



How Frequent is Frequent Enough? Exploring the Relationship between the Approach Taken to Identify Ride-sourcing Users and the Specification of Ride-sourcing Adoption Models

INTRODUCTION

The growth in the popularity of ride-sourcing services (such as those offered by Uber, Lyft, and DiDi) has prompted numerous investigations into the use of these services.

- A fundamental component of these studies is the identification of survey respondents as either *users* or *non-users*
- Respondents have been identified as

- users based on:
 - Whether they have ever used ride-sourcing
 - Whether they have used ride-sourcing during a specific period

- How often they use ride-sourcing
- Despite the number of studies on ride-sourcing adoption and use, the impacts of the approach taken to identify users has not been explored.

Key research questions:

- Does the approach taken to identify users affect the final specification of models of ride-sourcing adoption?
- Does the approach taken to identify users affect the impact of a given factor on the probability of a person being a user?

STUDY AREA AND DATA DESCRIPTION

2 data sets for collaboration??

The study used data from two web-based surveys that were administered to a random sample of market research panel members who live in *Toronto* and *Metro Vancouver*.

- Metro Vancouver:
 - Survey on the Effects of Ride-sourcing in the Vancouver Region (SERV)
 - Ride-sourcing was introduced in 2020

- Toronto:
 - Survey to Predict the Repercussions of the Introduction of Novel Transportation network services (SPRINT)
 - Ride-sourcing was introduced in 2014

- The surveys collected information on:
 - Personal and household attributes
 - Use of ride-sourcing services
 - Attitudes and perceptions
 - Choices in stated preference experiments

use iterative proportional fitting to determine sample weights

Survey	Survey Period	Sample Size
SPRINT	May 2019	822
SERV	March – May 2022	1,851

METHODOLOGY

Ride-sourcing adoption among SPRINT and SERV respondents was modelled using binary logistic regression:

- The logged odds of person i being a ride-sourcing user is:

$$L_i = \ln\left(\frac{P_i}{1 - P_i}\right) = \beta_0 + \beta'x_i$$

Where:

- P_i is the probability of person i being a ride-sourcing user
- β_0 is the intercept
- β is a vector of parameters

x_i is a vector of explanatory variables
Three models were estimated using each dataset, each with a different approach for classifying respondents as ride-sourcing users:

- Model 1:** respondents who have ever used ride-sourcing are classified as users
- Model 2:** respondents who use ride-sourcing at least once per year are classified as users
- Model 3:** respondents who use ride-sourcing at least once per month are classified as users

Model 1 was treated as a benchmark because the approach for classifying respondents was broadest among the three models. For each dataset, the models were compared based on:

- Their final specification
- The 95% confidence interval of the parameter estimates

COMPARISON OF RIDE-SOURCING USE

Ride-sourcing frequency is relatively consistent among the two sets of respondents. Among users, at least once per month was the most common response.

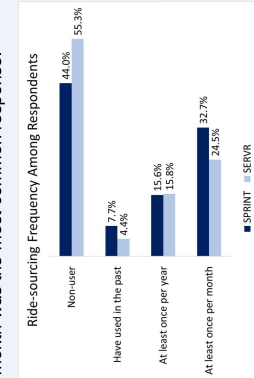


Figure 1: Comparison of ride-sourcing frequency

(Mitra et al., 2019): "Results of the ZINB model indicated that the determinants of adoption of on-demand ride-hailing services were different from the determinants of the frequency of use of these services among older adults."

did NOT use cluster analysis or latent class analysis so we did not have to make additional assumptions on the data and since I hadn't done it

FINAL MODEL SPECIFICATIONS

Table 2: Final specification of SERV models

Variable	Model 1: Has ever used ride-sourcing		Model 2: Makes >1 trip/year		Model 3: Makes >1 trip/month	
	Coef.	OR	Coef.	OR	Coef.	OR
Age (Yrs)	-0.023	0.98	-0.023	0.98	-0.023	0.98
Employed (full- or part-time) [0/1]	0.285	1.34	0.209	1.23	0.201	1.22
Employed (full- or part-time) [0/1]	-	-	[0.289, 0.302]	0.396	-	-
Master's degree or higher [0/1]	0.275	1.32	0.287	1.33	0.270	1.31
Has a driver's license [0/1]	0.792	2.21	0.774	2.10	0.891	2.44
Has a transit pass [0/1]	0.289	1.33	0.282	1.32	0.286	1.33
Has access to a bicycle (own or borrowed) [0/1]	0.250	1.28	0.244	1.27	0.244	1.27
Uses bike-sharing services at least once per year [0/1]	1.244	3.47	1.238	3.46	1.238	3.46
Uses car-sharing services at least once per year [0/1]	0.378	1.46	0.372	1.45	0.372	1.45
Household size	-0.079	0.92	-0.079	0.92	-0.079	0.92
Agrees ride-sourcing services would be a convenient way to travel [0/1]	0.718	2.05	0.711	2.03	0.711	2.03
Profers own car when making a trip [0/1]	0.476	1.61	0.469	1.60	0.469	1.60
Agrees new experiences are important [0/1]	-0.405	0.67	-0.412	0.66	-0.412	0.66
Enjoys trying new products and technologies [0/1]	0.283	1.33	0.277	1.32	0.277	1.32
Considers about safety when considering ride-sourcing services [0/1]	-	-	-0.416	0.66	-	-
Uses the internet or their smartphone to check when bus/train will arrive at least once per month [0/1]	0.287	1.80	0.280	1.80	0.280	1.80
Means of transportation (at least once per month) [0/1]	0.256	1.29	0.249	1.28	0.249	1.28
Uses the internet or their smartphone to choose a ride-sourcing service [0/1]	-	-	-	-	-	-
Shops online (at least once per month) [0/1]	0.216	1.24	0.210	1.23	0.210	1.23
Uses mobile payment technology (at least once per month) [0/1]	0.210	2.25	0.203	2.16	0.203	2.16
Uses food delivery services (at least once per month) [0/1]	-	-	-	-	-	-
Considers safety measures	-	-	-	-	-	-
AIC	165.693		151.351		124.286	
Nagelkerke's pseudo-R ²	0.363		0.401		0.401	
Nagelkerke's pseudo-R ²	0.258		0.364		0.364	
AUC	0.787		0.893		0.806	

Table 3: Final specification of SPRINT models

Variable	Model 1: Has ever used ride-sourcing		Model 2: Makes >1 trip/year		Model 3: Makes >1 trip/month	
	Coef.	OR	Coef.	OR	Coef.	OR
Age (Yrs)	-0.066	0.94	-0.066	0.94	-0.066	0.94
Full-time employed [0/1]	0.732	2.08	0.719	2.04	0.719	2.04
Has a driver's license [0/1]	0.524	1.69	0.524	1.69	0.524	1.69
Has a transit pass [0/1]	0.747	2.11	0.724	2.09	0.724	2.09
High income (> \$100k/year) [0/1]	0.457	1.58	0.443	1.57	0.443	1.57
Household size	-0.128	0.88	-0.132	0.87	-0.132	0.87
Lives in a house (single or semi-detached) [0/1]	-0.615	0.54	-0.626	0.53	-0.626	0.53
Number of vehicles	0.408	1.50	0.400	1.47	0.400	1.47
Lives with at least four students [0/1]	-1.655	0.19	-1.682	0.18	-1.682	0.18
Lives with at least one part-time worker [0/1]	0.339	1.40	0.336	1.39	0.336	1.39
Lives in Toronto-East York district [0/1]	0.415	1.52	0.403	1.49	0.403	1.49
Places importance on the safety when considering public transit [0/1]	0.233	1.26	0.231	1.25	0.231	1.25
Places importance on safety when considering public transit [0/1]	0.288	1.48	0.288	1.48	0.288	1.48
Places importance on parking availability when considering public transit [0/1]	-	-	-	-	-	-
Intercept	1.37	3.93	1.237	3.46	1.237	3.46
Considers safety measures	-	-	-	-	-	-
AIC	906.695		912.086		860.695	
Nagelkerke's pseudo-R ²	0.225		0.221		0.195	
Nagelkerke's pseudo-R ²	0.350		0.359		0.305	
AUC	0.803		0.806		0.792	

study results are useful especially to inform urban planners conducting basic descriptive analyses who need to define users and also for ADVANCE RESULTS frequency

The final specification of the adoption models differed based on approach taken to identify ride-sourcing users.

- Explanatory variables that were statistically significant in Model 1 were not significant in Model 2 and/or Model 3
- Model 2 and Model 3 include explanatory variables that were not statistically significant in Model 1

The impacts of explanatory variables on the odds of a person being a ride-sourcing user differ based on the approach taken to identify ride-sourcing users.

- Most parameter values in Model 2 and Model 3 were outside of the 95% CI of the values in Model 1
- Notable exceptions were observed for education, full-time employment, and the use of smartphones to check transit schedules in the SERV model

DISCUSSION AND IMPLICATIONS

Limiting adoption to so-called "frequent" users might cause the model to omit important information from users. Despite the similarities in explanatory variables, the approach taken to identify users can influence the specification of models of ride-sourcing adoption. Moving forward, studies on the adoption of ride-sourcing (and other emerging mobility services) should:

- Be mindful of the impacts of the approach taken to identify users on the results of the subsequent analysis
- Consider omitting those who use ride-sourcing sporadically (less than 1 trip/year)
- Develop criteria to identify users that are not overly restrictive

Based on the results of this analysis, it is recommended that those who use ride-sourcing at least once per year be the criterion for identifying users. broad enough to capture enough individuals while excluding individuals who are occasional users

- The approach taken to identify users can influence the specification of models of ride-sourcing adoption
- Criteria to identify ride-sourcing users should omit those who use these services sporadically
- Future studies on ride-sourcing adoption could develop separate models for frequent and infrequent users