

CISIL Report: Question 2

Joseph T. Ornstein & Suhan Kacholia

June 26, 2022

Data Challenge Question

How does the level and duration of incentive (i.e., fare subsidy) affect riders' long-term use of and enrollment in low-income transit programs?

Research Design

We compared transit riders who received a subsidized annual pass with those who received a stored value card. The study period is October 2020 through March 2022, following the reinstatement of fares in the King County Metro system.

Because the assignment of subsidies is not randomized, to estimate their causal effect we must rely on the (somewhat less satisfying) Selection on Observables assumption. We assume that potential outcomes are independent of treatment assignment conditional on the following confounding variables:

- Date when the card was issued
- Age of the cardholder
- Race of the cardholder
- Language spoken by the cardholder
- Median income in cardholder's Census tract
- Median age in cardholder's Census tract
- Percent white residents in cardholder's Census tract

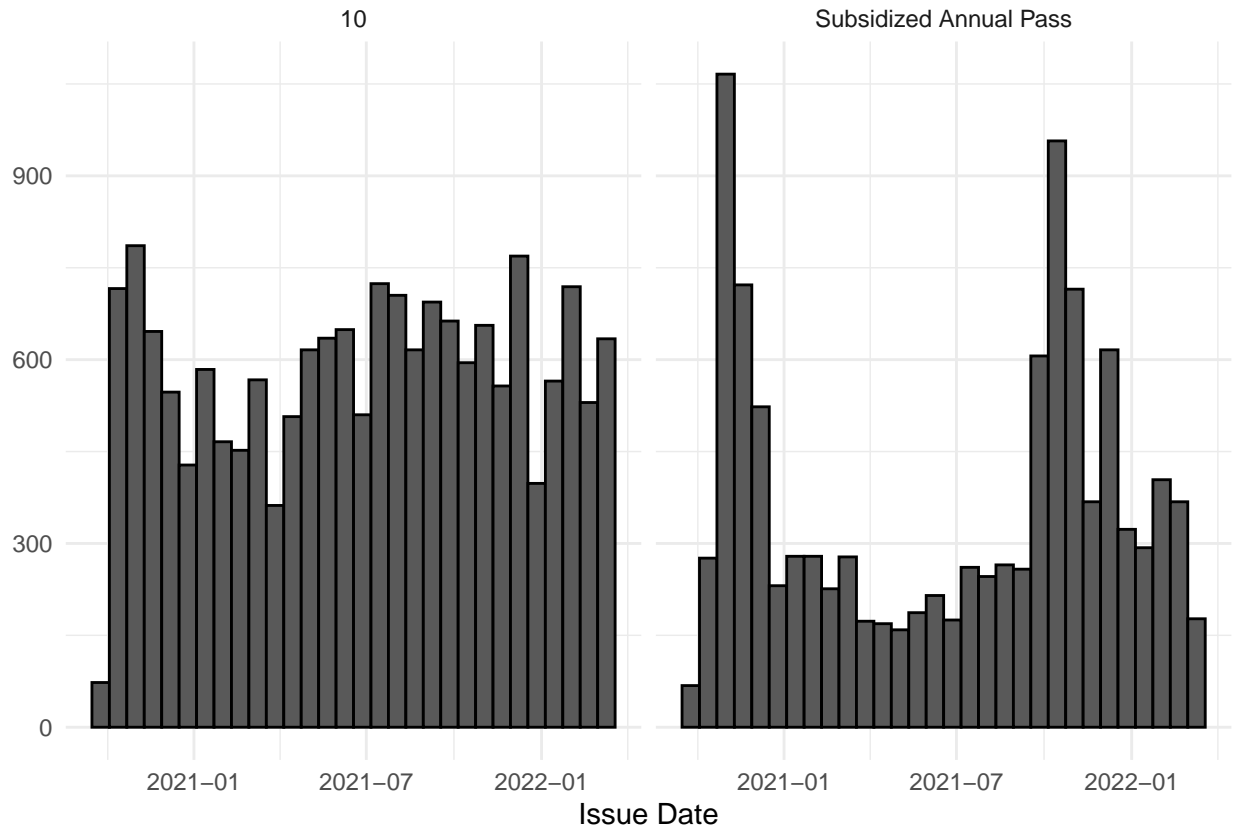
By matching on these characteristics, we can recover a more credible average treatment effect than a simple comparison of means.

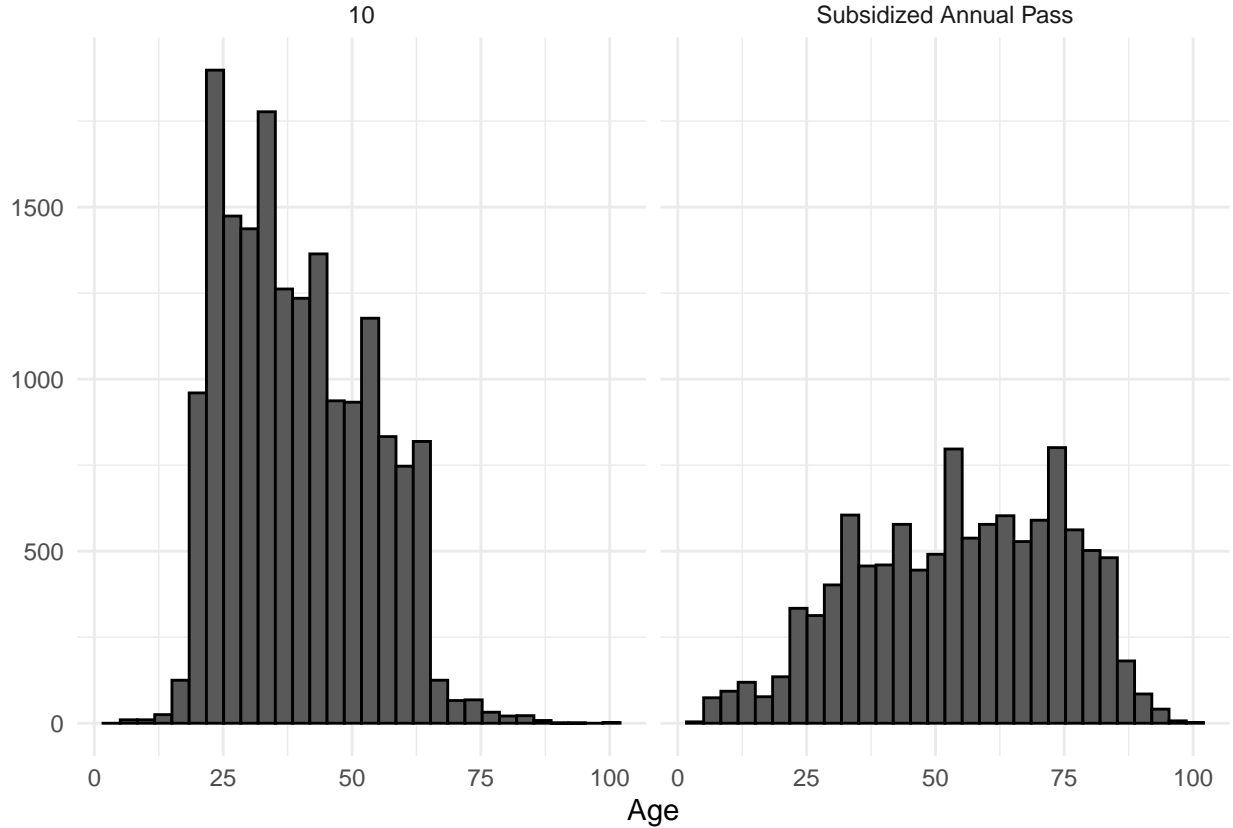
Supplementary Data

For each card ID, we compute two outcomes: the average weekly number of transit boardings, and the total dollar amount of individual sales (i.e. excluding organizational purchases) during the study period. For the independent variable, there were too few users in treatment conditions other than \$10 stored value and Subsidized Annual Pass, so we focus on those two levels. We also recode the language spoken variable to four levels: English, Spanish, Chinese, and Other. All code and data to replicate our results are available on our GitHub repository.

Methods

There are several significant covariate imbalances between the treatment and control groups. Riders who received a subsidized annual pass were significantly more likely to be senior citizens, less likely to speak Spanish, and their cards were more likely to be issued around the winter holiday season.





initial_load	Percent White	Percent Spanish-Speaking
10	37.37118	7.167943
Subsidized Annual Pass	28.54911	1.157769

To adjust for these imbalances, we create a matched control group using entropy balancing (Hainmueller 2012). This procedure selects a set of non-zero weights such that the first several statistical moments of the weighted control group’s covariates (listed above) are exactly equal those in the treatment group.

Here are the means of a selection of observed variables before and after matching.

Variable	Treated	Control (Before Matching)	Control (After Matching)
% White	26.177574	35.838691	26.177574
% Spanish-Speaking	1.000417	8.322618	1.000417
Mean Age	57.425872	39.720514	57.425872

Findings

After constructing the matched control group, we estimated the Average Treatment Effect on the Treated Group (ATT) using Weighted Least Squares. We find that riders who received a subsidized annual pass boarded transit roughly 0.54 more times per week during the study period than riders who only received a \$10 stored value card. These riders also spent \$123 less on passes and card loadings during the study period than the riders who received a \$10 preloaded card (as would be expected, since their pass allowed them to ride fare-free for a year).

	Weekly Boardings	Total Sales
(Intercept)	7.432 (0.077)	144.485 (2.156)
treated	0.538 (0.109)	−123.060 (3.049)
Num.Obs.	15 752	7106

References

Hainmueller, Jens. 2012. “Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies.” *Political Analysis* 20 (1): 25–46.