Title:
Using ‘Big Data’ to understand the impacts of Uber on taxis in New York City

Published venue and year:
Travel Behaviour and Society
2020

Number of citations:
44 (on Google Scholar)

Year of publication:
2021

Reputation of the venue:
Travel Behaviour and Society is an interdisciplinary journal publishing high-quality original papers which report leading edge research in theories, methodologies and applications concerning transportation issues and challenges which involve the social and spatial dimensions.

Problem discussed:
Does Uber's expansion in New York City affect the well established taxi market in New York City. If so, to what extent?

Research questions addressed:
Does Uber's popularity affect taxi rides in New York City?
Methodology: Time-Series Regression Models
Data Analysis Techniques: Proxy for Uber Popularity, Control Variables, Data Aggregation

The changes in quality of taxi services, and whether this is influenced by Uber.
Methodology: Regression Analysis, Normalization
Data Analysis Techniques: Control Variables

Novelty of the contributions:
Methodological Innovations: They used media data as a proxy. This avoids Uber's proprietary data restrictions.he study using NYTimes articles counts as a proxy for Uber's popularity and public interest.

Multi-Dataset Integration: Combined taxi trip records, media data, weather data, subway disruptions, and economic indicators—a rare holistic approach in transportation studies. This is also an Empirical Contribution.

Relevance to Your Project:

This paper is consistent with my RQ3. The baseline method is Media Proxies for Ride-Hailing Popularity, and the use of NYTimes articles by newspapers is considered a proxy for Uber's popularity, which is consistent with our goal of collecting more objective data.

Complementary Findings:

The paradox of service quality: This paper finds that an increasing number of complaints actually indicate a change in behavior, such as higher customer expectations.

Pros and Cons Discussion:

Con1: Soundness

Relying on media data as a proxy for Uber's popularity can bring potential bias. As well, media reports may not reflect the actual number of passengers.

Con2: Verifiability

Due to proprietary UberMovement data and undisclosed preprocessing steps, the authenticity of the data is limited.

Con3: Presentation

Lack of dynamic visualization, which fails to better highlight demand hotspots and temporal changes in analysis (compared to dynamic visualization).