Harvard Data Science Review • Issue 7.3, Summer 2025

Navigating Progress: Enhancing Public Transit for More Equitable Communities via Interpretable Causal Inference

Gaurav Rajesh Parikh[1] Jenny Huang[2] Albert Sun[3] Lesia Semenova[4] Cynthia Rudin[1]

[1]Department of Computer Science, Duke University, Durham, North Carolina, United States of America,

[2]Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts, United States of America,

[3]Epiq, New York, United States of America,

[4]Department of Computer Science, Rutgers University, Piscataway, New Jersey, United States of America

The MIT Press

Published on: Jul 31, 2025

DOI: https://doi.org/10.1162/99608f92.1476e26a

License: Creative Commons Attribution 4.0 International License (CC-BY 4.0)

ABSTRACT

Public transit can be vital in providing mobility for low-income communities. King County Metro, which serves the greater Seattle area, is one of the largest transit providers focused on equitable public transit. They have rolled out innovative subsidy-based incentive programs for low-income riders with the goal of increasing equitable access to public transit, and in turn improving access to jobs, housing, education, and other essential resources. However, due to the observational nature of data on ridership behavior, the effectiveness of such innovative policies is difficult to evaluate.

Our work proposes an interpretable pipeline for policymakers to evaluate the impact of transport subsidies on long-term ridership and reenrollment patterns. We identified and combined seven data sets from a variety of sources, including individual-level enrollment and ridership data, geographic data on access to jobs and transit, and demographic data, to perform observational causal inference to evaluate one of King County Metro's largest programs implemented in 2020: the subsidized annual pass (SAP). We used FLAME, a recent interpretable machine learning–based causal inference matching method, to construct high-quality matched groups and identify features that were important predictors of ridership and reenrollment.

Applying our pipeline to the data shows that the SAP is effective in increasing long-term reenrollment and ridership. We observe that treatment effects are more pronounced in the Asian population, especially the Chinese-speaking population, in individuals ages 65 and older, and individuals with better access to jobs and public transit. While increases in ridership are modest and may be impacted by factors such as COVID-19 effects, nonrandom selection into treatment or hidden confounding, the key insight is that we observe heterogeneity in the outcomes for different subpopulations. Insights from this work can be used to inspire improved future study designs to study public transportation policy decisions and their impact on different population groups.[1]

Keywords: interpretability, causal inference, transportation, mobility, machine learning

Media Summary

Enhanced access to public transit can significantly improve quality of life by enabling access to better paying jobs, better health care, schools, and other resources. This is especially vital for low-income communities that may heavily depend on public transportation for their daily commutes. King County in the greater Seattle area is one of the most innovative public transit providers that has been taking a data-centered approach to improve and increase transit access and ridership. Specifically, they collected individual-level data on ridership and transit pass enrollment, as well as provided subsidies to people below a certain income level. However, since a controlled experiment was not performed, it was difficult to evaluate the effectiveness of their innovative

policies on enhancing uptake of public transit. We studied ridership and reenrollment (two important aspects that measure the success of public transit) for one of King County Metro's largest programs implemented in 2020—the subsidized annual pass (SAP)—by combining data sets from a variety of sources (including access to jobs, transit, and demographics). We constructed matched groups of individuals who were similar in relevant ways, but were different in which transportation incentive they were offered. This mimics a controlled experiment under the standard assumptions of causal inference that each unit has an equal chance to be part of the treated or control group and that our matching variables can explain away the confounding and allow us to assess the effect of the SAP program. We found that SAP is effective both in increasing reenrollment and long-term ridership (ridership after 1 year of receiving the SAP incentive). For purposes of designing new subsidy programs under budget constraints, policymakers may be interested in knowing which subpopulations benefited most during an incentive program, in order to prioritize these individuals for future programs. To that end, we note that the largest estimated effects occurred in populations that have higher access to public transit and jobs. The estimated effects are also more pronounced in the Asian population and in older individuals above the age of 65. While lots of external factors might affect our study results, including effects of COVID-19, seasonal changes, and missingness in data on ridership and reenrollment, making our findings challenging to validate short of being able to conduct a real experiment, our work aims to 1) showcase that it is possible for different effects to exist for different target populations and 2) inspire further research, including randomized controlled trials, to identify policy impacts and promote more equitable public transportation policy decisions. Our analysis provides a foundation for policymakers to understand how transport subsidies impact long-term ridership patterns.

---

1. Introduction

Access to high-quality affordable transportation is important for making city resources more equitable and promoting fair access to job opportunities, education, and other essential services (Jenkins et al., 2020). Recognizing this, King County Metro created the One Regional Card for All - Low Income Fare Transit (ORCA-LIFT)reduced-fare program, a program that offers a variety of transit subsidy incentives to King County Metro users with income below 200 % of the federal poverty level (Constantine, 2020). As part of the ORCA-LIFT program, Metro introduced the Subsidized Annual Transit Pass (SAP) in October of 2020. These passes were made available to all residents in the three-county region with income at or below 80% of the federal poverty level and who are enrolled in certain state cash benefit programs (Constantine, 2020).

An important objective for King County Metro has been to understand whether the SAP program would enhance access to resources and opportunities, making urban transit more equitable for individuals in the lowest income bracket. However, forming and evaluating policies to enhance equity is not easy. Challenges in obtaining quality data where causal effects can be disentangled make it difficult to decide on and allocate financial resources for future public transit initiatives (Alonso González et al., 2022).

Nonetheless, the problem of identifying and quantifying the effect of public transportation interventions is so important that the Causal Inference for Social Impact Lab (CISIL) at Stanford University and King County Metro created a set of open-ended data challenges, inviting scholars to disentangle complex transit data with a broadly defined goal to identify interventions that can have the highest impact in improving equity in public transportation (Center for Advanced Study in the Behavioral Sciences [CASBS], 2022). The CISIL Data Challenges were aimed as a novel approach to 'crowdsource' research on increasing equitable commuting so that consolidated findings can be used to make recommendations to policymakers.

To quantify causal effects, one would ideally conduct a randomized controlled trial (RCT), where selection into the treatment group is randomized (Fisher, 1992). However, RCTs are often logistically challenging and expensive to run. Instead, King County released a large amount of *observational* data. Within this study, all residents in the three-county region with income at or below 80% of the federal poverty level and who are enrolled in certain state cash benefit programs were eligible to receive the treatment, an SAP (Constantine, 2020). The eligible resident then needed to enroll, either online or at a card pick-up location, in order to claim the pass. Such self-selection into treatment introduced biases that make it difficult to answer any causal questions. One common approach to address causal questions using observational data is to match treated and control units in a way that ensures that the underlying background covariates are balanced (Chapin, 1949). If one also chose to match units exactly on raw covariate values, the treatment effects within each matched group can be interpreted as conditional average treatment effects. The added insights from exact matching approaches for causal inference (in contrast to response-surface-modeling based approaches) give a more granular causal analysis, providing policymakers with information about which subpopulations benefited the most from treatment (T. Wang et al., 2021). Being able to identify the subpopulations that benefited the most from a public policy initiative can be used to inform strategies for future resource allocation.

2. Building the Interpretable Pipeline

2.1. Data Sets Curation and Description

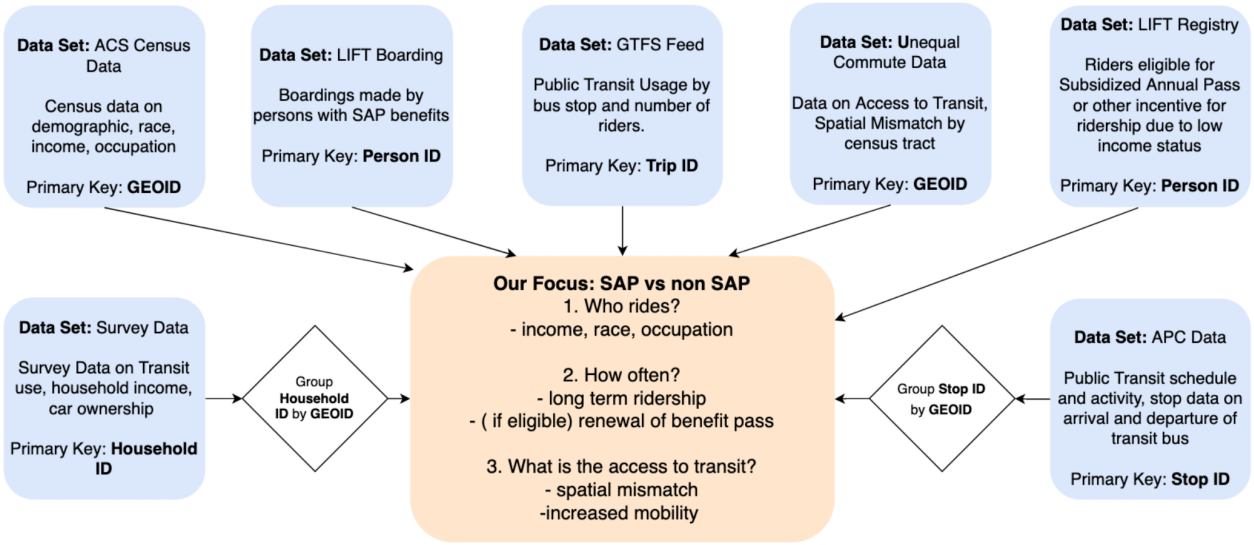


Figure 1. Data set descriptions and questions that we focus on in this work. Note that Person ID uniquely identifies each eligible person, Stop ID is defined as the unique ID for every bus stop or point from where transit may be available, Trip ID is defined as one trip made by a bus along a specific route, with outbound and inbound trips having unique Trip IDs. Household ID identifies household associated with each person. GEOID refers to the geographic location or census tract group to which a person's household belongs.

Given that our goals were open-ended, we first decided to find a variety of data sources that might help us address broad questions pertaining to access to transit and long-term outcomes. We assembled and connected data sets from a variety of sources including survey data, census data, and automatically collected transit data in the form of General Transit Feed Specification (GTFS) data on public transit usage. Please see Figure 1 for a summary of the data sets. Our data sets are useful for insights on *demographic information*, *ridership and reenrollment*, *and unequal commute*.

Demographics: We used the following data sets to understand the demographics of King County Metro users:

1. *American Community Census Data.* We use census tract data from the American Community Survey 5-year Data report from 2019 to look at statistics such as income and race for all county tracts in King County. This data is open source, and can be obtained from the U.S. Census Bureau (n.d.).
2. *Spatial Data.* We use spatial catalog data from King County Metro. This data is from a variety of sources on income, car ownership, transit use, household workplace location, and so on. This data is open source, and can be obtained from King County GIS Center Spatial Data Catalog (2019).
3. *Automatic Passenger Counting (APC) Data.* This contains public transit schedules and activity, as well as stop data including arrival and departure of transit buses. This data can be requested from the King County Metro/CISIL Data Challenge organizers.

Ridership and Enrollments: We used the following data sets to understand how riders use the King County Metro:

1. *ORCA-LIFT Card Registry.* ORCA-LIFT is a program created by the King County Metro to give different subsidy incentives to King County Metro users. This data set gives information on incentives, including monetary benefits and the SAP. There are 98,228 unique individuals in this data set. We considered users who enrolled for a SAP October 1, 2020, and April 1, 2021, as these were the individuals eligible, during the period of our study, for reenrollment for a SAP, which occurs annually. This data can be requested from the King County Metro/CISIL Data Challenge organizers (CASBS, 2022).
2. *ORCA-LIFT Lift Boardings*. This data set records the boardings made for ORCA-LIFT users (CASBS, 2022). We focus exclusively on boardings on the King County Metro (as opposed to boardings possible with the same transit card on other transit systems such as the light rail or water taxi in the area). In order to measure long-term use, we consider only the boardings made 6 months after registration. This data can be requested from the King County Metro/CISIL Data Challenge organizers.
3. *GTFS Feed.* This data set records public transit usage and consists of several data sets containing routes, fare information, and number of riders on every trip for all bus stops. Some of the data sets we find relevant are:
    - Trips: This data set captures every trip made by any form of public transit, including trip direction and route.
    - Stops: This data set records all stops in King County identified by location and unique ID.
    - Calendar: This data set records the dates of operation and service patterns that operate recurrently for every route serving in the King County region.

This data is publicly available from Sound Transit (2022).

Unequal Commute Data: We used census tract codes Federal Information Processing Standard/Geographic Identifier (FIPS/GEOID) to link the following data to estimate treatment effects regarding access to public transit:

- *Urban Institute's Unequal Commute Data set.* The Unequal Commute data gives data on job access via transit. We look at the differential effects of access to jobs and public transportation using open source, supplemental data from the Urban Institute (Jenkins et al., 2020). This helps to understand the difficulty for low-income riders to commute to their jobs.

2.2. Defining the Research Question and Study Design

After exploring several data sets, we identified two questions that could be addressed using our data—evaluating the long-term impact of the SAP incentive on ridership and reenrollment into the program. We decided on these questions because they were vital for King County Metro to evaluate the SAP program (CASBS, 2022).

We finally converged on a combination of Unequal Commute Data from the Urban Institute (Jenkins et al., 2020), American Census Survey data from the U.S. Census Bureau (n.d.), LIFT registry and boarding data containing information about SAP enrollment and ridership provided by King County Metro as part of the data challenges, and Automatic Passenger Counting (APC) data on public transit schedules identified by the Center for Advanced Study in the Behavioral Sciences (2022) (see Figure 1). Combining these data sets allowed us to assess differential impacts of the SAP program on different subgroups of the population in King County and make specific observations on unequal commute.

Using these data, we constructed two data sets for our study: one for reenrollment and another one for ridership. The reenrollment data set contained 26,073 entries, and ridership had 13,427 entries. The reenrollment data set contains more entries because it includes all individuals eligible for reenrollment, while ridership contains only individuals for whom ridership was observed. We used these data to develop variables describing overall job access, job access via transit, spatial mismatch, number of low-wage workers, job access via transit at third-shift hours (work shifts between 11 PM and 7 AM), language, race, age, and agency for reenrollment. The same attributes were used in addition to the stage of card for ridership. As a preprocessing step for matching, we converted our continuous covariates into categorical by binning their values into quartiles and assigning each quartile a score from 1 to 4. All of our covariates are described in Table E1.

2.3. Modeling Assumptions

In our study, we aim to estimate the causal effect of the SAP program on ridership. Since the data on SAP enrollment is observational rather than experimental, we cannot assume that the treatment (SAP enrollment) is assigned randomly. Instead, individuals may be selected for SAP in a way that correlates with their inherent likelihood of ridership. This is problematic because those who are more inclined to use public transit may be more likely to enroll in the SAP program, confounding our ability to attribute differences in ridership to the SAP program alone.

To overcome this issue, we employ a matching method for causal inference on observational data. Matching allows us to compare treated and untreated units (individuals with and without the SAP) that are similar across covariates—observable characteristics that might affect ridership. By ensuring that the characteristics of the treated and control groups are similar, we create a pseudo-experimental setup that allows us to estimate the treatment effect more reliably. More specifically, we are interested in estimating the **average treatment effect on the treated (ATT)** via **conditional average treatment effects (CATEs)**, which quantifies the expected effect of the SAP program on ridership for individuals with a given set of characteristics.

To perform causal inference and claim valid estimates of the SAP program's effect, we make several assumptions that are standard in causal inference:

- **Stable Unit Treatment Value Assumption (SUTVA)**: SUTVA has two components:

  - *No interference*: The treatment applied to one unit does not affect the outcome of another unit. In our context, one individual's SAP enrollment status does not impact the ridership outcome of others.
  - *Consistency*: Each individual has a well-defined treatment status and there is only one version of each treatment level. This means there is a consistent way in which the SAP program affects each enrolled individual.
- **Positivity (Overlap)**: This assumption requires that for each combination of covariates X, there is a positive probability of both treatment and control assignments.

$$0 < P(Z = 1 \mid X) < 1$$

  This ensures that within each subgroup defined by X, there are both treated and untreated units. If certain values of X correspond exclusively to either treated or untreated units, it would be impossible to make valid comparisons. To ensure this, we only included individuals who would have been eligible for the SAP based on their income levels before performing the matching.
- **Unconfoundedness (Ignorability)**: This assumption requires that, conditional on a set of observed covariates X, the treatment assignment Z (where $Z = 1$ for treated units and $Z = 0$ for untreated units) is independent of the potential outcomes, $Y(1)$ and $Y(0)$.

$$\{Y(1), Y(0)\} \perp Z \mid X$$

  where $Y(1)$ denotes the potential ridership outcome if the individual receives the SAP, and $Y(0)$ denotes the outcome if they do not. This means that given X, there are no unmeasured variables influencing both the treatment assignment and the outcome, which would otherwise bias our estimates. In practical terms, this implies we assume that all relevant factors affecting both SAP enrollment and ridership are captured by the demographic and socioeconomic data we have in our X. A sensitivity analysis was performed to assess the robustness of our results to potential unobserved confounding (see Appendix B).

Within each matched group, defined by similar covariates X^{REL}, we estimate the **Conditional Average Treatment Effect** (CATE), which is the difference between the average outcome of treated and control units within that group. Formally, the CATE for units with covariate values X^{REL} is given by:

$$\mathbb{E}[Y(1) - Y(0) \mid X = X^{\text{REL}}]$$

Here:

- $\mathbb{E}[Y(1) \mid X = X^{\text{REL}}]$: The expected outcome (e.g., ridership) if treated, conditional on covariates X^{REL}.
- $\mathbb{E}[Y(0) \mid X = X^{\text{REL}}]$: The expected outcome if untreated, conditional on the same covariates.

The difference between these expected values for treated and untreated units within each matched group gives an estimate of the causal effect of the SAP on ridership for individuals with characteristics X^{REL}. We construct X^{REL} as the empirical common support—the set of covariate values where both treated and untreated units are

present. This means X^{REL} is determined by the characteristics of the treatment group and the regions of overlap in the control group, which ensures compliance with the positivity assumption. Our target estimand is the ATT, and we compute a weighted average of the CATEs across matched groups containing treated units. This provides a treatment effect estimate that is representative of the population who actually received the SAP.

2.4. Performing Exact Matching and Evaluating Its Performance

We used a method known as Fast Large-scale Almost Matching Exactly (FLAME) to match individuals in the SAP program (treatment) with individuals in all other incentive groups (control) based on demographics such as age, race, and job access via transit (T. Wang et al., 2021). By matching units with similar covariates, this exact matching procedure attempts to achieve covariate balance between the treated and control groups by using as many relevant covariates as possible. The FLAME method determines the covariates that are most relevant to match on based on feature importance with a machine learning algorithm on a held-out training set. This method thus ensures that units are matched on the confounders most relevant to predicting our outcomes of interest: re-enrollment or ridership.

Not only is FLAME comparable in accuracy to black box algorithms, it is highly interpretable in that we are able to see the covariates each individual was matched on in Figure 2. After the matched groups are constructed using FLAME, CATEs can be calculated from each matched group by comparing the treatment and control outcomes of units in the matched group. As noted above, we estimate the ATT by averaging the CATEs across matched groups that contain treated units. The CATEs for individual matched groups can then be combined to estimate the average treatment effects for any desired subgroup. A key advantage of our approach is its interpretability; this analysis has the advantage of being interpretable for policymakers who would like to know which demographic groups within the population benefited most from the SAP program in order to understand how future resources should be allocated.

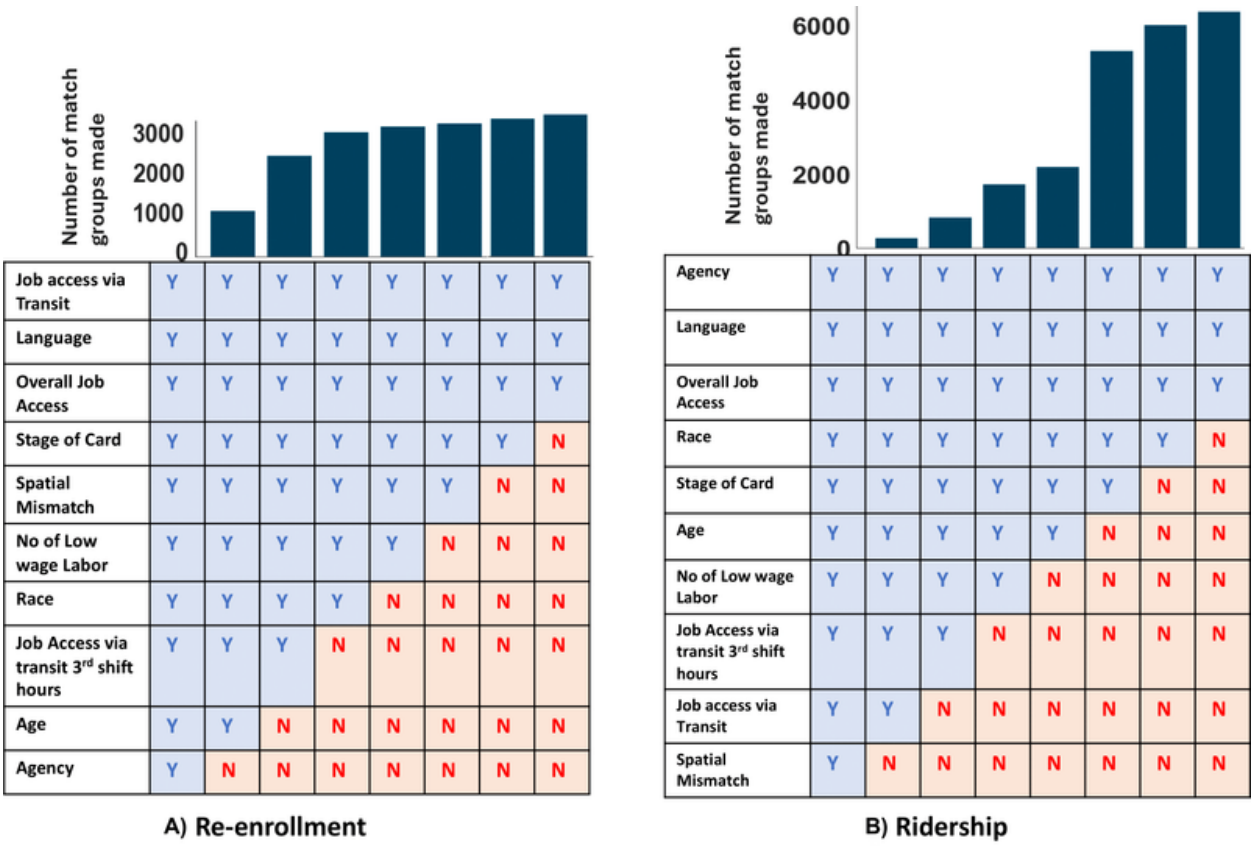


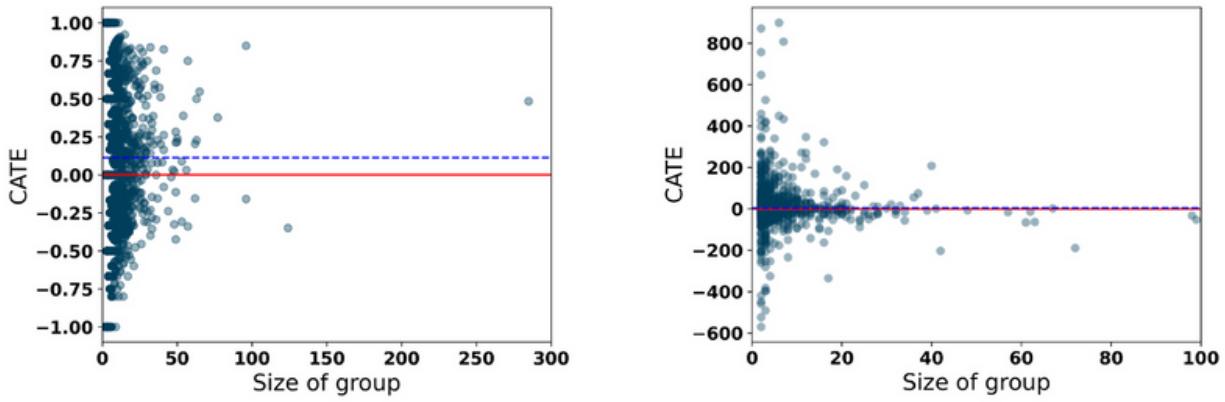
Figure 2. The covariates that matched groups were formed on against the number of matches made on those covariates for (a) reenrollment and for (b) ridership. 'Y' (yes) and 'N' (no) indicate whether or not matches were formed using the row's covariate at the column's iteration number. The columns represent iterations of Fast Large-scale Almost Matching Exactly (FLAME), starting from the left. Histograms above the tables show that a large number of units were matched on a large number of covariates, with the least important covariates dropped first (for example, agency for reenrollment in [a]) and more important covariates dropped later (for example, spacial mismatch for reenrollment in [a]).

In order to determine if an individual reenrolled, we checked to see if a particular user appears in the data set within a 6-month period after October 1, 2020. Since the SAP had been distributed starting October 1, 2019, this is the earliest that the SAP could have expired and reenrollment could have occurred. We allow a 6-month period for reenrollment to allow for some delays based on receiving the original SAP pass (i.e., receiving the SAP a few weeks later) and lags in discovering expiry and seeking enrollment. While some users reenrolled multiple times, we still consider this as reenrolling only once, as the duration of our study does not allow for multiple reenrollments due to each enrollment being valid for at least a year. In order to measure long-term ridership, we aggregated the total number of times individuals boarded (focusing exclusively on King County Metro and excluding any other transit programs in the area), starting 6 months after they registered for a SAP. We estimated the causal effect of SAP as follows: for ridership, given the treatment, we computed ATT and CATE estimates as the numerical increase in the number of rides. To support King County Metro's reenrollment goals, we estimated the CATE to measure how much more likely reenrollment is after the

treatment. We constructed 95% confidence intervals to quantify the uncertainty of these estimates. These intervals were generated using a hold-out procedure where we refit our model on bootstrap samples of the training set to learn a distance metric, which was then used to estimate treatment effects on the hold-out set through matching using the learned distance metric.

It is important to note that these ATT estimates are for people who were eligible for the SAP rather than the entire population of King County. Essentially, our ATT estimates capture the difference in outcomes for low-income communities given that they availed the SAP incentive or had to pay for public transit.

3. Our Findings and Analysis



(A) Reenrollment for all match groups (B) Ridership for all match groups

Figure 3. Conditional average treatment effects (CATEs) (each matched group is represented by a dot) and average treatment effect on the treated (ATT) (in lines) of subsidized annual pass (SAP) for reenrollment in (a) and ridership in (b). Using all matched groups, the estimated estimated ATT of SAP on reenrollment is 11.2% [10.7%, 11.8%], while the ATT of SAP is −1.613 [−2.35, −0.877] rides for ridership. SAP appears to increase both reenrollment and ridership, although given the variance of CATEs, this is not so obvious or reliable for ridership.

As matched groups on the left of the plots are smaller and less trustworthy, if we pick group size 20 as a threshold for a match group to be considered reliable and mask smaller groups and adjust our estimated ATT of SAP on reenrollment using the reliable matched groups we get an ATT of +0.164 [0.156, 0.173], which means users enroll on average 16.4% [15.6%, 17.3%] more often when given a SAP as compared to when they are given other incentives. Similarly, the estimated ATT of SAP on ridership using reliable matched groups is +3.968 [2.801, 5.135] rides during our time period. We try different thresholds of reliability and find that larger match groups strengthen our findings. To understand the scale, the median ridership for treated individuals is 36 and 21 for control individuals when excluding all riders with 0 rides.

3.1. Reenrollment

Overall, there was a *slight positive effect of SAP on ORCA-LIFT reenrollment*, as shown in Figure 3a.

The average treatment effect (ATE) is +0.112 [0.107, 0.118], which can be interpreted as: the SAP incentive causes on average 11.2% more individuals to reenroll as compared to those who were given other ORCA-LIFT incentives. The increased reenrollment rate for treated individuals shows that there is sustained demand for access to public transit. We also observe a heterogeneity of reenrollment rates across matched groups and on conditioning for certain covariates. With Hispanic individuals, we observe an estimated CATE of +0.039 [0.015, 0.063], while for Asians this is +0.171 [0.156, 0.186], showing that treatment has affected the outcomes of some populations more than others. Recognizing this heterogeneity is valuable because it can help determine how the policy might affect different sections of the population.

3.2. Ridership

Overall, there was a *slight positive effect of SAP on ridership*, as shown in Figure 3b.

A total of 119,101 rides were recorded for 3,350 riders who received the SAP card for an average ridership of about 35.55 rides per person enrolled in the program, while the control group (non-SAP recipients) was composed of 10,077 people who took 30.78 rides on average. Comparing median rides for riders who used transit at least once shows that treated units had a median of 36 rides, while control units had a median of 21 rides. After matching between treated and control units, we found that the SAP incentive modestly increased ridership by an additional +3.968 [2.801, 5.135] more trips on average per year on King County Metro compared to individuals that were given other incentives. We note that the number of overall rides is low as "physical commute trips (had) plummeted amid COVID-19 as most respondents (were) telecommut(ing) instead" (King County Metro, 2020).

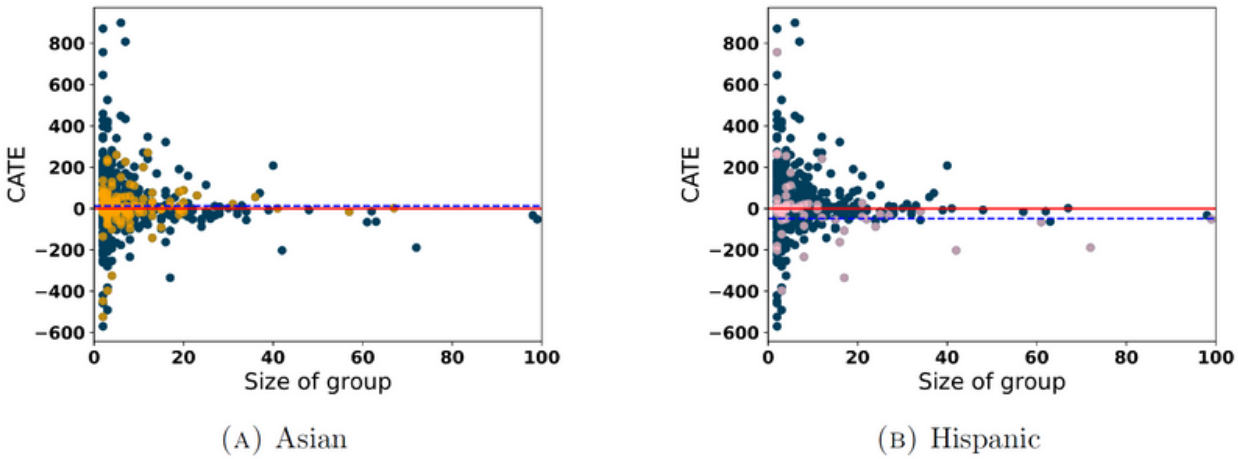


Figure 4. Treatment effect of subsidized annual pass (SAP) on ridership on Asian (a) and Hispanic (b) population. (a) Matched groups with individuals of Asian descent (dotted in yellow) showed a more positive estimated conditional average treatment effect of SAP on ridership compared to other races, +11.657 [9.004, 14.31]. (b) Matched groups with Hispanic individuals (dotted in pink) show a negative conditional estimated average treatment effect of SAP on ridership, −29.124 [−33.139, −25.11]. While the average treatment effect on the treated (ATT) values are smaller, it is worth considering that there could be differential effects for subpopulations.

Again we note that the effects across groups vary substantially, with Black populations exhibiting a CATE of +20.255 [14.431, 26.08] rides, while Hispanic populations have a negative −48.27 [−54.478, −42.063] rides as shown in Figures 4a and 4b. In Appendix C, we show that the moderate increase in ridership, if proven to be sustained through the SAP incentive mechanism, could have a positive effect on the overall economic landscape of King County when accounting for gains in productivity of workers due to improved commutes (Jenkins et al., 2020).

We plot the distribution of the number of rides for treatment and control units in Appendix D, showing that the treatment units overall tended to take slightly more rides.

Finally, although the average treatment effect estimates were positive, the sizes of these effects were quite small. We recommend that more data be collected in order to reach a more conclusive understanding of the effectiveness of the SAP program.

3.3. Significant Attributes for Predicting Ridership and Reenrollment

Since we used FLAME, an interpretable matching method, we can observe in which order covariates were dropped for every iteration of the algorithm. Figure 2 shows this order for reenrollment (Figure 2a) and ridership (Figure 2b). The most important characteristics in predicting reenrollment include *job access via transit, language, and overall job access*. Here, spatial mismatch is the number of job opportunities within a 30-minute commute from a job seeker divided by the sum of the competition for those jobs (Jenkins et al., 2020).

For ridership, the most important characteristics include *agency, language, and overall job access*. We provide an explanation of all attribute meanings in Table E.1 in Appendix E.

3.4. Influence of Covariates on Reenrollment and Ridership

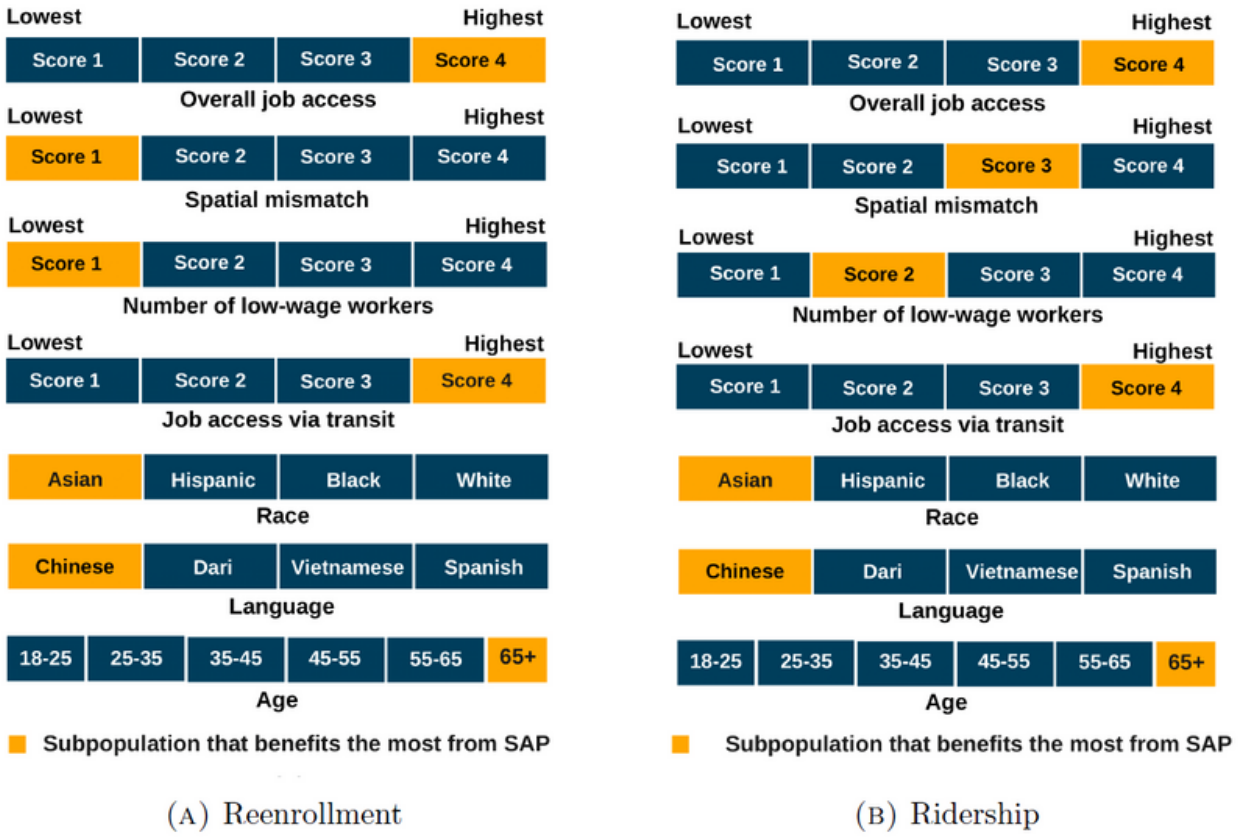
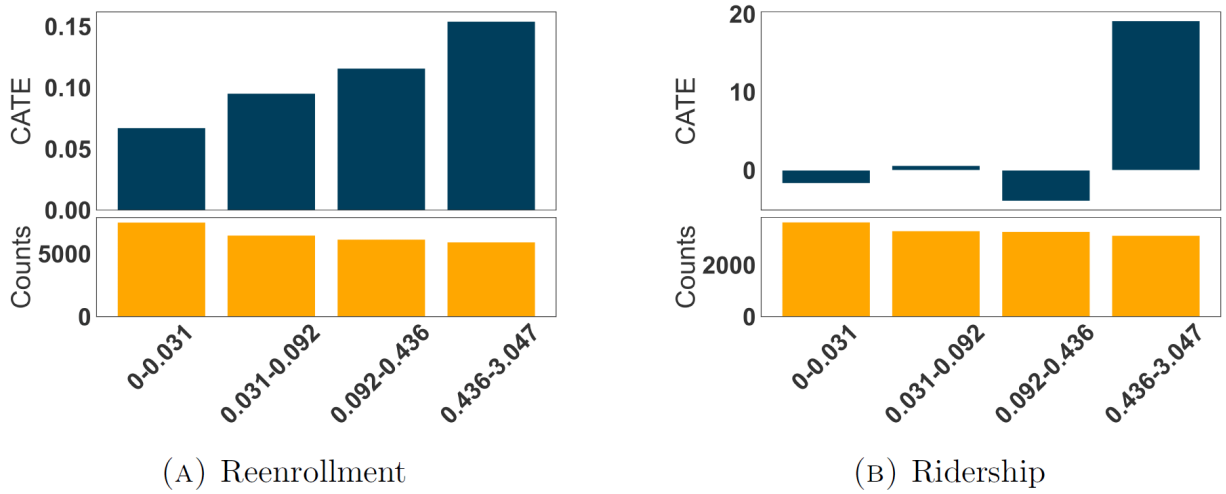


Figure 5. Characteristics of the groups who saw the greatest increase in (a) reenrollment and (b) ridership from receiving the subsidized annual pass (SAP) incentive. (These results are summarized from Figures G1–G2 in Appendix G.) Recall that scores are determined as follows: Bins are created by breaking the entire population into quartiles for the given continuous covariate and assigning each quartile a number from 1 to 4.

For subgroups among the 10 covariates, we computed the CATE for reenrollment (see Figure G1) and ridership (see Figure G2). We summarize these figures in Figure 5 for reenrollment (a) and ridership (b). We found that *older people, people who identify as Asian (and particularly speakers of Mandarin), people who are already long-term users of the metro, and people who live in areas with highest job access tend to be the individuals who benefited most from the free passes*. We also observe that for long-term ridership, *agency* appears to be important. This suggests that King County Metro's treatment assignment may not have been random since the agency where the card was received should not affect ridership more for the SAP incentive than other incentives, controlling for other variables. Nonetheless, since cards were handed out without any detailed guidelines (CASBS, 2022), the procedure for distributing cards might have varied across agencies and more careful treatment assignment should be carried out if an experiment were to be conducted to validate these

findings in the future. Moreover, it could be possible that all cards at some agencies might have been handed out to regular riders (those most likely and most frequent to use transit) as opposed to infrequent riders, so that after matching, our results indicate that agency plays an important role. Interestingly, we observed that the Hispanic population showed a large negative CATE for ridership (Figure G2c) while all of the other race groups had positive CATEs. Figure 4 further highlights this heterogeneity between races.

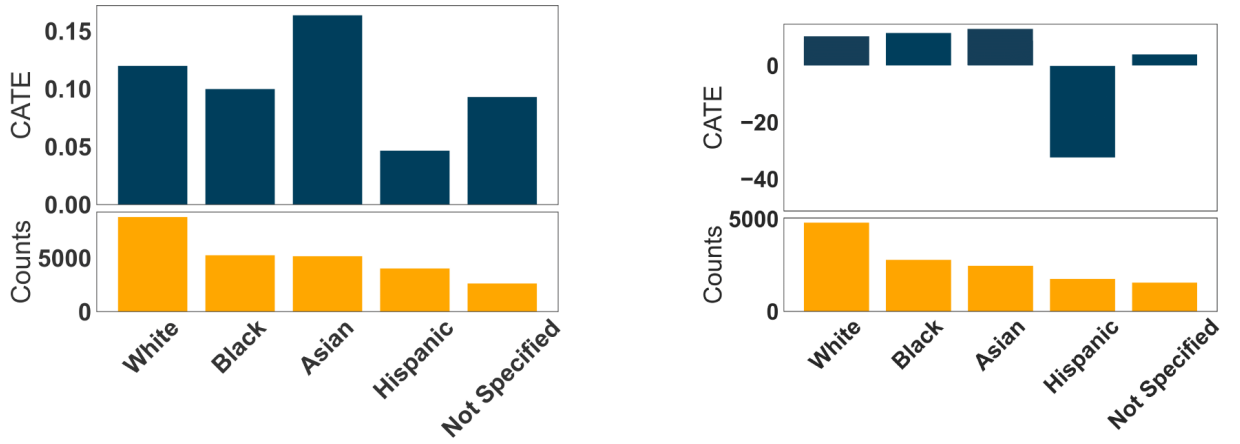
In terms of age, we observed higher positive treatment effects of SAP on reenrollment and ridership for older individuals. More specifically, the treatment effects of those in the oldest age group of 65+ are moderately higher than those in the 55–65 age group for ridership (Figure G2c) and reenrollment (Figure G1c).



(A) Reenrollment (B) Ridership

Figure 6. Conditional average treatment effects (CATEs) for job access via transit on reenrollment (a) and ridership (b). For reenrollment, the CATE values and 95% confidence intervals are: Quartile 1: 0.067 [0.055, 0.08], Quartile 2: 0.095 [0.082, 0.108], Quartile 3: 0.116 [0.103, 0.128], Quartile 4: 0.154 [0.145, 0.163]. For ridership, the CATE values and 95% confidence intervals are: Quartile 1: −1.673 [−3.483, 0.138], Quartile 2: 0.591 [−1.49, 2.672], Quartile 3: −3.936 [−6.134, −1.737], Quartile 4: 19.119 [16.266, 21.973]. Individuals with higher job access via transit are more likely to ride King County Metro as a result of enrolling in SAP than individuals with lower job access via transit. (Recall that job access via transit was binned into quartiles and scores between 1 to 4 are assigned to these quartiles. Meaning that, for Quartile 1 with values in range [0,0.031], we assign score 1, for Quartile 2 [0.031, 0.092], score 2, for Quartile 3 [0.092, 0.436], score 3, and for Quartile 4 [0.436, 3.047], score 4). Note that since not all units are matched, the final distribution of scores is not exactly uniform.

Individuals belonging to census tracts with the highest job access via transit and overall job access showed a greater effect of SAP on reenrollment and ridership based on CATEs (Figure 6 and Figure H1 in the appendix).



(A) Reenrollment (B) Ridership

Figure 7. Conditional average treatment effects (CATEs) for race on reenrollment (a) and ridership (b). For reenrollment, the CATE values and 95% confidence intervals are White: 0.127 [0.118, 0.137], Black: 0.098 [0.085, 0.111], Asian: 0.171 [0.156, 0.186], Hispanic: 0.039 [0.015, 0.063], Not Specified: 0.081 [0.068, 0.093]. Asian individuals are more likely to ride King County Metro as a result of enrolling in subsidized annual pass (SAP) than individuals of other races. For ridership, the CATE values and 95% confidence intervals are White: 8.451 [6.684, 10.217], Black: 8.921 [5.743, 12.099], Asian: 11.657 [9.004, 14.31], Hispanic: −29.124 [−33.139, −25.11], Not Specified: 4.059 [1.54, 6.578].

Individuals who are Asian showed a greater effect of SAP on reenrollment and ridership, meaning that there is a higher positive CATE for Asian individuals (Figure 7). The matched groups with individuals of Asian descent showed higher CATE estimates, compared to matched groups of White individuals. We also note that the size of matched groups with Asian individuals tended to be large (Figure 4a), giving us reliability in the CATE estimates due to the larger size of the groups. Additionally, Chinese-speaking individuals have shown a greater effect of SAP on reenrollment and ridership (see Figures G1e and G2e,). Among individuals of different races, Hispanic individuals show a negative conditional estimated average treatment effect of SAP on ridership as we show in Figure 4b. Interestingly, for reenrollment, the estimated average treatment effect of SAP for Hispanic riders was positive (Figure 7a).

4. Related Work

There are several topics and methods related to our work.

4.1. Supplementing Smart Card Data

Smart card data is known to support detailed research of user behavior in transit systems (Agard et al., 2006; Bagchi & White, 2005), and transit researchers have a growing interest in using these data sets to understand user travel patterns (Koutsopoulos et al., 2019). Many researchers have leveraged smart card data to research user behavior in relation to differential pricing, subway crowding, and the vulnerability of different transit

systems (Anupriya et al., 2020; Halvorsen et al., 2016; Zhang et al., 2021). However, none of these previous works leverage smart card data to analyze social and economic equity issues for fare pricing subsidies. Moreover, these previous works often use smart card data with only one or two adjacent data sets or in isolation. Instead, we use a suite of seven different data sets connected by census tract identifiers to understand how SAP affects different races and those with varying degrees of access to public transit.

Our project makes a novel effort to leverage population data from (1) U.S. Census's American Community Survey and (2) Urban Institute's Unequal Commute Data (Jenkins et al., 2020; U.S. Census Bureau, n.d.). In doing so, we are able to better analyze the effects of different incentives on different subgroups of the population, such as racial groups and groups with different job access via transit from their place of residence.

4.2. Causality Research in Transportation

Transportation researchers generally use traditional methods such as linear regression and propensity score matching to perform causal analysis (Schiff et al., 2017). These approaches have problems. First, linear regression is almost always misspecified: a linear model may not perform well when the reality is nonlinear. Matching approaches are nonparametric, which mitigates this problem (T. Wang et al., 2021). Second, propensity score matches do not allow us to estimate *conditional* average treatment effects, nor are their matches interpretable. Propensity score matching projects the data to one dimension (namely, propensity, which is the probability of receiving the treatment), so matches are made on individuals who may be completely different from each other in feature space. FLAME, unlike previous causal estimation techniques, does not run into the problem of misspecification and can faithfully provide CATE estimates.

4.3. Machine Learning in Transportation

Many transportation researchers are beginning to leverage traditional machine learning techniques such as support vector machines, decision trees, and deep neural networks to conduct data analysis of transportation patterns (Bhavsar et al., 2017; S. Wang et al., 2021). Recent work such as that of Zheng et al. (2021) even uses deep neural networks to analyze transit accessibility, an issue that is central to our article. Unfortunately, these techniques are black box machine learning models, making it difficult for users to understand why those models make the decisions they do. Attempts to interpret the gradients or saliency maps of black box models may not result in faithful interpretations (Rudin, 2019). Further, it is likely that neural networks will not provide any benefit over interpretable models since the outcomes are not a deterministic function of the inputs (see Semenova, Chen, Parr, & Rudin, 2023; Semenova, Rudin, & Parr, 2022, on the relationship between noise in the outcomes and the existence of simple-yet-accurate models). On the other hand, FLAME is inherently interpretable, and is designed for observational causal inference. Instead of matching on propensity score, FLAME makes almost-exact matches using as many important covariates as possible. This allows us to create interpretable high-quality matches between units.

4.4. Equitable Access in Public Transit

Social scientists, policy researchers, and sociologists have long analyzed how racial disparities exist for American workers to commute to work (Gautier & Zenou, 2010; McKenzie, 2015). Our research builds on this literature to understand how a subsidy plan can have differential effects on the aforementioned historically marginalized groups.

5. Discussion

We find that low-income incentive programs such as subsidized annual passes are effective for incentivizing ridership in low-income populations. Additionally, we find that individuals from census tracts with the highest access to jobs benefit the most from the SAP incentive.

5.1. Limitations in the Data and Future Steps to Improve Data Collection

We list below the limitations of the King County Metro data.

First, the SAP program launched in late 2020, giving us only 2 years' worth of data to analyze long-term ridership and reenrollment, which limited the scope of our analysis. We encourage future analyses to look into the motivation behind reenrollment when longer-term data are available.

Second, we did not have the necessary data to ensure that our treatment and control groups had the same income. Specifically, only individuals at or *below* 80% of the poverty level are eligible to enroll in the SAP (treatment), whereas other individuals in the ORCA-LIFT Card Registry data set may be *up to* 200% of the federal poverty level (control). Hence, although we know that both our control and treatment groups are comprised of only low-income individuals, individuals in the control group could have been individuals who did not qualify for the SAP program. As we lacked data on income, which determined treatment eligibility, we had to perform our analysis as if all control units were eligible. Future data sets would benefit from including income data about individuals.

5.2. Broader Impact

Since our analysis is reproducible and each matched group can easily be checked for trustworthiness, it allows us to suggest policy recommendations whose efficacy may be determined due to our interpretable modeling choice. Cost permitting, we recommend that reduced- and free-fare programs be implemented and expanded in neighborhoods and areas with high job access via transit and job hubs, such as inner cities. Our work communicates to King County Metro stakeholders that the SAP program has been effective and builds a case for expansion of the SAP program. We also recommend including a control group when implementing programs to allow efficacy to be measured more easily.

As King County Metro is one of the leaders in public transit policy, the work leading out of this analysis may be closely followed by other jurisdictions and presented in national transportation policymaking forums

(CASBS, 2022). If these free-fare programs are successful, it could increase access to:

- Jobs: Increased access to jobs leads to upward economic mobility. Furthermore, during the COVID-19 pandemic, public transportation was critical for essential workers to commute to jobs in areas such as health care (Davis & Stacy, 2021; Haltiwanger et al., 2018).
- Libraries: Previous research has shown that people are more likely to visit public libraries if they have an easier time traveling to them. Libraries are important for improving community well-being and alleviating social inequity by providing community members the opportunity to access information, services, and educational programs (Allen, 2019).
- Schools: Transportation is often a barrier to high-quality education for low-income families. Many low-income families say that they would send their children to better schools in choice-heavy cities if they had accessible transportation to schools (Chingos & Blagg, 2017).
- Health care: Studies have shown that 25% of missed appointments for low-income individuals was a result of lack of public transportation (Silver et al., 2012). Health care research indicates that this leads to poor health care outcomes, particularly for low-income families (Syed et al., 2013).

Increased job access via transit can have broader implications for equity:

- Air pollution: Increased access to more equitable transit can reduce air pollution, especially in congested metropolitan areas, having a positive effect on the health of all community members (Skillen & Abramsohn, 2021), specifically improving living conditions for communities of color and low-income communities.
- Safety: Fatal motor accidents are the leading cause of commute-related deaths in the United States according to the U.S. Bureau of Transportation Statistics (2022). Moreover, road accidents disproportionately affect lower income communities (Lazo et al., 2022). Increased job access via public transit can make roads safer, leading to fewer accidents and motor vehicle–related deaths (Skillen & Abramsohn, 2021).

Follow-up observational studies (particularly long term) would be useful to monitor the effectiveness of the free-fare programs. While King County was not able to release data on their randomized control trials for previous cash subsidy programs, we believe that some long-term RCT studies on the effectiveness of the SAP program specifically would support the evidence-based policy expansion of SAP. Code to reproduce and validate our results is available at the following GitHub link.

5.3. Designing Future Studies

A limitation of our study design is the lack of true randomization in treatment assignment. Randomized controlled trials (RCTs) are widely regarded as the ‘gold standard’ for evaluating causal effects (Schiff et al., 2017), as they eliminate confounding by randomly assigning treatment. In our context, an ideal RCT would involve randomly selecting transit card numbers for SAP eligibility, allowing us to measure the intent-to-treat (ITT) effect of the subsidy. The ITT estimates the effect of being offered the SAP incentive on ridership and reenrollment, regardless of whether recipients fully complied with the program.

However, due to the lack of randomization in our setting, we rely on observational data and use covariate-matching methods to estimate the ATT. This estimand reflects the causal effect of SAP participation on those who actually received the subsidy. Our matching strategy adjusts for observable differences between treated and untreated individuals to approximate a pseudo-randomized comparison. Nevertheless, this approach assumes no unmeasured confounding and is sensitive to treatment compliance behaviors.

In contrast, the ITT effect from an RCT would remain unbiased even under noncompliance. Noncompliance could arise in our setting from SAP misuse—such as card sharing—or if recipients lost access and did not reenroll. Unfortunately, we do not observe data that would allow us to identify or model compliance directly. Still, the ATT and CATEs estimated in our analysis provide valuable insights into the realized impact of SAP among actual participants, which is particularly relevant for understanding how the program performs in real-world conditions.

Acknowledgments

Thank you to the Causal Inference for Social Impact Lab's Data Challenge, in collaboration with the Center for Advanced Study in the Behavioral Sciences at Stanford University and King County Metro. The CISIL Data Challenge is supported by the Alfred P. Sloan Foundation, the Knight Foundation, and SAGE Publishing. We also thank the organizers (in particular, Wendy Martinez) of the American Statistical Association Data Challenge Expo 2022.

This study relies in part on data that were produced by an experiment conducted by Brough et al. (2022) in collaboration with King County Metro.

Disclosure Statement

Authors have no financial or nonfinancial disclosures to share for this article.

References

Agard, B., Morency, C., & Trépanier, M. (2006). Mining public transport user behaviour from smart card data. *IFAC Proceedings Volumes*, *39*(3), 399–404. https://doi.org/10.3182/20060517-3-FR-2903.00211

Allen, J. (2019). Mapping differences in access to public libraries by travel mode and time of day. *Library & Information Science Research*, *41*(1), 11–18. https://doi.org/10.1016/j.lisr.2019.02.001

Alonso González, M., Jonkeren, O., & Wortelboer-van Donselaar, P. (2022, August). *Equitable Transport Policy* [Research note]. Kennisinstituut voor Mobiliteitsbeleid. https://english.kimnet.nl/site/binaries/site-content/collections/documents/2022/08/08/equitable-transport-policy/KiM_Research+note+equitable+transport+policy_def.pdf

Anupriya, Graham, D. J., Hörcher, D., Anderson, R. J., & Bansal, P. (2020). Quantifying the ex-post causal impact of differential pricing on commuter trip scheduling in Hong Kong. *Transportation Research Part A: Policy and Practice*, *141*, 16–34. https://doi.org/10.1016/j.tra.2020.08.011

Bagchi, M., & White, P. R. (2005). The potential of public transport smart card data. *Transport Policy*, *12*(5), 464–474. https://doi.org/10.1016/j.tranpol.2005.06.008

Bhavsar, P., Safro, I., Bouaynaya, N., Polikar, R., & Dera, D. (2017). Machine learning in transportation data analytics. In M. Chowdhury, A. Apon, & K. Dey (Eds.), *Data analytics for intelligent transportation systems* (pp. 283–307). Elsevier. https://doi.org/10.1016/B978-0-12-809715-1.00012-2

Brough, R., Freedman, M., & Phillips, D. C. (2022). Experimental evidence on the effects of means-tested public transportation subsidies on travel behavior. *Regional Science and Urban Economics*, *96*, Article 103803. https://doi.org/10.1016/j.regsciurbeco.2022.103803

Center for Advanced Study in the Behavioral Sciences (CASBS). (2022). *Causal inference for Social Impact Lab's data challenge*. Stanford University. https://casbs.stanford.edu/causal-inference-social-impact-lab-s-data-challenge

Chapin, F. S. (1949). *Experimental Designs in Sociological Research*. Harper.

Chingos, M. M., & Blagg, K. (2017). *Student transportation and educational access*. Kennisinstituut voor Mobiliteitsbeleid. https://www.urban.org/research/publication/student-transportation-and-educational-access

Cinelli, C., & Hazlett, C. (2020). Making sense of sensitivity: Extending omitted variable bias. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *82*(1), 39–67. https://doi.org/10.1111/rssb.12348

Constantine, D. (2020, October 15). *King County Metro and Sound Transit offer new subsidized annual pass for riders in greatest need*. King County. https://kingcounty.gov/elected/executive/constantine/news/release/2020/October/15-Metro-fares.aspx

Davis, C., & Stacy, C. P. (2021, January 4). Four lessons for cities to help advance equitable transportation during the COVID-19 pandemic and beyond. *Urban Wire*. https://www.urban.org/urban-wire/four-lessons-cities-help-advance-equitable-transportation-during-covid-19-pandemic-and-beyond

Fisher, R. A. (1992). Statistical methods for research workers. In S. Kotz & N. L. Johnson (Eds.), *Breakthroughs in statistics: Methodology and distribution* (pp. 66–70). Springer. https://doi.org/10.1007/978-1-4612-4380-9_6

Gautier, P. A., & Zenou, Y. (2010). Car ownership and the labor market of ethnic minorities. *Journal of Urban Economics*, *67*(3), 392–403. https://doi.org/10.1016/j.jue.2009.11.005

Haltiwanger, J., Hyatt, H. R., & Spletzer, J. R. (2018). Job displacement and the duration of joblessness: The role of spatial mismatch. *The Review of Economics and Statistics*, *100*(2), 203–218. https://doi.org/10.1162/REST_a_00707

Halvorsen, A., Koutsopoulos, H. N., Lau, S., Au, T., & Zhao, J. (2016). Reducing subway crowding: Analysis of an off-peak discount experiment in Hong Kong. *Transportation Research Record*, *2544*(1), 38–46. https://doi.org/10.3141/2544-05

Jenkins, W., Stacy, C., Stern, A., Blagg, K., Su, Y., Noble, E., Rainer, M., & Ezike, R. (2020, October 6). *The Unequal Commute: Examining inequities in four metro areas' transportation systems*. Urban institute. https://www.urban.org/features/unequal-commute

King County GIS Center Spatial Data Catalog. (2019). *King County ACS Select Indicators*. https://www5.kingcounty.gov/sdc?Layer=demographic_base_area

King County Metro. (2020, August). *COVID-19 Impacts Survey Results*. https://cdn.kingcounty.gov/-/media/king-county/depts/metro/documents/about/data-and-reports/2020/metro-covid-recovery-report.pdf?rev=929b37f5245d4b449ac2f76c5fcbd0d2&hash=6EF2269468A1DC8FF25D345B5E806648

Koutsopoulos, H. N., Ma, Z., Noursalehi, P., & Zhu, Y. (2019). Transit data analytics for planning, monitoring, control, and information. In C. Antoniou, L. Dimitriou, & F. Pereira (Eds.), *Mobility patterns, big data and transport analytics* (pp. 229–261). Elsevier. https://doi.org/10.1016/B978-0-12-812970-8.00010-5

Lazo, L., Jayaraman, S., & Moriarty, D. (2022, February 23). D.C. traffic deaths at 14-year high with low-income areas hardest hit. *The Washington Post*. https://www.washingtonpost.com/transportation/2022/02/23/dc-traffic-deaths-highest-record/

McKenzie, B. (2015, December 16). *In what ways are D.C.-area neighborhoods with rail transit different from other neighborhoods?* United States Census Bureau. https://www.census.gov/newsroom/blogs/research-matters/2015/12/in-what-ways-are-d-c-area-neighborhoods-with-rail-transit-different-from-other-neighborhoods.html

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, *1*(5), 206–215. https://doi.org/10.1038/s42256-019-0048-x

Schiff, A., Wright, L., & Denne, T. (2017). *Ex-post evaluation of transport interventions using causal inference methods* (research report 630). NZ Transport Agency. https://www.nzta.govt.nz/assets/resources/research/reports/630/630-ex-post-evaluation-of-transport-interventions-using-causal-inference-methods.pdf

Seattle Office of Labor Standards. (2015). *Minimum Wage Ordinance. Chapter 14.19 (Ord. 124960, § 45, 2015.)* https://www.seattle.gov/laborstandards/ordinances/minimum-wage

Semenova, L., Chen, H., Parr, R., & Rudin, C. (2023, December 10–16). *A path to simpler models starts with noise* [Poster presentation]. NeurIPS 2023: Thirty-seventh conference on neural information processing systems, New Orleans, LA, United States. https://openreview.net/forum?id=Uzi22WryyX

Semenova, L., Rudin, C., & Parr, R. (2022). On the existence of simpler machine learning models. In *FAccT '22: Proceedings of the 2022 ACM conference on fairness, accountability, and transparency* (pp. 1827–1858). Association for Computing Machinery. https://doi.org/10.1145/3531146.3533232

Silver, D., Blustein, J., & Weitzman, B. C. (2012). Transportation to clinic: Findings from a pilot clinic-based survey of low-income suburbanites. *Journal of Immigrant and Minority Health*, *14*(2), 350–355. https://doi.org/10.1007/s10903-010-9410-0

Skillen, W. H., & Abramsohn, E. E. (2021, July 29). *Public transportation in the US: A driver of health and equity* [Health policy brief]. Health Affairs. https://www.healthaffairs.org/content/briefs/public-transportation-us-driver-health-and-equity

Sound Transit. (2022). *King County Metro GTFS Feed Data* . https://www.soundtransit.org/GTFS-KCM/google_transit.zip

Syed, S. T., Gerber, B. S., & Sharp, L. K. (2013). Traveling towards disease: Transportation barriers to health care access. *Journal of Community Health*, *38*, 976–993. https://doi.org/10.1007/s10900-013-9681-1

U.S. Bureau of Transportation Statistics. (2022). *Transportation fatalities by mode* [Data set]. U.S. Department of Transportation. https://www.bts.gov/content/transportation-fatalities-mode

U.S. Census Bureau. (n.d.). *American Community Survey 5-year data (2009-2020)*. Retrieved April, 2022, from https://www.census.gov/data/developers/data-sets/acs-5year.html

Wang, S., Wang, Q., Bailey, N., & Zhao, J. (2021). Deep neural networks for choice analysis: A statistical learning theory perspective. *Transportation Research Part B: Methodological*, *148*, 60–81. https://doi.org/10.1016/j.trb.2021.03.011

Wang, T., Morucci, M., Awan, M. U., Liu, Y., Roy, S., Rudin, C., & Volfovsky, A. (2021). FLAME: A fast large-scale almost matching exactly approach to causal inference. *Journal of Machine Learning Research*, *22*(1), Article 31.

Zhang, N., Graham, D. J., Hörcher, D., & Bansal, P. (2021). A causal inference approach to measure the vulnerability of urban metro systems. *Transportation*, *48*(6), 3269–3300. https://doi.org/10.1007/s11116-020-

10152-6

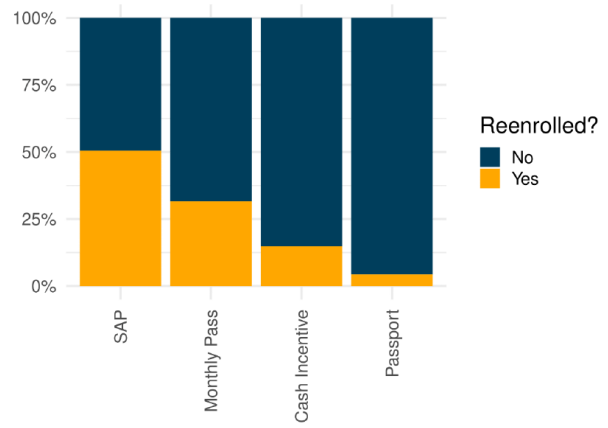
Zheng, Y., Wang, S., & Zhao, J. (2021). Equality of opportunity in travel behavior prediction with deep neural networks and discrete choice models. *Transportation Research Part C: Emerging Technologies*, *132*, Article 132. https://doi.org/10.1016/j.trc.2021.103410

Appendices

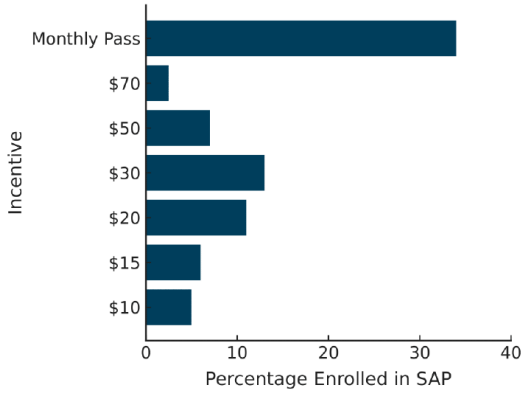
Appendix A. Additional Exploratory Data Analysis

A.1. Among All Incentives, Subsidized Annual Passes (SAP) Showed the Highest Reenrollment

During our initial data exploratory analysis, we found that individuals receiving SAP had the highest percentage of reenrolling in their respective program (see Figure A1a). That is why we were interested in exploring the causal relation between receiving SAP and reenrollment, accounting for confounding factors like race, income, and access to transit to explore how effective SAP had been.



(A) Individuals receiving subsidized annual passes (SAP) have the highest percentage of reenrolling in their respective ORCA-LIFT program.



(B) Impact of different incentives on enrollment in SAP. A high percentage of users who received the monthly pass incentive later enrolled in SAP compared to users who received other incentives.

Figure A1. Exploratory data analysis to explore incentive types and percentage of individuals receiving a given incentive.

A.2. Influence of ORCA-LIFT Incentives on Subsidized Annual Pass (SAP) Enrollment

Prior to the launch of the SAP incentive program, King County policymakers were interested in how ORCA-LIFT incentives impacted enrollment in SAP. Although our main analysis did not focus on this question, through preliminary exploratory analysis, we observe that those who received ‘monthly passes’ as their

incentive showed the highest percentage of SAP enrollments (see Figure A1b). However, we note that we do not have data on individual income, hence are not able to make conclusive statements about whether the individuals analyzed were at or below 80% of the poverty line (which is the SAP enrollment eligibility criteria). That is, we do not know who was eligible for SAP enrollment, so we cannot answer the question of how ORCA-LIFT incentives impacted SAP enrollment.

Appendix B. Sensitivity Analysis

B.1. Binning Strategy

To ensure robustness in our causal inference results, we ran the Fast Large-scale Almost Matching Exactly (FLAME) algorithm on alternative binning strategies, particularly for continuous variables. We experimented with decile binning as an alternative to quartile binning for these continuous variables. This finer grained approach allowed us to observe whether smaller bin intervals significantly altered downstream causal inference results. Our findings indicate that results remained consistent when using decile binning, which suggests that the main insights are stable across different binning strategies and robust to the level of granularity.

We chose to present the binning strategy that maximizes interpretability of results. Age was grouped into bins defined by the intervals: [18, 25, 35, 45, 55, 65+]. This choice reflects common age brackets that are widely understood and allows for clear distinctions across different life stages, which enhances interpretability in our final analysis.

For other continuous variables, we opted to categorize data into quartiles. To validate our binning approach, we generated histograms for each variable and visually inspected the distributions to confirm that the bins were logical and reasonably spaced. Our chosen bin widths maintained interpretability without compromising the granularity needed for meaningful analysis.

B.2. Robustness to Unobserved Confounding

To further assess the resilience of our estimated treatment effects to potential unobserved confounding, we conducted a robustness value check, following the approach proposed in Cinelli and Hazlett (2020). The robustness value quantifies the minimum association that an unobserved confounder would need to have with both the treatment and the response to nullify an observed effect estimate. Specifically, it represents the threshold at which an unobserved confounder could reduce the estimated effect of the treatment on the outcome to zero.

For the sensitivity analysis, we fit an ordinary least squares (OLS) regression using the response variables (1) enrollment count and (2) ridership. The treatment variable was defined as an indicator for receiving a subsidized annual pass (SAP), and all observed covariates from the initial analysis were included as control variables. The robustness value derived in this setup offers insight into the potential impact of omitted variables on the observed treatment effect.

Our analysis yielded a robustness value of 0.241 for the treatment effect of SAP on reenrollment. This result implies that any unobserved confounder would need a partial correlation of at least 0.241 with both the treatment (SAP) and the outcome (enrollment count) to completely explain away the observed effect. For the treatment effect of SAP on ridership, we obtained a robustness value of 0.114. Thus, an unobserved confounder would need a partial correlation of at least 0.114 with both the treatment and the response (total boardings) to nullify the estimated effect of SAP. These robustness values indicate that our findings are moderately resilient to potential omitted variable bias. The relatively high threshold for the enrollment outcome suggests that the treatment effect for SAP on reenrollment is less sensitive to unobserved confounders compared to the effect on ridership. Therefore, while unobserved confounders could influence our results to some extent, they would need to exhibit considerable associations with both the treatment and response to substantially alter the observed effects. This analysis gives us confidence that our treatment effect estimates are relatively robust to the presence of unmeasured confounding factors.

Appendix C. Cost Analysis

While there is a cost associated with providing subsidized annual pass (SAP) enrollment, primarily attributed to the loss in collection of ridership fares due to the incentive, this cost may be outweighed by the considerable gains in productivity resulting from faster commutes. With an estimated economic cost of around $240,000 of the SAP, based on a cost per ride of $2 for the 119,101 rides recorded for SAP riders, the enhanced productivity from faster mobility may surpass this expense since even a gain of 10 minutes per ride for each of the rides would translate to close to an additional 19,850 productive hours annually, which at a city minimum wage of $17.27 would amount to $342,809 (Seattle Office of Labor Standards, 2015). By increasing mobility, the actual economic impact might be much greater because it could allow people to commute to better paying jobs, meaning that the program may indirectly lead to an improvement in household incomes for those who receive treatment.

The ability to enhance mobility, especially for underserved and economically vulnerable communities, underscores the program's potential to contribute positively to the community and the economy at large. Without knowing all the true costs, it is difficult for us to make a complete financial argument, but we advocate that—cost permitting—this intervention could significantly improve equity, especially for marginalized populations who benefit from this incentive.

Appendix D. Plot for Distribution of Ridership Between Treated and Controlled Units

In Figure D1, we now show the density plot of ridership for treated and control units. We observe that treatment shifts the distribution to the right, which is more noticeable at lower values in the distribution.

Due to the COVID-19 pandemic, King County has reported that commuting frequency plummeted substantially. Compared to close to half before the pandemic, less than 1 in each 10 respondents of a survey

sent out by King County Metro were riding the Metro at least once a month as of August 2020. The survey also noted that usage has dropped significantly in all major county sub-areas (King County Metro, 2020).

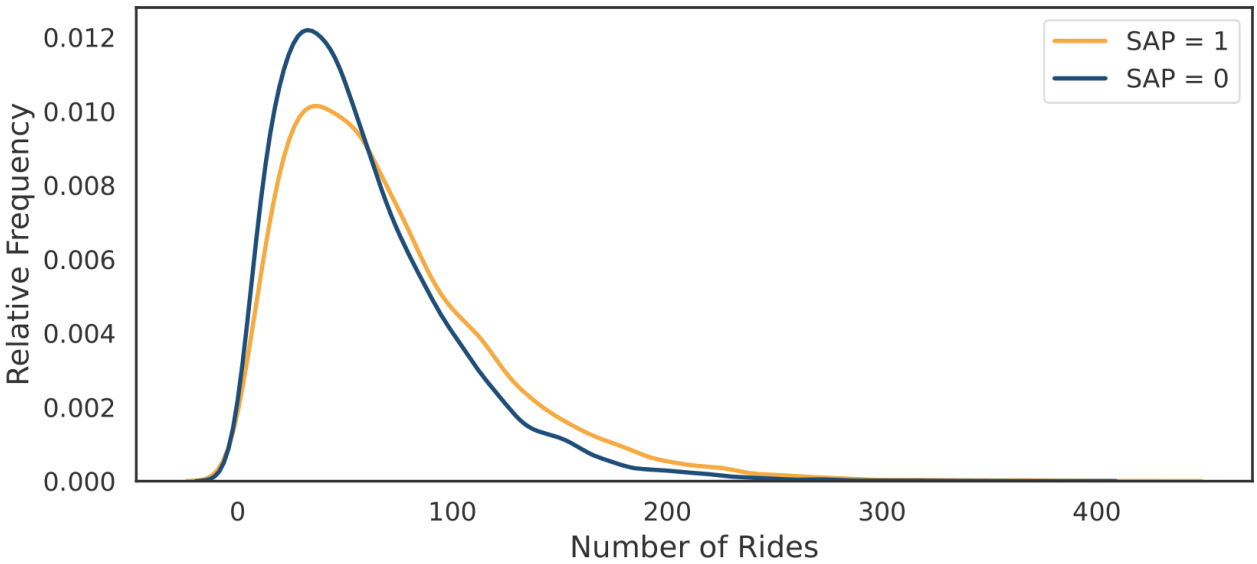


Figure D1. Distribution plot of ridership for treated and control units truncated to 1,000 rides to exclude right-tail outliers. SAP = subsidized annual pass.

Appendix E. Data Information

Table E1 shows descriptions of the variables used from Jenkins et al. (2020) as covariates. We used these covariates to better understand how SAP affected different populations in relation to their access to public transit.

Note that the "Number of low-wage workers" variable is the raw count of low-wage laborers in a census block rather than a normalized quantity, such as the fraction of low-wage workers in a particular census. We did this so as not to include the number of non-low-wage individuals in our calculations. (A small percentage of low-wage workers in a wealthy population would lead to a very low percentage, even if the number of low-wage workers in that area is large.)

Table E1. Descriptions of variables that appear in this article.

Feature	Data Set	Definition
Geographic Identification	Unequal Commute	Federal Information Processing Standard code for block group to identify geographic location.
Overall job access	Unequal Commute	Normalized measure of overall access to jobs for low-wage job seekers

Number of low-wage workers	Unequal Commute	Count of low-wage labor force in block groups from Longitudinal Employer-Household Dynamics (LODES) data.
Job access via transit	Unequal Commute	Normalized measure of jobs accessible within 30 minutes via public transit at peak hours.
Job access via transit at 3rd shift hours	Unequal Commute	Normalized measure of jobs accessible within 30 minutes via public transit at 3rd shift hours.
Spatial mismatch	Unequal Commute	Normalized measure of spatial mismatch between low-wage job seekers and jobs.
Agency	Card Registry	Agency where the card was issued.
Stage of card	Card Registry	Number of cards the rider has held. Cards expire every 2 years.
Race	Card Registry	Race of rider.
Language	Card Registry	Primary language that rider speaks.
Age	Card Registry	Rider's age group, binned as 18-24, 25-34, 35-44, 45-54, 55-64, and 65+.

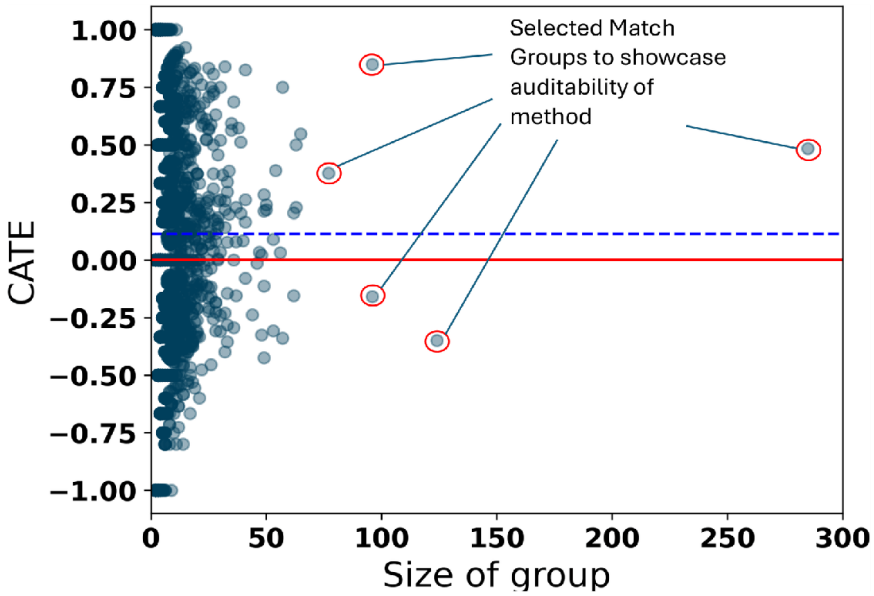
Appendix F. FLAME Matched Groups Interpretability

Fast Large-scale Almost Matching Exactly (FLAME) is fully interpretable in that it allows us to see and audit every single match group to validate findings. Our findings for every subgroup can be validated, allowing our method to be used easily by practitioners. Also, in the case where the matched groups are found to be problematic, our method permits easy troubleshooting. To illustrate the interpretable matched groups, Figure F1 shows several matched groups from FLAME for reenrollment. It is important to note that the *larger* a matched group is, the *greater certainty* and *less variability* we have in determining the group's treatment effect.

Let us look at an example row. For the first row in Figure F1, we show matched group number 1099 with 285 individuals. All individuals in this group are Asian, speak Chinese, are age 65+, in the 4th quartile for Overall job access, 4th quartile for Job access via transit, 2nd quartile for Spatial mismatch, and 3rd quartile for the Number of low-wage workers.

Some individuals in this group received the treatment, some did not. The collective conditional average treatment effect (CATE) estimate for this group is +0.484, suggesting that an individual in this group who

received the SAP treatment is likely to enroll +0.484 more times than the individuals in this group who did not receive the subsidized annual pass (SAP) treatment.

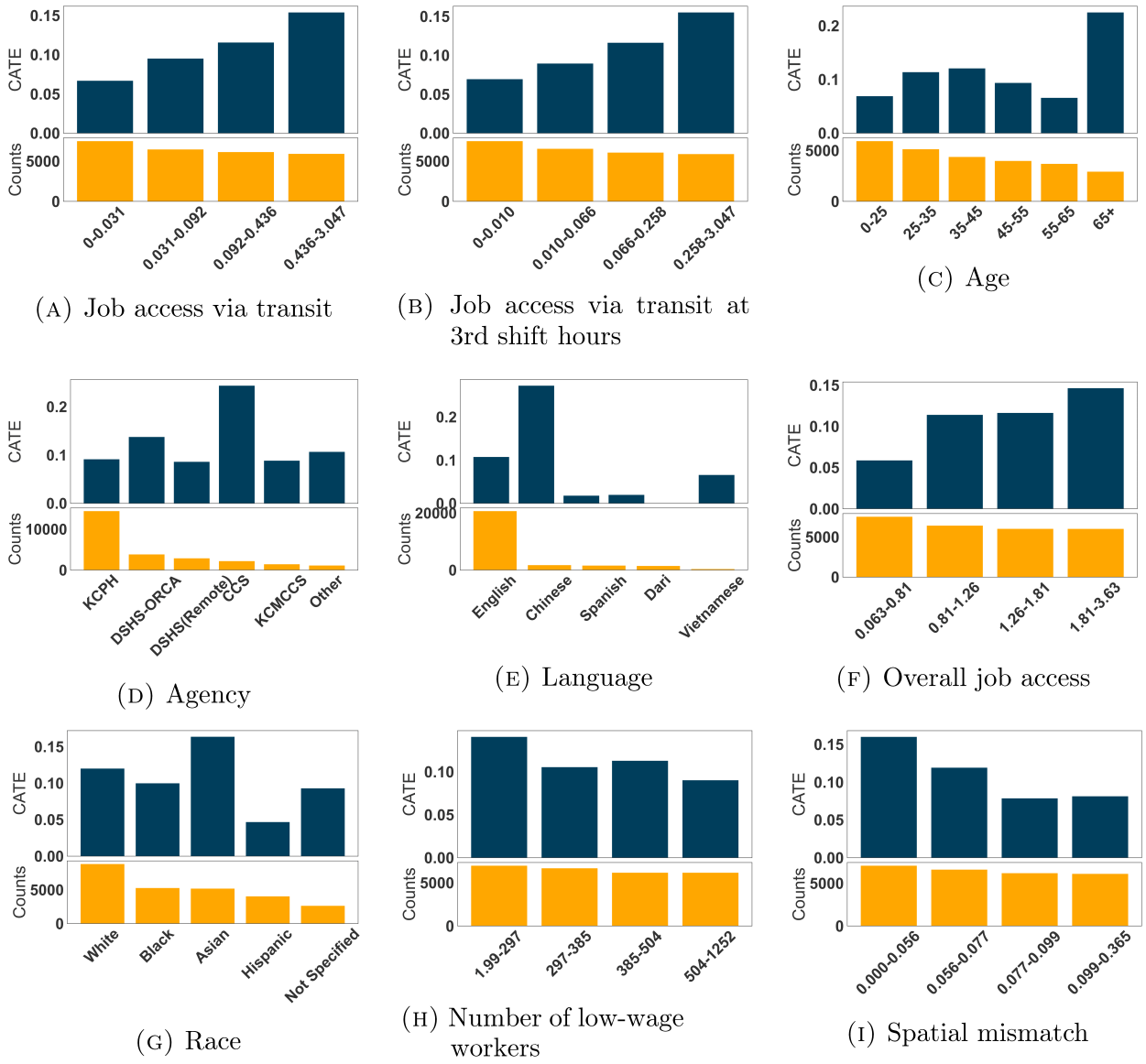


Match Group Number	Units in Match Group	Age	CATE	Race	Language	Overall job access	Job access via transit	Spatial mismatch	No. of low-wage workers
1099	285	65+	+0.484	Asian	Chinese	Q4	Q4	Q2	Q3
3225	124	-	-0.350	-	English	Q4	Q4	Q3	Q4
1094	96	65+	+0.849	Asian	Chinese	Q4	Q4	Q4	Q1
1373	96	25-35	-0.159	White	English	Q4	Q4	Q1	Q2
2435	77	65+	+0.377	Asian	Chinese	Q4	Q4	Q1	Q2

Figure F1. Characteristics of the selected fully matched groups created when computing the effect of subsidized annual pass (SAP) on reenrollment from receiving the SAP incentive using Fast Large-scale Almost Matching Exactly (FLAME). CATE = conditional average treatment effect.

Appendix G. Conditional Average Treatment Effects

Figures G1–G2 plot estimated conditional average treatment effects for reenrollment and long-term use (ridership after 6 months), respectively. These figures also show the number of people in each matched group.



(A) Job access via transit

(B) Job access via transit at 3rd shift hours

(C) Age

(D) Agency

(E) Language

(F) Overall job access

(G) Race

(H) Number of low-wage workers

(I) Spatial mismatch

Figure G1. Conditional average treatment effect (CATE) on *reenrollment* (dark blue) and number of people in each group (yellow). Using Fast Large-scale Almost Matching Exactly (FLAME), we created matched groups using the nine covariates above, specifically for subsidized annual pass (SAP) recipients eligible to reenroll in SAP. We exclude stage of card because most riders are on the initial stage of card so there are no heterogeneous effects.

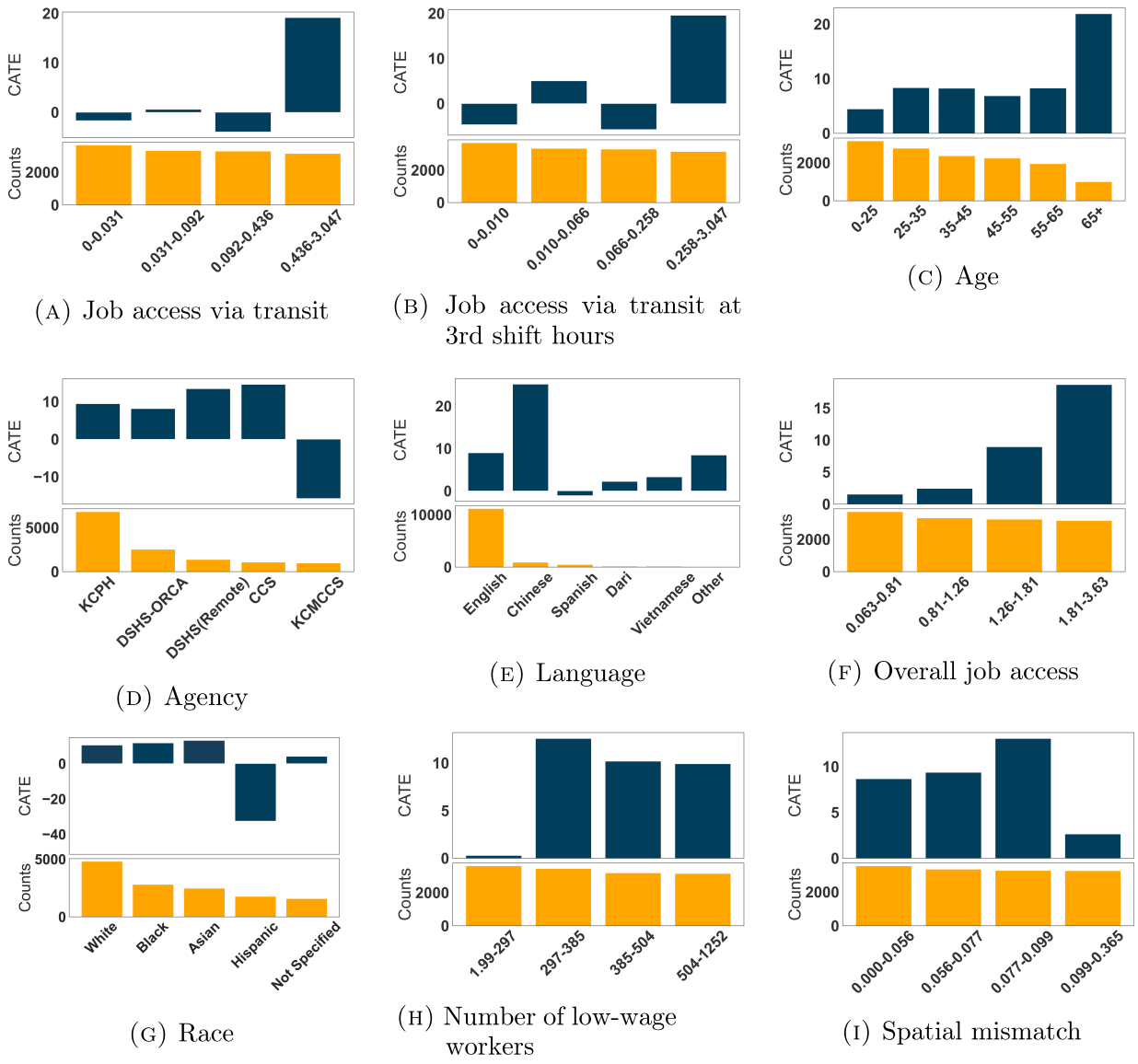
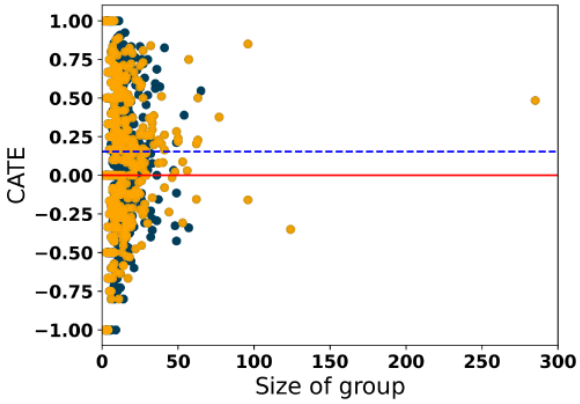


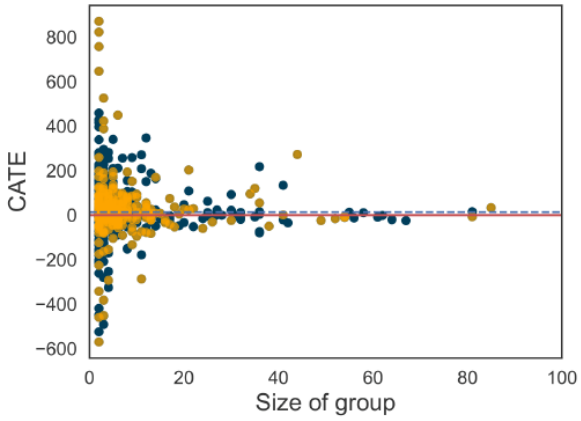
Figure G2. Conditional average treatment effect (CATE) on *ridership* (dark blue) and number of people in each group (yellow). Using Fast Large-scale Almost Matching Exactly (FLAME), we create matched groups using the nine covariates above, specifically for subsidized annual pass (SAP) recipients eligible to reenroll in SAP. We exclude stage of card as most riders are on the initial stage of card so there are no heterogeneous effects.

Appendix H. Effects of Subsidized Annual Pass (SAP) by Job Access via Transit

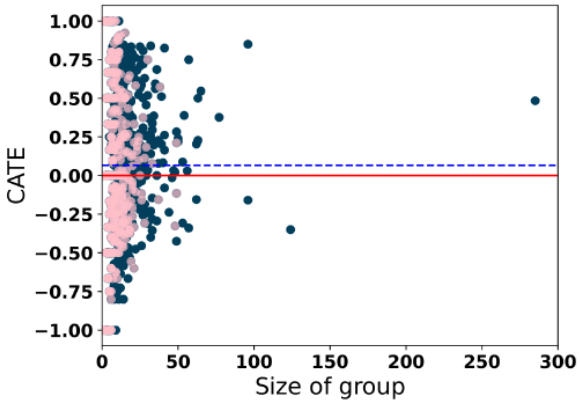
See Figure H1 for the conditional average treatment effect (CATE) of job access via transit on reenrollment and ridership.



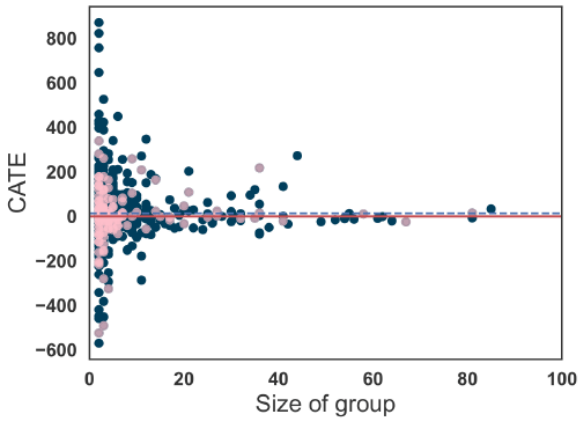
(A) Reenrollment, top quartile of job access via transit



(B) Ridership, top quartile of job access via transit



(C) Reenrollment, bottom quartile of job access via transit



(D) Ridership, bottom quartile of job access via transit

Figure H1. Effects of subsidized annual pass (SAP) on reenrollment (a, c) and ridership (b, d) by job access via transit. (a, b) Matched groups containing individuals in census blocks of top quartile of job access via transit (in orange) show the largest positive conditional average treatment effect (CATE) of +0.147, or a 14.7 percent higher reenrollment on average (a) and +14.86 more rides (b). (c, d) Matched groups of individuals belonging to census tracts with the bottom quartile of job access via transit (in pink) show a smaller treatment effect of +0.061 (c) and +1.26 rides (d).

©2025 by Gaurav Rajesh Parikh, Jenny Huang, Albert Sun, Lesia Semenova, and Cynthia Rudin. This article is licensed under a Creative Commons Attribution (CC BY 4.0) International license, except where otherwise indicated with respect to particular material included in the article.

Footnotes

1. This paper appeared at 2020 NeurIPS Workshop on Causality for Real-World Impact. Work presented in this paper won the 2022 American Statistical Association Data Challenge Expo Student Competition. ↩

References

- Agard, B., Morency, C., & Trépanier, M. (2006). Mining public transport user behaviour from smart card data. *IFAC Proceedings Volumes*, *39*(3), 399–404. https://doi.org/10.3182/20060517-3-FR-2903.00211 ↩
- Allen, J. (2019). Mapping differences in access to public libraries by travel mode and time of day. *Library & Information Science Research*, *41*(1), 11–18. https://doi.org/10.1016/j.lisr.2019.02.001 ↩
- Alonso González, M., Jonkeren, O., & Wortelboer-van Donselaar, P. (2022, August). *Equitable Transport Policy* [Research note]. Kennisinstituut voor Mobiliteitsbeleid. https://english.kimnet.nl/site/binaries/site-content/collections/documents/2022/08/08/equitable-transport-policy/KiM_Research+note+equitable+transport+policy_def.pdf ↩
- Anupriya, Graham, D. J., Hörcher, D., Anderson, R. J., & Bansal, P. (2020). Quantifying the ex-post causal impact of differential pricing on commuter trip scheduling in Hong Kong. *Transportation Research Part A: Policy and Practice*, *141*, 16–34. https://doi.org/10.1016/j.tra.2020.08.011 ↩
- Bagchi, M., & White, P. R. (2005). The potential of public transport smart card data. *Transport Policy*, *12*(5), 464–474. https://doi.org/10.1016/j.tranpol.2005.06.008 ↩
- Bhavsar, P., Safro, I., Bouaynaya, N., Polikar, R., & Dera, D. (2017). Machine learning in transportation data analytics. In M. Chowdhury, A. Apon, & K. Dey (Eds.), *Data analytics for intelligent transportation systems* (pp. 283–307). Elsevier. https://doi.org/10.1016/B978-0-12-809715-1.00012-2 ↩
- Brough, R., Freedman, M., & Phillips, D. C. (2022). Experimental evidence on the effects of means-tested public transportation subsidies on travel behavior. *Regional Science and Urban Economics*, *96*, Article 103803. https://doi.org/10.1016/j.regsciurbeco.2022.103803 ↩
- Center for Advanced Study in the Behavioral Sciences (CASBS). (2022). *Causal inference for Social Impact Lab's data challenge*. Stanford University. https://casbs.stanford.edu/causal-inference-social-impact-lab-s-data-challenge ↩
- Chapin, F. S. (1949). *Experimental Designs in Sociological Research*. Harper.

↩

- Chingos, M. M., & Blagg, K. (2017). *Student transportation and educational access*. Kennisinstituut voor Mobiliteitsbeleid. https://www.urban.org/research/publication/student-transportation-and-educational-access

↩

- Cinelli, C., & Hazlett, C. (2020). Making sense of sensitivity: Extending omitted variable bias. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *82*(1), 39–67. https://doi.org/10.1111/rssb.12348

↩

- Constantine, D. (2020, October 15). *King County Metro and Sound Transit offer new subsidized annual pass for riders in greatest need*. King County. https://kingcounty.gov/elected/executive/constantine/news/release/2020/October/15-Metro-fares.aspx

↩

- Davis, C., & Stacy, C. P. (2021, January 4). Four lessons for cities to help advance equitable transportation during the COVID-19 pandemic and beyond. *Urban Wire*. https://www.urban.org/urban-wire/four-lessons-cities-help-advance-equitable-transportation-during-covid-19-pandemic-and-beyond

↩

- Fisher, R. A. (1992). Statistical methods for research workers. In S. Kotz & N. L. Johnson (Eds.), *Breakthroughs in statistics: Methodology and distribution* (pp. 66–70). Springer. https://doi.org/10.1007/978-1-4612-4380-9_6

↩

- Gautier, P. A., & Zenou, Y. (2010). Car ownership and the labor market of ethnic minorities. *Journal of Urban Economics*, *67*(3), 392–403. https://doi.org/10.1016/j.jue.2009.11.005

↩

- Haltiwanger, J., Hyatt, H. R., & Spletzer, J. R. (2018). Job displacement and the duration of joblessness: The role of spatial mismatch. *The Review of Economics and Statistics*, *100*(2), 203–218. https://doi.org/10.1162/REST_a_00707

↩

- Halvorsen, A., Koutsopoulos, H. N., Lau, S., Au, T., & Zhao, J. (2016). Reducing subway crowding: Analysis of an off-peak discount experiment in Hong Kong. *Transportation Research Record*, *2544*(1), 38–46. https://doi.org/10.3141/2544-05

↩

↩

- King County GIS Center Spatial Data Catalog. (2019). *King County ACS Select Indicators*.

↩

- King County Metro. (2020, August). *COVID-19 Impacts Survey Results*. https://cdn.kingcounty.gov/-/media/king-county/depts/metro/documents/about/data-and-reports/2020/metro-covid-recovery-report.pdf?rev=929b37f5245d4b449ac2f76c5fcbd0d2&hash=6EF2269468A1DC8FF25D345B5E806648

↩

- King County Metro. (2020, August). *COVID-19 Impacts Survey Results*. https://kingcounty.gov/en/dept/metro/about/data-and-reports/-/media/depts/metro/accountability/reports/2020/covid-19-impacts-survey-results-aug2020.pdf

↩

- Koutsopoulos, H. N., Ma, Z., Noursalehi, P., & Zhu, Y. (2019). Transit data analytics for planning, monitoring, control, and information. In C. Antoniou, L. Dimitriou, & F. Pereira (Eds.), *Mobility patterns, big data and transport analytics* (pp. 229–261). Elsevier. https://doi.org/10.1016/B978-0-12-812970-8.00010-5

↩

- Lazo, L., Jayaraman, S., & Moriarty, D. (2022, February 23). D.C. traffic deaths at 14-year high with low-income areas hardest hit. *The Washington Post*. https://www.washingtonpost.com/transportation/2022/02/23/dc-traffic-deaths-highest-record/

↩

- McKenzie, B. (2015, December 16). *In what ways are D.C.-area neighborhoods with rail transit different from other neighborhoods?* United States Census Bureau. https://www.census.gov/newsroom/blogs/research-matters/2015/12/in-what-ways-are-d-c-area-neighborhoods-with-rail-transit-different-from-other-neighborhoods.html

↩

- Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, *1*(5), 206–215. https://doi.org/10.1038/s42256-019-0048-x

↩

- Schiff, A., Wright, L., & Denne, T. (2017). *Ex-post evaluation of transport interventions using causal inference methods* (research report 630). NZ Transport Agency.

↩

- Seattle Office of Labor Standards. (2015). *Minimum Wage Ordinance. Chapter 14.19 (Ord. 124960, § 45, 2015.)* https://www.seattle.gov/laborstandards/ordinances/minimum-wage

↩

- Semenova, L., Chen, H., Parr, R., & Rudin, C. (2023, December 10–16). *A path to simpler models starts with noise* [Poster presentation]. NeurIPS 2023: Thirty-seventh conference on neural information processing systems, New Orleans, LA, United States. https://openreview.net/forum?id=Uzi22WryyX

↩

- Semenova, L., Rudin, C., & Parr, R. (2022). On the existence of simpler machine learning models. In *FAccT '22: Proceedings of the 2022 ACM conference on fairness, accountability, and transparency* (pp. 1827–1858). Association for Computing Machinery. https://doi.org/10.1145/3531146.3533232

↩

- Silver, D., Blustein, J., & Weitzman, B. C. (2012). Transportation to clinic: Findings from a pilot clinic-based survey of low-income suburbanites. *Journal of Immigrant and Minority Health*, *14*(2), 350–355. https://doi.org/10.1007/s10903-010-9410-0

↩

- Skillen, W. H., & Abramsohn, E. E. (2021, July 29). Public transportation in the US: A driver of health and equity [Health policy brief]. Health Affai

↩

- Sound Transit. (2022). *King County Metro GTFS Feed Data* . https://www.soundtransit.org/help-contacts/business-information/open-transit-data-otd/otd-downloads

↩

- Syed, S. T., Gerber, B. S., & Sharp, L. K. (2013). Traveling towards disease: Transportation barriers to health care access. *Journal of Community Health*, *38*, 976–993. https://doi.org/10.1007/s10900-013-9681-1

↩

- U.S. Bureau of Transportation Statistics. (2022). *Transportation fatalities by mode* [Data set]. U.S. Department of Transportation. https://www.bts.gov/content/transportation-fatalities-mode

↩

- U.S. Census Bureau. (n.d.). *American Community Survey 5-year data (2009-2020)*. Retrieved April, 2022, from https://www.census.gov/data/developers/data-sets/acs-5year.html

↩

-

↩

- Wang, T., Morucci, M., Awan, M. U., Liu, Y., Roy, S., Rudin, C., & Volfovsky, A. (2021). FLAME: A fast large-scale almost matching exactly approach to causal inference. *Journal of Machine Learning Research*, *22*(1), Article 31.

↩

- Zhang, N., Graham, D. J., Hörcher, D., & Bansal, P. (2021). A causal inference approach to measure the vulnerability of urban metro systems. *Transportation*, *48*(6), 3269–3300. https://doi.org/10.1007/s11116-020-10152-6

↩

- Zheng, Y., Wang, S., & Zhao, J. (2021). Equality of opportunity in travel behavior prediction with deep neural networks and discrete choice models. *Transportation Research Part C: Emerging Technologies*, *132*, Article 132. https://doi.org/10.1016/j.trc.2021.103410

↩