

Plan Overview

A Data Management Plan created using DMPTool

DMP ID: <https://doi.org/10.48321/D10S8X>

Title: Transit dependents, choice riders, and service criticality: an analysis of the determinants of bus ridership in the Philadelphia Region

Creator: Erick Guerra - **ORCID:** [0000-0002-7769-2581](https://orcid.org/0000-0002-7769-2581)

Affiliation: University of Pennsylvania (upenn.edu)

Data Manager: Han Qiao, Andrew Tzuyuan Wang, Qi Si, Xuezhu Zhao

Project Administrator: Han Qiao, Andrew Tzuyuan Wang

Funder: United States Department of Transportation (DOT) (transportation.gov)

Grant: 69A3551747111

Template: Digital Curation Centre

Project abstract:

This report presents the results of two interrelated projects on transit ridership in the SEPTA region. The first focuses on developing and testing an empirically based theory of transit-dependency using a predictive model of transit mode choice in the Philadelphia region. The second examines station-level shifts in transit use in response to the Covid-19 outbreak in the Philadelphia region and how these correspond with the distribution of transit dependents throughout the region.

Part I

As local, state, and federal agencies began investing substantial resources into subsidizing transit in the 1960s and 70s, public documents argued that transit agencies should focus on attracting choice riders instead of dependent riders, who have no alternatives and use transit regardless of service quality. After six decades, the definitions, uses, and implications of the terms choice and dependent rider have remained consistent in the academic and professional literature. These definitions, however, lack a strong theoretical grounding or empirical evidence to support them. Using travel diary data from the Philadelphia region, I estimate discrete choice models to identify choice riders, who I define as those who have close to a 50% probability of choosing between a car or transit for a given trip. The Philadelphia region, which has a diverse range of transit users and transit services, is an ideal place to develop and fit an empirical model of choice ridership. Attributes assumed to be associated with dependent riders, such as a lack of a car, low income, and being a racial or ethnic minority, are much more prevalent among choice riders than the general metropolitan population. Choice riders are also diverse, with a mix of racial backgrounds, income levels, educational attainment, and access to private cars. Transit dependency, by contrast, is rare. The lowest and highest income residents generally only choose transit when service quality is high, and transit is cost- and time-competitive with the car.

Part II

The Covid-19 pandemic outbreak led to a long-lasting shock that jeopardizes health, economy, education, cultural and social activities around the world. People adjusted their travel behaviors according to new lifestyles as social distancing and travel restrictions were being implemented to prevent the spread of the virus. Huge declines in all modes of transportation were seen across the world and buses were among the most impacted modes of transportation. In this paper, we try to understand how different types of places lost bus ridership at the beginning of the pandemic by unpacking how the impact of demographics, socioeconomics and land use factors on bus ridership changed during the initial few weeks of the pandemic outbreak. We adopt Philadelphia as a case context and utilize a mixed-effect multilevel linear regression model to reveal the underlying correlations. Results show that factors negatively correlated with bus ridership (i.e. income, precipitation) became stronger in driving bus ridership and factors positively correlated with bus ridership (i.e. job accessibility, population density, parking cost, transfer station, weekday) became weaker after the pandemic outbreak. The result emphasizes the challenges that transit agencies face, especially during the immediate period after a society-wide change, and sheds light on future transit network planning and policies in providing more resilient and equitable travel mode choices in challenging times.

Start date: 07-01-2020

End date: 06-30-2023

Last modified: 07-12-2023

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customize it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Transit dependents, choice riders, and service criticality: an analysis of the determinants of bus ridership in the Philadelphia Region

Data Collection

What data will you collect or create?

We generated two different data sources for this plan. The first relies on publicly available household travel survey data matched to secondary Census sources as well as travel times, travel costs, and parking price estimates from the Delaware Valley Regional Planning Commission. We do not have permission to share these data skims publicly but will share random data samples removed of all identifying features to other researchers on request.

The second relies on proprietary data from two automated passenger counting firms with contracts with SEPTA. These data have been aggregated by week and by Census tract but cannot be made publicly available without explicit authorization from SEPTA. Random data samples removed of all identifying features will be provided to other researchers on request.

How will the data be collected or created?

Question not answered.

Documentation and Metadata

What documentation and metadata will accompany the data?

The report and paper include full descriptions and sources for the data used in this project.

Ethics and Legal Compliance

How will you manage any ethical issues?

There are no ethical concerns with the shared data. All data are anonymous and shared without identifiers.

How will you manage copyright and Intellectual Property Rights (IP/IPR) issues?

Only publicly available data and will be re-shared. Randomly sampled de-identified data will also be made available upon request.

Storage and Backup

How will the data be stored and backed up during the research?

The data are stored and saved on a folder in Box. I will also create a backup hard copy on a USB harddrive.

How will you manage access and security?

Data are only available to the research team.

Selection and Preservation

Which data are of long-term value and should be retained, shared, and/or preserved?

All data will be retained and preserved.

What is the long-term preservation plan for the dataset?

Storage on Box and in a harddrive.

Data Sharing

How will you share the data?

Randomly sampled and de-identified data will be shared upon request by other researchers.

Are any restrictions on data sharing required?

Yes, the traffic skims and the APC data are not publicly available and may not be shared.

Responsibilities and Resources

Who will be responsible for data management?

Erick Guerra

What resources will you require to deliver your plan?

No additional resources are required.
