr. Lisa Servon, whose active commitment to recruiting a diverse pool of equity-informed students is the reason I pursued this degree at the University of Pennsylvania. To my classmates and colleagues, who continue to inspire me daily and many of whom have become unexpected friends throughout this journey. I give special thanks to Dr. Megan Ryerson, whose empathy, support, and unwavering belief in me has allowed me to complete this program despite many setbacks along the way. Finally, I give my greatest thanks and dedication to my family, who instilled creativity and a love in learning in me that is the foundation of everything I do.
Abstract

Transportation networks in major U.S. cities are built and managed around a very narrow subset of transit riders: able-bodied, solo men traveling to and from a white-collar job in the city center. Despite the evidence of women as primary users of public transportation, the U.S. planning field has lagged in incorporating this knowledge into practice. This thesis attempts to address this gap in transportation accessibility planning by evaluating the Southeastern Pennsylvania Transportation Authority (SEPTA) public transportation network in Philadelphia in terms of personal safety. The study employs a geographic information systems (GIS) analysis to illustrate how safety metrics could be incorporated into accessibility measurements as a practical method of evaluating transit networks from a female-centered perspective.

**Key words:** gender mobility, safety, feminist planning, accessibility, public transportation
Table of Contents

ACKNOWLEDGEMENTS ........................................................................................................... II

LIST OF TABLES & FIGURES .................................................................................................... V

I. INTRODUCTION .................................................................................................................. 1

II. REVIEW OF LITERATURE .................................................................................................. 3

  MEASURING ACCESSIBILITY IN TRANSPORTATION PLANNING ........................................... 3
  WOMEN, URBAN SPACE, AND SAFETY ................................................................................. 6
  TRANSPORTATION SAFETY MEASURES ............................................................................. 7
  DETERMINING ADDITIONAL SAFETY METRICS ................................................................. 9

III. METHODOLOGY ............................................................................................................... 10

  DATASETS ............................................................................................................................ 11
    AMERICAN COMMUNITY SURVEY 5-YEAR ESTIMATES 2013-18 ........................................... 11
    SPATIAL DATASETS ........................................................................................................... 11
  METHODS ............................................................................................................................. 13
    K-MEANS CLUSTERING ......................................................................................................... 13
    SPATIAL ANALYSIS ............................................................................................................. 13

IV. RESULTS ........................................................................................................................... 15

  K-MEANS CLUSTERING OF CENSUS TRACTS .................................................................... 15
  EVALUATING SAFETY MEASURES ....................................................................................... 20
    VISIBILITY ........................................................................................................................ 20
    PEDESTRIAN ENVIRONMENT ............................................................................................ 28

V. DISCUSSION ......................................................................................................................... 31
List of Tables & Figures

Table A: Datasets .................................................................................................................. 12
Table B: Results of k-means clustering model ..................................................................... 17

Figure 1 .................................................................................................................................... 17
Figure 2 .................................................................................................................................... 19
Figure 3 .................................................................................................................................... 21
Figure 4 .................................................................................................................................... 22
Figure 5 .................................................................................................................................... 24
Figure 6 .................................................................................................................................... 25
Figure 7 .................................................................................................................................... 27
Figure 8 .................................................................................................................................... 32
Figure 9 .................................................................................................................................... 32
Figure 10 ................................................................................................................................. 32
Figure 11 ................................................................................................................................. 32
I. Introduction

Transportation networks in major U.S. cities were historically built and managed around a very narrow subset of transit riders: able-bodied, solo men traveling to and from a white-collar job in the city center. Despite the rapid rise of female labor force participation beginning in the mid-20th century, resulting in a greater share of the female population using public transportation, transportation agencies have continued to plan their networks and service patterns primarily around peak-hour, solo commuters. The emergence of accessibility planning in the transportation field began to shift the paradigm from efficient movement to ease of reaching destinations. However, planning institutions in the United States still rarely include gender-specific data in their accessibility evaluations. Extensive literature in the field has documented that women all over the world use public transportation more than men, most often for non-work trips during off-peak hours (Hasson and Polevoy, 2011). Yet despite the evidence that women have become primary users of public transportation in cities, and that their travel patterns diverge from men’s, few U.S. transit agencies have practically included their specific needs in accessibility assessments.

The Los Angeles County Metropolitan Transportation Authority (LA Metro) published a groundbreaking report in 2019, “Understanding How Women Travel”, that examined the specific needs of women on their system. This report is the first of its kind from a transportation agency in the United States, emphasizing how far behind the U.S. transportation field is in normalizing gendered analyses for making policy decisions.
Findings in this report confirm anecdotal claims made by the many women who regularly use public transportation: they plan and change their trips largely because of safety concerns, cost, and ability to travel comfortably with small children. Safety in particular was cited as the number one barrier to riding public transportation, and most of the specific concerns were “center[ed] primarily around harassment and personal security” (p. 89). Reports investigating women’s mass transit use in the Global South echo these findings. A 2020 study by the Transport Gender Lab of the Inter-American Development Bank (IDB) found that 50% of public transit users studied were victims of an attack, and 31% of them have modified their routines for safety reasons. In Buenos Aires, Santiago, and Quito, researchers found that most users felt insecure on public transportation and had significantly higher levels of concern for their safety compared to men (FIA Foundation 2017). Compounding these concerns is the increased likelihood that a woman using public transportation is traveling with a child, due to continued trends of women taking on primary care roles in their households (Glynn 2018; Hasson & Polevoy 2011).

If a subset of transit users is changing routes or avoiding use altogether due to safety, it can be suggested then that a transit network or mode is inaccessible. However, gender-specific demographics and data related to safety are rarely considered in transportation accessibility assessments. The clear evidence that public transportation networks are not universally accessible for women (and children) should be of great importance to transportation professionals that seek to practice equitable planning. In Philadelphia, Pennsylvania, the Southeastern Pennsylvania Transportation Authority (SEPTA) has made few moves collect and analyze gender-specific ridership data, despite
reporting as having one of the highest female ridership bases among large transit agencies in U.S. cities in (Saksa 2015). According to a 2018 customer survey, SEPTA’s pre-COVID-19 ridership was 61% female and 39% male. SEPTA’s network may be lacking in its ability to adequately serve this large population of women, but conventional data collection and analyses are missing out on capturing these possible inequities.

This study seeks to examine the SEPTA network from the perspective of female transit riders by analyzing safety around transit stops that are likely to have a high percentage of female riders traveling with a child. I perform a spatial analysis to determine how indicators of safety can be operationalized, visualized, and thus incorporated into transportation accessibility evaluations. This paper contains five sections. Following the introduction, the second section reviews the existing literature on topics of measuring accessibility, transportation safety metrics, and gender-informed perspectives on safety and public space. The third section describes the methods of the current study and includes summaries of the data sources used; the fourth section summarizes the results for each analysis. Section five concludes the paper with key takeaways, implications of this research on policy and planning practice, limitations of the study, and an agenda for future research.

II. Review of Literature

Measuring Accessibility in Transportation Planning

Accessibility is most often defined as the ease of movement between places, or the ease of reaching destinations (Giuliano and Hanson 2017). The concept of planning for accessibility emerged in the late 20th century and has since grown to become a
mainstream topic for transportation planning and policymaking. Accessibility, as an abstract concept, evaluates the interrelationship of transportation systems and land use patterns, including the benefits or consequences of this interrelationship. For example, access to public transportation is associated with greater transit use, and good accessibility is associated with higher land values (Boisjoly and El-Geneidy 2017; Chen et al., 2018; El-Geneidy and Levinson 2006). These interdependencies are of great importance to transportation planners and policymakers, who use these models to determine system performance and forecast ridership.

Hansen (1959) introduced the concept of accessibility with his seminal piece, ‘How accessibility shapes land use’, in which he provides an empirical examination of the interdependence between transportation systems and land use patterns. Here he provides an early definition of accessibility—“the potential for interaction”—and an operationalized measurement in the form of the equation

\[ A_i = \sum_j O_j d_{ij}^{-b} \]

where:

- \( A_i \) represents the accessibility of person \( i \);
- \( O_j \) represents the number of opportunities at a distance of \( j \);
- \( d_{ij} \) is a cost measure between \( i \) and \( j \);
- and \( b \) is a distance decay function

4
This equation measures accessibility spatially by totaling the number of potential opportunities from a person’s home and determines how easily they can be reached by applying a cost measure, such as travel time or physical distance (Giuliano and Hanson 2017). Several accessibility formulas have subsequently been developed by transportation researchers and professionals, spanning many dimensions of accessibility and each with their own set of inputs. The major quantitative methodological categories of these measures include isochrone, gravity-based, and utility based. While isochrones and gravity-based methods provide simple outputs based on travel time or cost, utility-based methods (e.g., activity- or trip-based models) are disaggregate measures and can capture unique travel behaviors like trip-chaining (Dong, et al. 2006; National Academy of Sciences 2020).

To communicate these quantitative measures to policymakers, planners have developed accessibility indices to evaluate transportation systems and land use within defined geographic areas. An accessibility index is a comprehensive measure that incorporates a set of variables that represent access to or are relevant to a transportation network. These variables often include a combination of measures like service frequency, coverage, and wait times (Kerrigan and Bull 1992); demographic and land use data, such as job density and socio-economic characteristics (Ryus et al. 2000); and other spatial data unique to a network’s area of service. As a recent example, the metropolitan planning organization (MPO) for the Delaware Valley Region, the Delaware Valley Regional Planning Commission (DVRPC) evaluated transit access to essential services by vulnerable populations with an accessibility index. This index combined measures of
service frequency, walk time, number of essential services in an area of analysis, and number of destinations accessible by transit within 45 minutes (Delaware Valley Regional Planning Commission 2020).

Despite the growth in methodological approaches to transit accessibility planning, operationalized measures that included elements of safety or comfort were rarely encountered in the literature reviewed. Rather, research on index development for transportation level-of-service (LOS) has been the main arena of incorporating variables like comfort, safety, and even user perception (Das and Pandit 2014). A notable exception is the transit level-of-service (TLOS) index developed by Ryus et al. (2000), which includes a measure of pedestrian comfort in environments around transit stops. The authors of this study equate transit availability (“the opportunity to use transit service at a particular location”) with accessibility; more specifically, they assert that a safe and comfortable pedestrian environment is a critical metric of availability.

This study aims to acknowledge this large gap in methodological literature on accessibility planning and explore environmental measures of safety that are particularly important to women who may be traveling with children. In the following section, I will discuss how feminist and gender-informed academic research on urban space has documented the unique influence of safety on women’s experience in urban spaces, including public transportation.

*Women, Urban Space, and Safety*
Researchers across a broad scope of social sciences have documented the differing experiences of women and children in public space over time. Women have been denied access to public space throughout history, confined to domestic “spheres” of society that were separate from the public, urban spheres of men (Spencer-Wood 2006). As more women began to enter the workforce in Western cultures and occupy more public space, the inequities set into the built environment became alarmingly apparent. Emerging feminist perspectives on geography, urban theory, and planning from the 1970s century began to expose the heteronormative, gender-segregated nature of urban environments and the planning practices that created them (Beebeejaun 2016). Writings from the latter half of the 20th century continued to explore how gender influences the experience of public space, notably in the areas of mobility and perceptions of safety. For example, geographer Gill Valentine (1990) explored how social relations and power have greater influence over how women feel in a space than the actual design. Since then, a number of scholars and planning professionals have investigated the idea of planning “nonsexist city”, with the intent of gender equality (Hayden 1980). Despite the wealth of literature documenting the barriers and constraints women face in urban space, practical methods for addressing these inequities have been lacking in the planning field, especially in the areas of transportation.

Transportation Safety Measures

Assessing transportation safety has primarily been the domain of engineers in the transportation field. However, safety has become increasingly important to planners and policymakers as major cultural shifts have brought greater awareness to the connections
between public health, technology, personal safety, and transportation. In the broadest sense, safety in the U.S. transportation realm is primarily measured by documenting crashes, injuries and fatalities that occur on roadways or intersections. This measure, “killed or seriously injured” (KSI), is a critical number used to inform the installation of safety interventions for non-motorized users as outlined in the Manual on Uniform Traffic Control Devices (MUTCD). Created by the Federal Highway Administration, the MUTCD sets the federal design requirements for state and local transportation entities to follow. Killed or seriously injured statistics have been used in aggregate by transportation professionals to inform pedestrian safety policies, such as a Vision Zero High Injury Network map (City of Philadelphia 2017). However, KSI is a flawed metric on its own as it does not account for encounters between pedestrians, cyclists, and vehicles that were or felt unsafe but did not result in an injury or fatality (Federal Highway Administration 2017; Ryerson et al. 2020).

Loading, unloading, and road crossings to and from transit stops are where unsafe pedestrian interactions are most likely to occur (Cafiso, Di Graziano, and Pappalardo 2013). For example, Hedelin, Bunketorp, and Bjonstig (2002) revealed that in Goteburg, Sweden, three-fourths of those involved in public transportation incidents were injured at transit stops or in pedestrian crossings. Safety metrics (inclusive of KSI) have thus been developed to measure pedestrian safety, such as road crash risk exposure at the macro- and microscopic levels. At the macroscopic level, pedestrian road crash risk exposure has been estimated as the count of crashes or injuries per the population of pedestrians (European Road Safety Observatory 2008); as distance traveled walking (Lee and Abdel-
Aty 2005) or as the number of road crossings on a road segment (Papadimitrou, Yannis, and Golias 2012). Microscopic level analyses have proposed additional estimations, such as a combination of pedestrian and traffic volume at specific road sections and times (Lassare et al. 2007) or the count of vehicles encountered by pedestrians while crossing a lane. While this review was not exhaustive of all existing transportation safety metrics, it does highlight those commonly used and most relevant in relation to the current study. As noted, the transportation realm focuses largely on counts of crashes and injuries to evaluate safety for motorized and non-motorized travelers. Evaluating the road environment for pedestrians who will be walking to and from transit stops is necessary to determining network accessibility, and as discussed in the previous section, is critically important for women traveling with children on public transportation. However, additional safety metrics are needed in addition to crash and injury counts in order to gain a better understanding of other ways transit environments can be unsafe and inaccessible.

**Determining Additional Safety Metrics**

Harassment has emerged as one of the leading concerns of women regarding mobility in public space and on public transportation. As noted in the introduction of the current study, the Transport Gender Lab of the Inter-American Development Bank (IDB) found that 50% of public transit users they studied were victims of harassment. Unfortunately, sexual crimes in transit environments (such as staring, touching, and groping) ultimately go unreported (Ceccato 2017; Gekoski et al. 2015). Presence of benches, shelters, and lighting are elements related to improved feelings of safety commonly cited by female survey respondents in academic and institutional research on mass transit public safety.
(LA Metro 2019; FIA Foundation 2017; Transport Gender Lab 2020). Research examining perceptions of neighborhood safety of children and young adults also includes elements like lighting and the presence of vacant lots and buildings as indicators of safety and visibility, which impact the likelihood of traveling by foot or engaging in physical activity (Rossen et al 2012; Mota et al 2005).

Based on this literature, I compiled a list of indicators that could be obtained from a variety of open-source platforms to use for this analysis. The first set of indicators are largely centered around physical indicators of safety (or perceived unsafety). This includes the location of bus shelters or benches, the location of streetlights, and the location of vacant lots and buildings. The second set of indicators are related to the pedestrian environment and feelings of safety as a pedestrian. These indicators include the location of dangerous roadway segments, the location of bus shelters or benches, and “hot spots” of pedestrian crashes along road segments.

III. Methodology

The primary research method for this study is a spatial analysis using the programming software Python and geospatial analysis software ArcGIS Pro. I use demographic data from the 2013-18 American Community Survey 5-year estimates to perform a k-means clustering algorithm and geographically identify clusters of female-headed households with children and without access to a vehicle. As there is no existing disaggregate data available on female transit ridership in Philadelphia, identifying these clusters of tracts is a proxy estimate of transit ridership of women traveling with children.
I then layer spatial datasets onto these clusters using data from OpenDataPhilly (an online repository of open-source data for the Philadelphia region), to visualize environmental safety variables near transit stops, and explore why incorporating this type of data could be beneficial for future quantitative accessibility evaluations.

Datasets

American Community Survey 5-year Estimates 2013-18

Clustering algorithms work best with large spatial datasets of points. Thus, it is necessary to use the smallest geographic unit possible that the U.S. Census Bureau publishes with reliable sample data. I chose to use the 2013-2018 American Community Survey (ACS) 5-year estimates to provide the most accurate analysis within a set time period. One difficulty of using these datasets is accurately matching the variables (block group number, gender of householder, presence of children in household, access to a vehicle) between the correct tables. Block group-level data is also not as reliable as larger geographic units, such as tract, and therefore these datasets have larger margins of error. Using this 5-year estimate will allow for direct comparison in future research.

Spatial Datasets

Table 1 outlines the datasets used in the spatial analysis and their source. Each dataset was chosen based on the literature review examining women’s safety concerns in public space and on mass transit networks, as well as the availability of publicly available data. These variables were then grouped into two categories—visibility and pedestrian environment—as discussed in the literature review.
Table A

Datasets

<table>
<thead>
<tr>
<th>Data</th>
<th>Description</th>
<th>Source</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vacant Property Indicators</td>
<td>Polygon data of vacant lots</td>
<td>OpenDataPhilly</td>
<td>Visibility</td>
</tr>
<tr>
<td>Bus Shelters</td>
<td>Spreadsheet of bus shelter geocoordinates</td>
<td>SEPTA</td>
<td>Ped Env. /Visibility</td>
</tr>
<tr>
<td>Street Poles</td>
<td>Point data of active street poles (lights)</td>
<td>OpenDataPhilly</td>
<td>Visibility</td>
</tr>
<tr>
<td>High Injury Network 2017</td>
<td>Line data of High Injury Network street network</td>
<td>OpenDataPhilly</td>
<td>Pedestrian Environment</td>
</tr>
<tr>
<td>Philadelphia Crash Incidents</td>
<td>Point data of crash incidents in Philadelphia County</td>
<td>OpenDataPhilly</td>
<td>Pedestrian Environment</td>
</tr>
<tr>
<td>2013-2017</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEPTA Stop Ridership 2019</td>
<td>Point data of SEPTA transit stops</td>
<td>SEPTA GIS Portal</td>
<td>-</td>
</tr>
</tbody>
</table>
Methods

K-means Clustering

Clustering is an unsupervised machine learning method of data analysis that determines optimal groupings of points within large datasets. Data points within these groupings are similar to one another but divergent from points in other clusters. Clustering methods are useful for investigating the structure of large datasets, particularly data points with spatial representations (Na et al. 2010; Dabbura 2018). The k-means clustering algorithm is one of the most utilized clustering methods because of its simplicity and ease of implementation. It attempts to group data points closely together to minimize the sum of squared distance between data points. K-means requires a predetermined $k$ number of clusters, which greatly affects its performance; however, a heuristic like the elbow method can be used to find the best fitting $k$ number of clusters for a dataset (Yuan and Yang 2019). Spatial clustering is used often in transportation planning and geography to find patterns in spatial datasets using demographic and transportation-related data (Anderson 2009; Pulugurtha et al. 2003) To identify spatial locations of the target analysis group (female-headed households with children and without access to vehicles), I performed a k-means clustering algorithm in Python using the scikit-learn package and block group-level demographic data from the 2013-2018 American Community Survey 5-Year Estimates, queried from the United States Census Bureau API (Sander et al. 1996; Kriegel et al. 2011; Scikit-learn 2021)

Spatial Analysis
To analyze the distribution of street lighting in Philadelphia, I generated a kernel density estimation of the streetlight point dataset in ArcGIS Pro to visualize “hot spots” of street poles, which would represent areas of greater streetlight density. Kernel density estimation (KDE) is one of the most popular methods of spatially analyzing points events, also called point pattern analysis (PPA). The purpose of this method is to generate a smooth 2-D visual of point events over space, where the aggregation of points close in space is computed into a density estimation (Xie and Yan 2008). Estimating “hot spots” of point events is used widely in planning research, such as in analyses of traffic accidents (Erdogan et al. 2008; Anderson 2009), crime incidents around bus stops and shelters (Loukaitou-Sideris 2007), and more recently, traveler perception of transportation system performance (Pennetti et al. 2020). For this analysis I utilized the kernel density estimation tool in the Spatial Analyst Extension of ArcGIS Pro. I then overlaid layers of the major SEPTA transit lines and the census tracts from the k-means clustering analysis to compare streetlight density and public transportation routes between all tracts in Philadelphia County. For the individual tract analysis, I layered vacant lots and properties as well as all bus stops within a half-mile walking distance from the census tract to examine the transit environment. To analyze the pedestrian environment, I developed another “hot spot” estimation of pedestrian crashes using the Optimized Hot Spot Analysis tool in ArcGIS Pro. This tool creates a map of statistically significant “hot” and “cold” spots in a dataset of point events by using the Getis-Ord Gi* statistic (Esri 2020). Once a set of points indicating pedestrian crash hot spots was generated, I layered datasets relating to transit ridership, High Injury Network segments,
and bus shelter locations together to provide a fuller picture of the pedestrian environment.

IV. Results

K-means clustering of Census tracts

The results of the Moran’s I calculation are presented in Table A. Three variables of analysis—percent of family households with female householders and no husband present, percent of children in households, and percent of occupied housing units with no access to vehicles—were used for this calculation. The purpose of calculating Moran’s I is to determine if the variables of interest are spatially related before utilizing the means clustering algorithm. If the datapoints (census block groups) are not spatially autocorrelated within the dataset, the k-means model would not perform reliably.

Each variable was calculated with data from the 2013-2018 American Community Survey 5-year estimates at the census block group level. The calculation was performed using the PySAL v.2.4.0 package in Python. Interpretation of Moran’s I follows that of interpreting standard correlation coefficients. Values closer to +1 indicate positive spatial autocorrelation, values below -1 indicate negative spatial correlation, and values close to 0 indicate random spatial autocorrelation. Based on the calculation of Moran’s I, each variable of analysis exhibits moderately positive spatial autocorrelation. The percent of occupied housing units with no access to vehicles had the highest correlation coefficient of the three variables, suggesting households without access to vehicles are far more spatially autocorrelated than the remaining variables. As the coefficient of the Moran’s I calculation for each variable determined moderate spatial autocorrelation, the k-means
test performed moderately well at grouping together census block groups by the three variables of analysis. Figure 1 presents a map plotting the results of the k-means clustering algorithm. Colored polygons represent a block group in Philadelphia County; each block group is colored by cluster, ranging from 0 to 3. I chose to group the data into four clusters based on the results of an elbow plot, as discussed in the methods section.

The mean value for each variable of analysis within each cluster is shown in Table B. The geographic areas (block groups) this study is most interested in are those with the highest proportions of each variable of analysis. Cluster 0 had the highest mean value for presence of the three variables and was chosen as the cluster for further analysis. As the Moran’s I coefficient values were not greater than 0.6 for each of the variables, the clusters determined by the k-means model were not overwhelmingly well-defined and showed similarities across clusters. It is possible this is because the three variables of analysis are somewhat correlated with one another.

The Cluster 0 dataset includes 33 census tracts. A spatial dataset was created by merging the data from the Cluster 0 dataset with a dataset of all census tracts in Philadelphia County, queried from the U.S. Census Bureau API. The proportion of Cluster 0 block groups within a tract was calculated by dividing the total number of Cluster 0 block groups by the total number of block groups of those tracts. The final selection was made by selecting census tracts within the 4th quantile of the merged dataset’s distribution. These census tracts are outlined in red in Figure 3. These are census tracts where over 83.3% of their block groups are in Cluster 0. Like many of the block groups within Cluster 0, the majority of the selected census tracts are located in the
North and Northeastern Philadelphia neighborhoods. One is located in West Philadelphia, and two are located in Southwest Philadelphia. Although the spatial autocorrelation of the variables was moderate, and the k-means clusters were not perfectly defined, the selection process resulted in a clear clustering within geographic areas of Philadelphia County. It can be suggested that households with the variables of analysis are more likely to be located in the North, Northeast, West, and Southwest areas of Philadelphia.

**Table B**

Results of k-means clustering model

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Percent of family households with female householders, no husband present</th>
<th>Percent of children in households</th>
<th>Percent of occupied housing units with no access to vehicles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster 0</td>
<td>45.8%</td>
<td>43.2%</td>
<td>49.5%</td>
</tr>
<tr>
<td>Cluster 1</td>
<td>10.4%</td>
<td>24.6%</td>
<td>15.5%</td>
</tr>
<tr>
<td>Cluster 2</td>
<td>5.6%</td>
<td>10.4%</td>
<td>49.5%</td>
</tr>
<tr>
<td>Cluster 3</td>
<td>25.4%</td>
<td>39%</td>
<td>19.5%</td>
</tr>
</tbody>
</table>

**Figure 1**

Results of $k$-means clustering of Philadelphia block groups
Figure 2
Clustered Philadelphia block groups & final census tracts of analysis
Evaluating Safety Measures

Visibility

Figure 3 depicts a streetlight density estimation within the bounds of Philadelphia County, overlaid with the 2010 census tract boundaries. Streetlight density in Philadelphia County appears to be the most concentrated in Center City and South Philadelphia, as well in pockets throughout the Lower North, North, Riverwards, and West Philadelphia neighborhoods. Areas of moderate density include the Southwest, Northeast, and Northwestern areas of Philadelphia County. Large pockets of low density consist mostly of large parks or wooded areas. A majority of the selected census tracts (represented as dashed outlines) are located in areas of high-to-moderate streetlight density. Streetlight density among the major SEPTA transit lines—the high-speed Broad Street and Market Frankford rail lines, as well as the trolley routes—is varied.

Additional indicators to be evaluated alongside streetlight density are presented in Figure 4; these indicators include the location of covered bus shelters and transit stops within half-mile radius of the selected census tracts. The “Top Ridership” indicator in Figure 4 represents the top fifteen stops with the highest total boardings and alightings in both directions. Transit stops—including bus stops—within a half-mile of the selected census tracts are largely located in areas of moderate to high streetlight density. None of the top fifteen stops are located in areas of low streetlight density. Areas of concern include tracts located near large spots of low streetlight density in the southwest, northeast, and northwest. Bus shelters are sparsely distributed among stops within a half-mile of selected census tracts. The census tract furthest south has the greatest
concentration of bus shelters compared to all of the selected census tracts and is notably located in an area of low streetlight density. None of the stops identified as having the highest ridership of all selected census tracts overlap with a bus shelter point.

**Figure 3**
Kernel density estimation of streetlights in Philadelphia
Figure 4

Kernel density estimation with transit stops & shelters
Figures 5-7 present an overview of visibility indicators for the Upper North, Southwest, and West census tracts. In addition to the location of bus shelters and transit stops within a half-mile, I have also included data of the location of vacant lots and buildings. In the Upper North area of Philadelphia (Figure 5), streetlight density is fairly well distributed along street segments. Census tracts near areas of significant low streetlight density are located in the Hunting Park, Crescentville, and Germantown East neighborhoods. Land use in the Hunting Park and Crescentville census tracts are largely light industrial or commercial retail; the street poles dataset from OpenDataPhilly does not include commercial street poles, resulting in some areas presenting as low streetlight density. The tract in Germantown East is located near a wooded arboretum. Streetlight density among transit stops in this geographic area is varied; many stops are located in well-lit areas, while others, such as those in the Hunting Park census tract, are in “cold” spots of light density. This geographic area has very few bus shelters compared to the sheer count of stops that were selected for this analysis. Vacant lots are spread out across this analysis area; however, there is a large concentration of vacant lots near or in tracts in Germantown East and Upper Kensington with multiple transit stops nearby. As referenced in the literature, the presence of vacant lots near transit stops can impact perceptions of safety. Vacant lots represent “dark” or “blind” spots, even in areas of adequate street lighting or pedestrian activity.
Figure 5
Kernel density estimation with safety indicators for Upper North census tracts
Figure 6

Kernel density estimation with safety indicators for West Philadelphia census tracts
In the Southwest, both tracts are located in moderate-to-low areas of streetlight density. The tract furthest south, approximately located in the Eastwick neighborhood, appears to be in a particularly dark section of the city. However, this tract is located near a large shopping center with a surface parking lot. Lighting for commercial lots is not included in the street pole dataset used in this analysis. This area of analysis has a number of transit stops with bus surrounding both census tracts, mostly placed along Lindbergh Boulevard. Vacant lots are relatively prevalent in this area of analysis, particularly the large empty lot in the northern census tract. There are also a number of transit stops near these vacant lots. Compared to the Upper North and Southwest, the west census tract is located in an area of consistently high streetlight density (Figure 7). Even with the presence of large greenspace nearby, the environment around the census tract and the transit stops within a half-mile are likely to be well lit a majority of the time. The most notable observation about this area is the number of vacant lots and buildings within the tract itself and the surrounding transit service area. Along 52nd street, a major transit corridor for West Philadelphia, has a number of vacant lots and buildings located directly near transit stops. Only one bus shelter is located along this corridor. As this area of West Philadelphia is mostly dense residential, a consistent presence of vacant buildings could be perceived as unsafe for a female rider traveling at night. Although this area appears to
be relatively well lit, vacant areas can become blind spots at night and feel unsafe for those waiting at a transit stop for an extended period of time.

This analysis of visibility in relation to transit stops sought to answer how indicators like lighting and presence of vacant lots and buildings could be used to determine if stops on a transit network could be perceived as unsafe or inaccessible. The use of kernel density estimation for streetlight density was useful for looking at larger
patterns throughout Philadelphia County, but was not as useful for evaluating safety at the neighborhood level. A dataset of streetlight illumination may be more useful for looking at closer geographies and evaluating the quality of illumination in relation to transit stops. Incorporating the presence of vacant lots and buildings did provide some insights into targeting specific transit stops or hubs, and the potential of this metric is discussed in-depth in the conclusion.

**Pedestrian Environment**

Pedestrian safety is of critical importance to boarding and alighting transit users, especially those who may be traveling with small children. Figure 8 presents an overview of the indicators selected for pedestrian environment and safety. These indicators include High Injury Network (HIN) segments, the location of bus shelters, and pedestrian crash hotspots. Pedestrian crashes—vehicular crashes that involved one or more pedestrians—are largely concentrated in Center City as well as on North Broad Street and West Roosevelt Boulevard in the North and Northeastern areas of Philadelphia. Many of the identified high ridership stops are located on High Injury Network segments and intersect with a pedestrian crash hotspot. Bus shelters do not appear to have any particular pattern related to the prevalence of pedestrian crashes, and as noted previously, are sparsely distributed.

The multiple hotspots crashes on Broad and along Roosevelt are of most importance to this analysis, as they are located within a mile or less of census tracts in the current study’s focus. Figure 9 presents an overview of the pedestrian environment for the Upper North area of analysis. This area has a significant concentration of pedestrian
crash hotspots and HIN segments within a one-mile buffer around the selected census tracts. There are several high ridership stops along North Broad Street less than a quarter mile from a pedestrian crash hotspot; none of these transit stops have bus shelters. However, there are bus shelters present at another pedestrian crash hotspot further south on North Broad. Overall, the pedestrian environment of this area of analysis rates very poor for the average transit user due to the potential risk of exposure.

As we know from the clustering analysis, this area of Philadelphia is likely to have a high proportion of female transit users traveling with children. And as discussed in the literature review, it has been documented that women perceive safety differently than men in urban space and may perceive certain situations or environments as particularly unsafe. Thus, this geographic area is of critical importance to further investigate transit users’ perceptions of safety and how that impacts their ability to access transit.

In the Southwest area of analysis, pedestrian crash hotspots are not nearly as concentrated as in the Upper North area. Figure 10 presents an overview of the pedestrian environment for this area of analysis. The hotspot analysis identified intersection of Island Avenue and Lindbergh Boulevard in Eastwick as a location with substantial clustering of crash incidents between 2013 and 2017. This intersection is a large commercial retail center with large surface parking lots, and it serves as a termination point for the Route 36 Trolley. There are also many HIN segments in this area of analysis as well, notably along Island Avenue, a major arterial of Southwest Philadelphia. Two trolley stops are located at this intersection, directly overlapping with pedestrian crash hotspots; only one of these stops is reported to have a bus shelter. A terminal transit stop
surrounded by commercial uses (in this case, a shopping plaza with a supermarket) is of
great importance to our female rider analysis group. Literature has documented that
women are more likely to be using transit for non-work trips like shopping; and,
considering a census tract of analysis borders this intersection and stops, it is highly
likely that households who use transit in this area are shopping at this commercial retail
plaza.

Figure 11 presents an overview of the pedestrian environment for the West area of
analysis. Compared to the Upper North and Southwest areas, there are not as many
indicators of pedestrian environment to analyze. There is a very high proportion of HIN
segments in this area, and multiple top ridership stops at the intersection of two HIN
segments. Pedestrian crash hotspots are very sparse; only one falls within a half-mile
radius of the selected census tract. As discussed in the previous section of the results,
there are few bus shelters in this area. Although this geographic area appears to be “safer”
than others analyzed in this section, the presence of HIN segments is an important
consideration, even though it does not measure exposure.

This analysis of the pedestrian environment sought to answer how indicators
related to pedestrian safety could be used to determine if stops on a transit network could
be perceived as unsafe or inaccessible to female riders traveling with children. The varied
results of the analysis suggests that the pedestrian environment could be perceived as
very unsafe by the rider analysis group and experienced as a barrier to accessibility. Each
area analyzed, while located in vastly different parts of the city, faced similar issues of
potential crash risk due to hotspots and a lack of bus shelters at critical points. I found
this analysis more enlightening of the potential to incorporate pedestrian safety-related metrics, as there were clearer patterns of potential crash exposure and unsafe pedestrian walking environments.

V. Discussion

The spatial analyses in the current study provide an exploratory evaluation of transit safety for census tracts in Philadelphia with high proportions of female-headed households, with children, and without access to a private vehicle. First, the k-means clustering method used in this study has shown that the algorithm is useful for estimating transit ridership characteristics, and that Philadelphia block groups with households of similar characteristics (female-headed with no husband present, children present, and without access to a private vehicle) are somewhat moderately spatially autocorrelated.

For the geographic areas evaluated, I found that the transportation environment could be perceived as particularly unsafe by women traveling with children. I hypothesize this due to the presence of vacant lots near transit stops, the absence of bus shelters (especially at key locations), and the prevalence of pedestrian injuries and fatalities along corridors near census tracts of this population. The spatial analysis methods, as well as the selected data sources, used for this study could be useful in further constructing quantitative methods to evaluate the safety of environments surrounding transit stops, stations, and other infrastructure.

Although the nature of this study was exploratory, the results (and limitations) suggest several takeaways for planning practice. First is the need for transportation
agencies and planning institutions to expand the range of socio-demographic characteristics included when evaluating demand and accessibility. Including gender distributions of households served by transit would allow transportation planners to understand the needs of users more accurately within specific geographic areas, or with similar socio-demographic characteristics. The $k$-means clustering method used in this study has shown that the algorithm is useful for creating an estimate of transit ridership characteristics, and that Philadelphia block groups with households of similar characteristics (female-headed with no husband present, children present, and without access to a private vehicle) are somewhat moderately spatially autocorrelated. Utilizing a statistical method like $k$-means to cluster spatially related data is already in practice by planners; however, the combination of gender and age of household member variables is a particularly novel set of variables and proved useful in determining specific socio-demographic clusters. An example of this method being put into practice comes again from the LA Metro transportation agency. Planners at LA Metro created a Transit Equity Score to evaluate transit need based on socio-demographic indicators for its 2020 NextGen Bus Plan. One of the key indicators of transit need included the number of single mothers per land acre and the number of school children (aged 10-19) per land acre. Including these two variables alongside other typical socio-demographic variables used in accessibility indices led to the selection of Transit Equity Focused Areas, communities determined to have the greatest mobility needs (LA Metro 2020). As accessibility indices rely heavily on their inputs, adding meaningful demographic information will likely lead to the discovery of communities that have been underserved but overlooked due to perspective bias.
Figure 8
Overview of environment safety indicators in Philadelphia
Figure 9
Pedestrian environment safety indicators in Upper North census tracts
Figure 10

Pedestrian environment safety indicators in Southwest Philadelphia census tracts
Figure 11

Pedestrian environment safety indicators in West Philadelphia census tracts
Second is the need for planning researchers and professionals to reconsider the way safety is defined, operationalized, measured, and incorporated into transportation policies and metrics. The analysis of pedestrian environment was successful in illustrating how a combination of specified socio-demographic data and existing measurements can be used to perform more critical evaluations of transportation and pedestrian infrastructure. For example, I was able to conclude that the trolley stops in Eastwick at a specific intersection may be particularly unsafe for the female caregivers who are likely using transit and live within a mile radius. Further investigation into this particular area could lead to the development of inclusive safety interventions that may otherwise have been deprioritized for other geographic areas. Making a case to communities and policymakers for specific or expedited safety interventions could be strengthened by arguments relating to safety for the women and youth who use transit in this area. However, continuing to rely entirely on motorized and pedestrian crashes and injuries as a singular measurement of safety limits planners’ ability to evaluate transportation environments from different perspectives. The indicators of safety used in the current study provided only a limited understanding of how unsafe transit stops and the environments may be. The need to develop quantified safety metrics beyond KSI was a major takeaway from this research. These metrics do not necessarily need to be drawn from quantitative data collection methods; survey, ethnography, and interview responses can be coded and operationalized into quantitative data points or inspire the creation of new quantitative data sets.
Limitations of this study are primarily related to data accuracy and availability. First, each of the safety-related measures—visibility, infrastructure, and security—were chosen based on studies and surveys of riders on other transportation systems. It is possible there are safety concerns specific to women using SEPTA within the Philadelphia area that cannot be known without collecting that primary data. Therefore, the results of this study do not provide a complete picture of what the actual users of transit in this geographic area perceive to be as unsafe; they also do not provide a complete picture of what women traveling with children on SEPTA perceive as unsafe or inaccessible. Second, all of the data used in this analysis is from publicly available sources. There may be more complete or updated datasets that are not available to the public, which would alter the results of this analysis. For example, the spreadsheet containing the geocoded addresses of SEPTA bus shelters was last updated in 2015. There may have been more bus shelters installed throughout the city in the past six years. Finally, the disruption of travel patterns due to the COVID-19 pandemic is not modeled in this analysis. These limitations emphasize the difficulty of evaluating the safety (or rather, perceived safety) of spatial environments based on quantitative data alone.
References

Bhat, Chandra R., Bricka, Stacey., La Mondia, Jeffrey., Kapur, Aarti., Guo, Jessica.Y. and Sudeshna Sen (2006). Metropolitan Area Transit Accessibility Analysis Tool; TxDOT Project 0-5178-P3; Texas Department of Transportation: Austin, TX, USA.


https://doi.org/10.1177/0885412220927691


Los Angeles County Metropolitan Transportation Authority (2019). Understanding How Women Travel.

http://doi.org/10.3828/tpr.2010.38


