Plan Overview

A Data Management Plan created using DMPTool

Title: Equity Effects of Rare Events on Transportation Network Company and Transit Riders

Creator: Destenie Nock

Affiliation: Carnegie Mellon University (CMU)

Funder: Mobility21 (Big Idea Proposal)

Funding opportunity number: 396

Template: CMU NSF Generic DMP

Project abstract:
In disadvantaged neighborhoods residents often must travel longer distances to and from transit stops, which creates challenges for human mobility. Transportation network companies (TNCs) (e.g., Uber) revolutionized mobility by detaching car access from car ownership. However, economic barriers to these services have limited mobility increases in disadvantaged neighborhoods. Here we will investigate the inequalities in TNC and public transit service across the demographic groups under normal and rare events, and create simulation tools to promote equitable mobility.

Start date: 07-01-2022

End date: 06-30-2023

Last modified: 02-24-2022

Equity Effects of Rare Events on Transportation Network Company and Transit Riders

---

Introduction

Many DMPs include an introduction. If your DMP includes an introduction, add it here.

This project will explore how TNC distribution and diffusion impacts equity under normal and rare conditions. This project is anticipated to generate data a wide variety of data that each have their own standards and needs. Given the interdisciplinary nature of this project, we will use a variety of data sources for each study, which we will now summarize.

Types of data produced

Types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project. Click on box size (small | medium | full) for detailed guidance.

Name	Data Type
Task 1: US Econometric Analysis	Public Archival Data
Task 1: Deep Dive City Data Analysis	Public Archival Data
Task 2: Qualitative Methods	Public Archival Data Participant Observation
Task 2: Simulation and Policy Assessment	Simulated Data

We now discuss each form of data in turn:

- **Public Archival data**. These data are from a variety of public sources
- **Participant Observations**. As part of understanding how infrastructure is integrated in each community, we may need to observe various meetings to understand who is involved, what tactics they use, and how effective these approaches are.
- **Simulated Data**. This is data that is extrapolated into the future based on anchoring in current data and/or statistically derived model estimates.

Data and metadata standards

Standards to be used for data and metadata format and content (where existing standards are absent or deemed inadequate, this should be documented along with any proposed solutions or remedies). Click on box size (small | medium | full) for detailed guidance.

For the public archival data, the formats will be those easily implementable in statistical programs such as R (CSV format) and STATA (DTA format). For participant observation, these will be recorded and saved as encrypted files of researcher observations or observations with informants and will be transcribed using

Microsoft Word software for ease of access.

Policies for access and sharing

Policies for access and sharing; Provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements. Click on box size (small | medium | full) for detailed guidance.

We will share public archival data electronically via a collaboration server, which will be housed at CMU. Initially, access will be limited to team members and will be protected with passwords, which will be changed regularly. Once data analysis is completed, we will make our cleaned versions of public U.S. datasets publicly available, insofar as they are in compliance with any data sharing policies that pertain to these public data sources. For participant observation data, we will work closely with CMU's Human Subjects Division to ensure that our protocols are IRB-approved and such data, when selectively released for publications or educational materials, is done so in an appropriately anonymized way. Any interview and observation data that contains individual-identifying information will only be kept on encrypted computers or hard disk drives with passwords changed regularly. We will preserve privacy by only sharing individual data in aggregate form. Moreover, the dissemination of results or any educational material based on these data will not reveal any such personal information.

We are committed to the free exchange of software and knowledge. As such, for the public archival and simulated data and resulting databases, we plan to make our code and all non-proprietary and non-privacy-sensitive data and coding for constructing the dataset available so scholars can replicate our analysis. We will make the dataset publicly available using (1) the journal that publishes the paper, when available and (2) our web server space at CMU, or by request if the dataset is too large for practical download via the Internet.

Therefore in short, within a reasonable time and request, our aim is for the public archival and simulated data to be shared with others following publication of the the findings derived from this data, so as to give project researchers the opportunity to benefit directly from their work.

However, for the qualitative data, these will be under IRB Human Subjects Review regulation which usually inhibits the dissemination of this data to preserve the anonymity and informed consent of the respondents involved. Therefore, we likely can only offer the aggregated insights that sufficiently protect anonymity and only after they appear in published work, so as to again give project researchers the opportunity to benefit directly from their work.

Policies for re-use, redistribution

Policies and provisions for re-use, re-distribution, and the production of derivatives. Click on box size (small | medium | full) for detailed guidance.

We discuss policies for re-use and redistribution in previous sections.

Plans for archiving & preservation

Plans for archiving data, samples, and other research products, and for preservation of access to them. Click on box size (small | medium | full) for detailed guidance.

We discuss plans for archiving data, samples, and other research projects, as well as their preservation, in previous sections. The only addendum to add is we plan to maintain encrypted access to these databases for at least 3 years following the completion of the project to ensure we can attend to reasonable requests for the public archival and simulated data, in line with data sharing policies that are inherent to such data. This is the mandated federal requirement for IRB-related activities (45 CFR 46.115(b)).

Software Sharing Plan

Some NSF solicitations require software sharing plans in the DMP. Please check with your specific solicitation for this requirement.

We discuss our software sharing plan in previous sections.

Planned Research Outputs

Data paper - "Transportation Network Company Ridership Sensitivity to COVID-19 Policies"

This paper will detail how COVID-19 impacted the service and supply of Uber and Lyft services in low and high income communities. Transportation Network Companies (TNC) like Uber and Lyft set out to fill gaps in transportation needs not filled by private vehicle ownership or public transit. We investigate how Covid-19 has shifted TNC ridership across community metrics like income and transit access because low income community areas may have more essential workers and fewer TNC alternatives like the ability to walk or bike to work. We measure the sensitivity of different populations' ability to respond to policies surrounding Covid (e.g., early stay-at-home orders) using a difference-in-difference regression. While the whole city receives the same stay at home order, some residents are more able to change their transportation behavior in response to the mandates. The treatment period represents the time after the stay-at-home order went into place and before TNC companies released their own Covid-19 guidelines for riding during May 10th-13th. High-income community areas showed a greater sensitivity to the stay-at-home order with a 51% greater decrease in TNC ridership during the stay at home order compared to low income community areas. This could indicate that high-income community areas are more able to adapt to stay-at-home orders because of the nature of their jobs or more access TNC alternatives like private vehicles and public transit.

Planned research output details

Title	Type	Anticipated release date	Initial access level	Intended repository(ies)	Anticipated file size	License	Metadata standard(s)	May contain sensitive data?	May contain PII?
Transportation Network Company Ridership Sensitivi ...	Data paper	2022-08-01	Open	None specified		None specified	None specified	No	No