Anna Duan
February 18, 2020
URBS-428-301

Research Methods Memo

In this memo, I first introduce our research question and the population that we will study. Next, I describe the three main types of data that we will use and how they contribute to our study. Following this, I explain our exploratory approach and how we will investigate correlations between our selected factors. Next, in the research design section, I discuss our considerations in choosing to conduct quantitative research and I outline the accessibility measurement model that we selected. Finally, I conclude with a discussion of the assumptions and constants in our research design.

1. Research Question

In this study, we will explore the variations of accessibility for older adults in different census tracts and socio-demographic groups by measuring their accessibility to senior centers via public transit. We define older adults as individuals ages 65 and up, and we will study older adults residing within Philadelphia's city limits. We further specify the socio-demographic attributes in this study as income level, race, sex, disability, and car ownership. As we aim to study the potential correlation between older adults' socio-demographic attributes, residential address, and their accessibility on a citywide level, we will not select individuals as participants in our study. Rather, we will take advantage of open-sourced data (see next section, *2. Data Sources*) available at the census tract level to analyze each tract as a unit. That is, we will explore the correlation between each census tract's *share* of older adults with each of the aforementioned attributes, the tract's accessibility to senior centers, and the tract's geographical location.

2. Data sources

We will use the following data sources in our study:

1. General Transit Feed Specification (GTFS) data from Southeastern Pennsylvania Transportation Authority (SEPTA)
2. Population data from 2018 American Community Survey (ACS)
3. Geospatial data from OpenStreetMap.org (OSM).

The following table outlines each source, the information it contains, and how we will use it.

Table 1. Data Sources and Descriptions

Data Source	Data Type	What it is	How we will use
Southeastern Pennsylvania Transportation Authority (SEPTA)	General Transit Feed Specification (GTFS)	- Open-sourced data type created in 2005 to record transit scheduling, route, and location information - Temporal component: transit schedules (departure and arrival times) and sometimes real-time updates - Spatial component: geographic coordinates of transport stations and routes	We will use SEPTA GTFS data for trip time, route options, travel time consistency, and station location information to measure older adults' accessibility to senior centers.
American Community Survey (ACS), Administered by U.S. Census Bureau	Population and housing data	- Housing and population survey administered annually by U.S. Census Bureau - Includes social, economic, housing, demographic information at the census tract level and higher	We will use ACS data on race, sex, disability, income, and car ownership in Philadelphia census tracts to explore the correlation between these attributes and older adults' accessibility to senior centers.
OpenStreetMap.org (OSM)	Geospatial data embedded in a map	- Open-sourced map which includes road and sidewalk network data and geocoded building, road network, transport station data	We will export road and sidewalk network data from OSM and use ArcGIS to model the distances that individuals must walk to and between transit stations.

3. Correlates and Exploratory Approach

This study takes an exploratory approach to examining how the aforementioned factors correlate with accessibility, and how accessibility to senior centers varies geographically in Philadelphia. Rather than hypothesizing about the relationship between certain attributes of a census tract (e.g. percentage share of minorities) and the tract's accessibility to senior centers, we aim to explore the reasons for differences in accessibility between social groups and geographies without making assumptions. This is because we want to avoid being influenced by expectations about specific correlations in our analysis, and also because previous work has found that assumptions about socio-demographic attributes and transport disadvantage are not always accurate. For instance, in a 2014 study of accessibility to healthy food retailers via public transit in Cincinnati, the researchers found that contrary to their expectation, areas with high concentrations of African Americans and low-income individuals did not always score lower in accessibility to supermarkets via public transit (Farber et al., 2014).

While we refrain from assumptions about the correlation of socio-demographic attributes with accessibility via public transit, we selected several factors to study because they have been cited in accessibility literature to be correlated with increased transport barriers for individuals (Lubitow et al., 2017; Nordbakke, 2013). The factors we have chosen are physical disability, racial minority, female, and household income below the federal poverty line, in terms of percentage share of a given census tract.

4. Research Design

We have chosen to use quantitative research for our study. This is because we are interested in finding *citywide* patterns in the correlations between geographical location, sociodemographic factors, and accessibility to senior centers. Using the increasingly available GTFS and ACS data, we are able to examine at a high resolution the citywide variations in accessibility and its correlates, with the end goal of making policy recommendations for transportation and senior center planning. In our research design, we have contemplated several accessibility measurement models which have been used in previous studies. The following table describes each model we considered and outlines its pros and cons.

Table 2. Accessibility Measurement Models

Model	Main framework	Typical data inputs	Pros (as relates to our study)	Cons (as relates to our study)
1. Person-based	Analyzes accessibility from the perspective of the individual, taking into account individual-level constraints (e.g. time budget, physical ability, money) (Geurs & van Wee, 2004; Nordbakke, 2013; Ryan et al., 2015).	Travel surveys (Surveys of household and individual level travel behavior such as transit journeys during a day)	Effective for examining individual level differences in abilities and needs such to evaluate a transit system's equity implications and accessibility for the transport disadvantaged	- We cannot use this as we do not have person-based data such as travel surveys for older adults in Philadelphia. - This model is difficult to operationalize on a large geographical scale.
2. Gravity/Potential-based	Models the accessibility from a designated zone to all other zones in the study area by producing an accessibility index for each zone (Ozelet et al., 2016; Karner, 2018).	-Coordinates of each designated zone -Population data -GTFS -OpenStreetMap	-Recognizes individual travel preferences by lowering accessibility indices as distance between zones increases -Incorporates land use (e.g. spatial distribution of facilities) and transport service into analysis	- This model is also difficult to operationalize as it combines land use, transport, and various weights on opportunities and travel costs/impedances - It further does not offer a way to look at how changing time budgets (e.g. an increase from 15 to 30 minutes) affects access to desired opportunities

3. Utility-based	Analyzes transportation-related decisions in terms of allocating limited resources to maximize benefit. This model is used to calculate the expected benefit or cost accrued from a given travel choice or infrastructure project (Farber & Fu, 2017; Geurs & van Wee, 2004).	-Trip data -Origin and destination coordinates -Researcher-generated "desirability" index of trips	Effective for systems-level economic evaluation of accessibility (e.g. measuring the increase in rides available after a new transportation infrastructure project)	- Hard to communicate as it is deeply rooted in microeconomic theory - Limited applicability to transport equity studies like ours as it measures economic benefit while our study looks at accessibility of those excluded from traditional analyses of transport system efficiencies and costs.
4. Cumulative Opportunity -based	Based on the assumption that time constraints limit individuals' accessibility, this model measures the number of spatial opportunities available to an individual within a specified timeframe (Benenson et al., 2017; Farber & Fu, 2017; Karner, 2018).	-Time constraints (may come from researcher or from data on travelers' time budgets) -Origin and destination coordinates -GTFS	Highlights the effect of time budgets on accessibility, reflecting the often time-constrained transport needs of disadvantaged groups	- Timeframe used in studies can be arbitrary, and may not reflect the time budgets of transit users - This model also does not account for limitations to the opportunities or capacity at destinations

After weighing the pros and cons of the models, we choose the gravity model as the most appropriate for our study. It offers a clear way to account for the impact of distance on accessibility, as it assigns to calculations a cost-of-travel value, a cost-sensitivity parameter (to account for how much travel cost impacts accessibility of an opportunity), and an impedance function to account for traffic and other transport inconsistencies. Further, it aligns with the data sources available to us as it uses land-use and transport data to model individuals' spatiotemporal access to opportunities in every zone.

In our study, we plan to implement the gravity approach in a manner similar to Karner's 2018 method (Karner, 2018). In this study, Karner used U.S. Census Bureau Longitudinal Employer Household Dynamics (LEHD) data, GTFS data, and OpenStreetMap.org (OSM). He imported OSM data into ArcGIS to create pedestrian and transit networks to model trips between geocoded origins and destinations, and found an impedance term for transit trips based on the consistency of trip times. Using these models and an accessibility formula, he first calculated a territorial accessibility index of each zone in the study area, representing the accessibility of citywide opportunities to residents in the zone. Then, he integrated population data from each zone to create a weighted accessibility index which accounts for the types of opportunities (jobs, in his case) demanded by each zone's residents. We will build on this model by incorporating parts of the cumulative opportunity model so that we can understand the impacts of time constraints on accessibility of senior centers. Further, we have identified a number of assumptions and constants for our study, informed by previous work and the capabilities and needs of older adults. We organized these around two themes: needs-related and measurement-related.

Needs-related

Previous studies have assumed an average walking speed of 4-5 kilometers per hour for transit users. For our study, as we are studying older adults, this number should be lower to reflect the reduced physical abilities of this group (Nordbakke, 2013). However, given the differing physical conditions of what scholars have called the "young-old" and the "old-old" populations (Ryan et al., 2015), we may need to further separate our population into sub-categories with different capability assumptions for each.

Another need-related assumption is transit dependence. One study defined transit dependence as relying on public transport as one's primary mode of travel (Fransen et al., 2015). Guided by Fransen et al., we also define transit dependence such that even if an individual does not own a car, if they live within a 5-10 minute walk of necessary facilities, they are not transit-dependent. Our last needs-related assumption is the acceptable walking distance to transit stations. We will need this in our study as we are considering transit trips in whole, including the walk to and in between stations. In Karner's 2018 study (Karner, 2018), he follows the Federal Transit Authority (FTA)'s recommendation of 0.25 miles from bus stops and 0.5 miles from rail stations as an acceptable walking distance. This means that he only considered trips where riders are not required to walk further. Once again, as our study concerns older adults, these distances must be lower.

Measurement-related

The first measurement-related constant we recognized is the time threshold for these measurements. Some studies have studied accessibility of destinations at times throughout the day (Farber et al., 2014; Fransen et al., 2015). Our study will follow another method, where we measure accessibility within a timeframe (e.g. 9am-11am for peak hours or 11am-1pm for off-peak hours) during a given day (Owen & Levinson, 2015). This leads to our next constant, which is the time buffer. This is the amount of time used to represent an acceptable trip time (Farber & Fu, 2017). In Farber and Fu's 2017 study, only trips which last less than 60 minutes are considered. For our study, we will expand this concept to measure accessibility under multiple time buffers. This will show us how accessibility varies based on one's time budget (For example, we could make 3 accessibility calculations, using 30, 45, and 60 minute time buffers).

References:

Aging Population–Focused Accessibility Assessment of Multimodal Facilities in Florida—Hidayet Ozel, Eren Erman Ozguven, Ayberk Kocatepe, Mark W. Horner, 2016. (n.d.). Retrieved February 17, 2020, from https://journals.sagepub.com/doi/abs/10.3141/2584-07

Benenson, I., Ben-Elia, E., Rofé, Y., & Geyzersky, D. (2017). The benefits of a high-resolution analysis of transit accessibility. International Journal of Geographical Information Science, 31(2), 213–236. https://doi.org/10.1080/13658816.2016.1191637

Farber, S., & Fu, L. (2017). Dynamic public transit accessibility using travel time cubes: Comparing the effects of infrastructure (dis)investments over time. Computers, Environment and Urban Systems, 62, 30–40. https://doi.org/10.1016/j.compenvurbsys.2016.10.005

Farber, S., Morang, M. Z., & Widener, M. J. (2014). Temporal variability in transit-based accessibility to supermarkets. Applied Geography, 53, 149–159. https://doi.org/10.1016/j.apgeog.2014.06.012

Fransen, K., Neutens, T., Farber, S., De Maeyer, P., Deruyter, G., & Witlox, F. (2015). Identifying public transport gaps using time-dependent accessibility levels. Journal of Transport Geography, 48, 176–187. https://doi.org/10.1016/j.jtrangeo.2015.09.008

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

Karner, A. (2018). Assessing public transit service equity using route-level accessibility measures and public data. Journal of Transport Geography, 67, 24–32. https://doi.org/10.1016/j.jtrangeo.2018.01.005

Lubitow, A., Rainer, J., & Bassett, S. (2017). Exclusion and vulnerability on public transit: Experiences of transit dependent riders in Portland, Oregon. Mobilities, 12(6), 924–937. https://doi.org/10.1080/17450101.2016.1253816

Nordbakke, S. (2013). Capabilities for mobility among urban older women: Barriers, strategies and options. Journal of Transport Geography, 26, 166–174. https://doi.org/10.1016/j.jtrangeo.2012.10.003

Owen, A., & Levinson, D. M. (2015). Modeling the commute mode share of transit using continuous accessibility to jobs. Transportation Research Part A: Policy and Practice, 74, 110–122. https://doi.org/10.1016/j.tra.2015.02.002

Ryan, J., Wretstrand, A., & Schmidt, S. M. (2015). Exploring public transport as an element of older persons' mobility: A Capability Approach perspective. Journal of Transport Geography, 48, 105–114. https://doi.org/10.1016/j.jtrangeo.2015.08.016