CISIL Report: Question 1

Joseph T. Ornstein & Suhan Kacholia

June 30, 2022

Introduction

The COVID-19 pandemic had dramatic economic and social impacts on cities and counties across the United States. The pandemic especially affected transportation within cities—some estimates indicate public transit ridership decreased by 73% (Liu, Miller, & Scheff, 2020). In response to the pandemic, many public transit agencies instituted changes to their systems. King County Metro in Washington, which services Seattle and surrounding areas, is one such agency. Between March 21, 2020 until October 1, 2020, King County suspended fares for its transit services to mitigate the transmission of COVID-19. On October 1st, 2020, King County re-introduced fares. Investigating the impact of the suspension and reintroduction of fares on subsequent ridership would shed light upon the general effects of free-fare programs. Determining how the effect on ridership in King County differed based on socioeconomic characteristics provides additional nuance on the effects of fare-free transit.

King County, Washington is the 12th largest county in America, with an estimated population of 2.25 million. It contains the city of Seattle, as well as its surrounding suburbs, including Bellevue, Kirkland, and Kent. However, King County also contains large swaths of rural and forested areas. In King County, 7% identify as Black, 19% identify as Asian, 9.9% identify as Hispanic, and 5.2% identify with two or more races. The median household income of King County is $99,158, and 7.6% of the population lives below the poverty line. However, this masks substantial heterogeneity in poverty within different areas of King County. The King County Metro, possessing an annual budget of around $1 billion, thus serves a large, socioeconomically diverse population.

COVID-19 impacted Seattle relatively early in the pandemic, prompting a swift response by the city. On March 11th, Seattle imposed its first COVID-19 restrictions. The pandemic steeply reduced mobility in King County. Based on cellphone data, people in King County visited 57% fewer census block groups

(CBGs) between February and April 2020, excluding their homes. The pandemic also impacted public transit ridership, which declined by approximately 74% (Brough, Freedman, & Philips, 2020). This change was heterogeneous: high-income residents shifted to other modes of transportation to a higher degree than low-income residents, though this difference dissipated with time. The ability to work remotely played a large role in this result, while changes in the supply of public transit played a relatively small role.

In response to the pandemic, King County Metro suspended its fares. This fare suspension started on March 21, 2020, and for the duration, all riders did not have to pay to use services provided by King County Metro, including buses, water taxis, and Access paratransit. On October 1, 2020, King County reintroduced fares on the Metro. This fare reintroduction could have impacted the incentives of riders, causing some to use Metro services less, or substitute for other modes of transport. This effect may have differed based on demographic factors - public transit mode substitution is easier for higher-income residents, for example.

In this study, we examine the effects of the reintroduction of fares on ridership on the King County Metro. We study how this difference varies by socioeconomic status, race, and age across different Census tracts within King County.

Data and Methods

We use data provided by 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 data contains automated passenger counts at the stop-level for Metro buses. Passenger counter systems are not installed on every bus, so the data represents a sample of overall system ridership.

We also use demographic information at the Census tract level, using the 2019 American Community Survey. We also include COVID-19 case count data in our model, accessed through the King County COVID-19 data dashboard. Lastly, we include data on average precipitation and temperature for each day of our analysis, accessed through the National Centers for Environmental Information's Climate Data Online service.

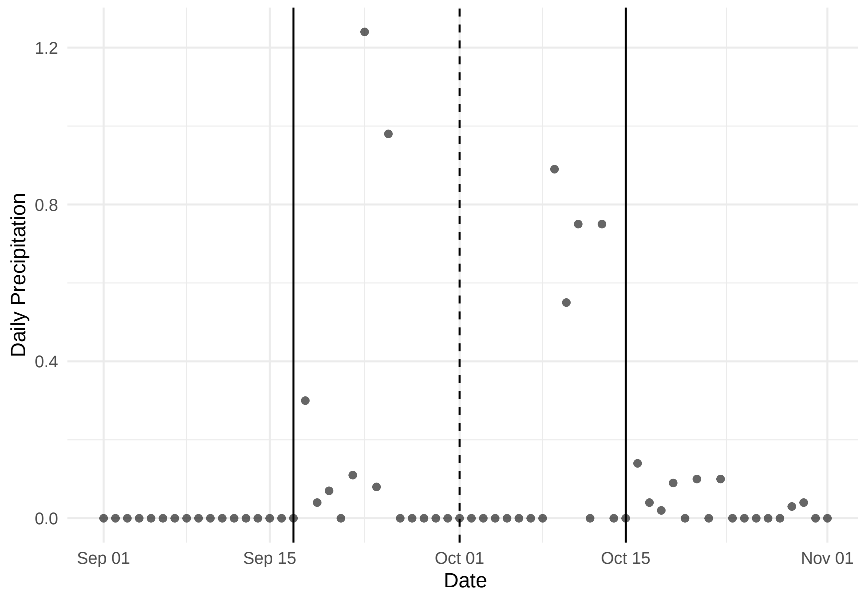
Research Design

To ensure that we are comparing ridership for the same services before and after the fare reinstatement, we analyze data at the trip level. Our outcome of interest is the number of boardings per trip, compared before and after the treatment. We conduct separate analyses for weekday and weekend routes, restricting our sample to trips with complete boarding data in the two weeks before and after the treatment.[1]

[1] King County notes that there were considerable service changes implemented on September 19, 2020.

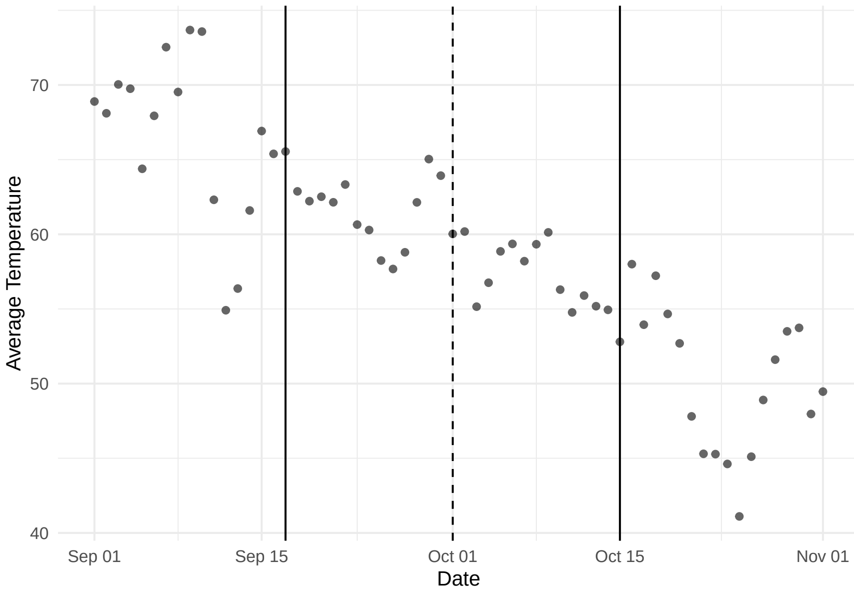
Because the fare reinstatement was implemented for every route on the same day, there is no comparison group we can use to estimate the causal effect. This means we cannot eliminate the possibility of time-varying confounders that might have affected ridership during this period. Perhaps ridership was unusually high in the two weeks prior to treatment – or unusually low in the two weeks post-treatment – for reasons other than the fare reinstatement. But to gain some confidence in the estimates, we can consider two important observable confounders: rates of COVID-19 and the weather.

Below, we display a graph measuring average precipitation in King County, with the two weeks surrounding the treatment date of October 1st marked with a dashed line:



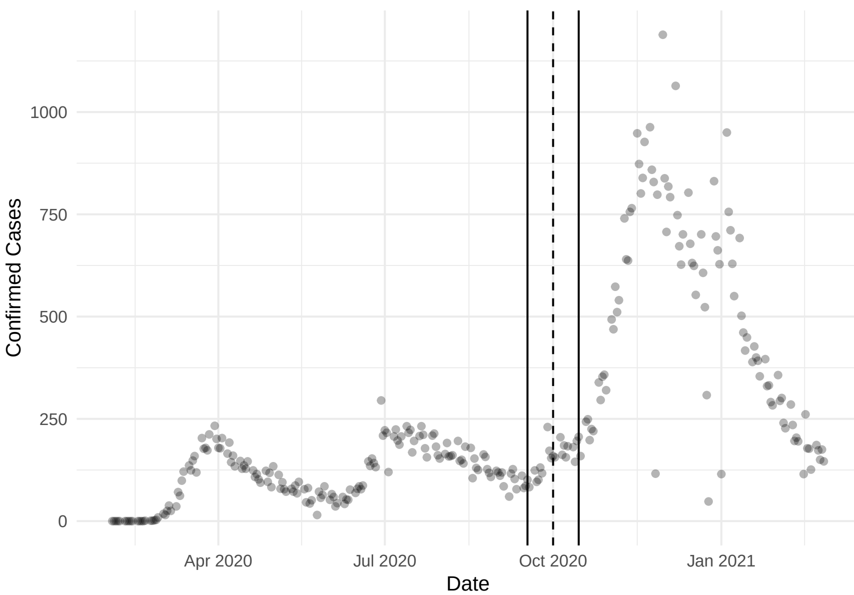
Rainier days could affect ridership, and there appear to be quite a few of them within the timeframe of the analysis. The two weeks surrounding the change appear more rainy than the weeks outside of the change, but there does not appear a noticeable pattern in precipitation levels between the two weeks on either side of the change themselves.

Next, we display a graph measuring average temperature in the county:



Temperatures steadily decreased between September 15th and October 15th, potentially affecting ridership.

Finally, the following graph displays confirmed COVID cases in King County in the two weeks surrounding the treatment:



There was a steady increase in cases during the timeframe of our analysis, which also may have impacted ridership.

To help alleviate concerns about these time-varying confounders, we include them as covariates in a series of fixed effects statistical models. We also include day-of-the-week as a variable, as ridership varies over the course of the week.

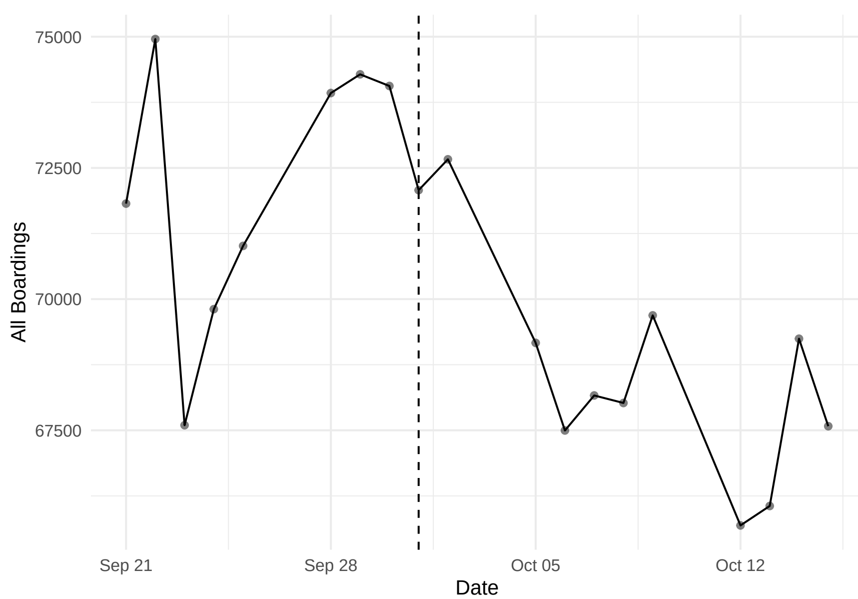
With these variables, we estimate a model of the following form, where Y_{it} denotes the number of boardings per trip, x_{it} is a treatment indicator, α_t denotes COVID cases per day, β_t denotes the average temperature per day, δ_t indicates the average precipitation per day, μ_i is a trip-level fixed effect, and γ_t is a day-of-the-week fixed effect:

$$Y_{it} = x_{it}\tau + \alpha_t + \beta_t + \delta_t + \mu_i + \gamma_t + \varepsilon_{it}$$

Standard errors are clustered at the trip-level.

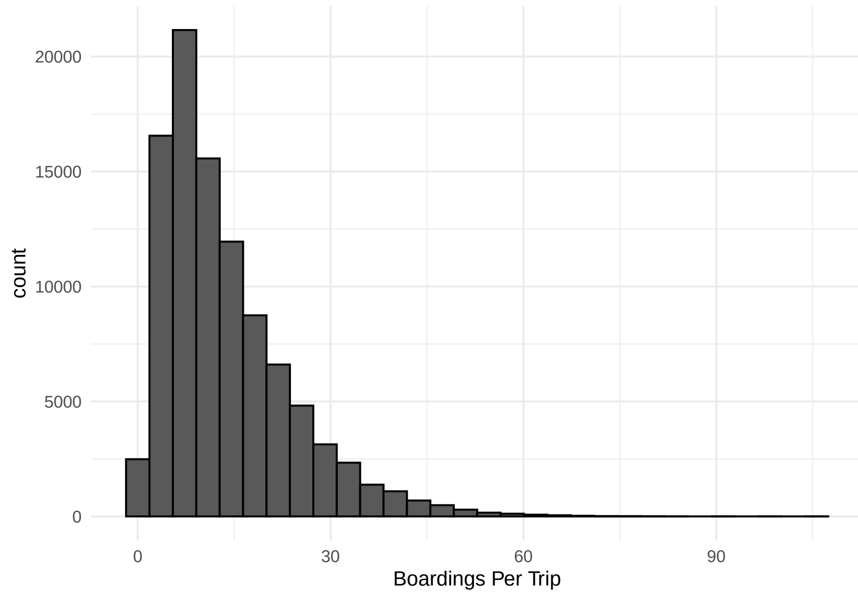
Results

The following is a graph displaying total passenger boardings in the four weeks of our analysis, with the treatment date of October 1st marked by a dashed line:



On an cursory glance, total boardings seem to decrease in the two weeks following the change. The peak occurs on September 22nd, with around 75,000 boardings, and reaches a minimum on October 12th, with below 66,250 boardings.

To compare ridership on similar services before and after the policy change, we restrict our analysis to 5,354 unique trip IDs which ran buses every day during the four week period. These buses completed a total of 97,775 unique trips during the period of study. We display the distribution of boardings per trip in the following graph:



The distribution is positively skewed, with a mean of 13.64 boardings per trip and median of 11.04.

To assess how the policy change affected ridership, we first estimate an Ordinary Least Squares regression of the passenger boardings against a dummy variable representing the treatment, the COVID-19 cases for that day, the average temperature for that day, and the precipitation, with trip ID and day of the week as fixed effects. Because ridership is a discrete count variable, we also estimate a Poisson regression using the same variables. The following table reports the coefficient estimates from those models.

	OLS	Poisson
Treatment	-0.639	-0.046
	(0.058)	(0.004)
COVID cases	0.002	0.000
	(0.001)	(0.000)
Average Temperature	0.085	0.006
	(0.009)	(0.001)
Precipitation	-0.663	-0.049
	(0.065)	(0.005)
Num.Obs.	97775	97756
RMSE	5.50	5.49

Both approaches yield very similar results. The first model suggests that following the fare reinstatement, there were 0.639 fewer boardings per trip. This is significant at the 99.9% level, and is equivalent to a roughly 4.5% decrease in ridership over the subsequent two weeks. The Poisson model also indicates a significant effect of the fare reintroduction on ridership, with an estimated marginal effect of 0.628 fewer boardings following the change. This is almost identical to the coefficient OLS model, adding confidence to our estimate. It is also equivalent to a roughly 4.5% decrease in ridership over the subsequent two weeks.

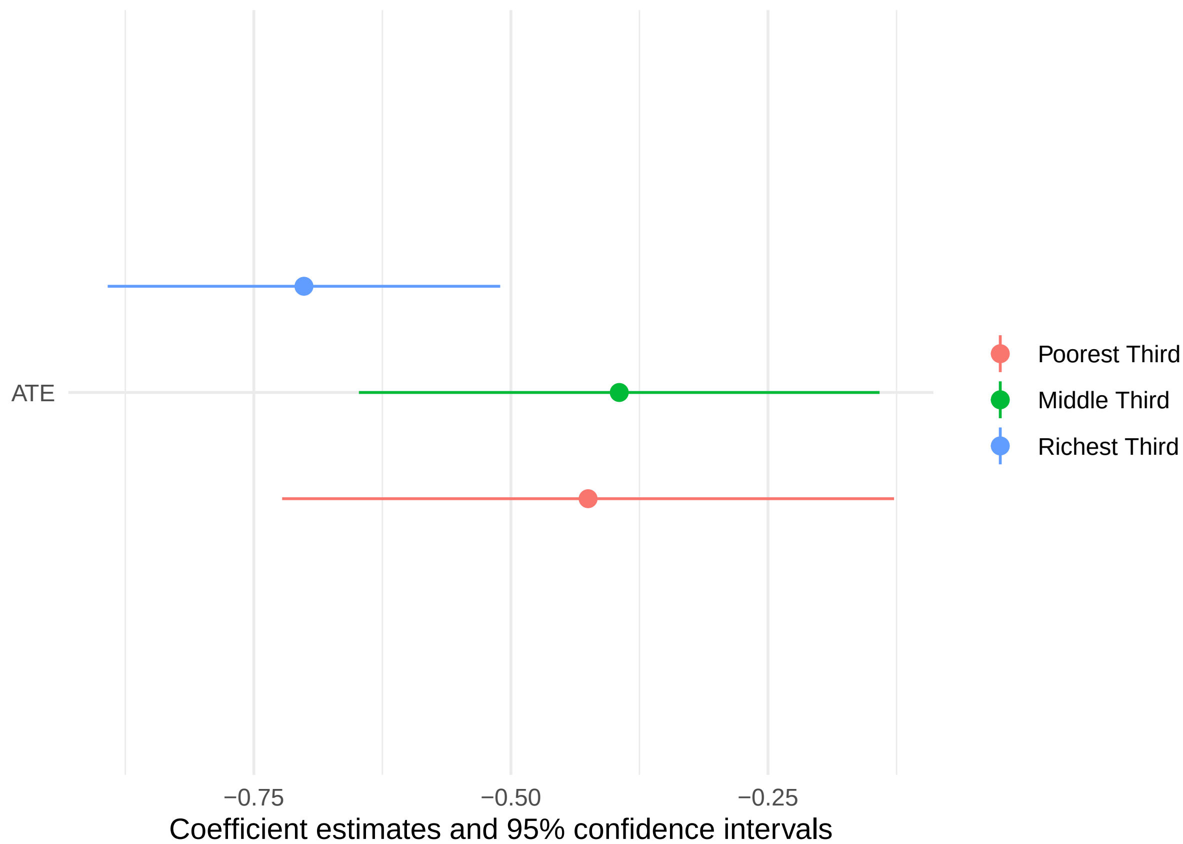
Heterogeneous Treatment Effects

The effect of the fare reintroduction on boardings could differ based on the demographic characteristics of a particular Census tract. We estimate these heterogeneous treatment effects by merging our trip data with demographic information provided by the American Community Survey, and adding the demographic information of the first stop to our model.

Socioeconomic group

We first estimate the impact of socioeconomic status on boardings. We divide Census tracts into three groups: High Income, Middle Income, and Low Income. Then, we compute the average treatment effect for each of these groups, using an Ordinary Least Squares regression. The resulting table and coefficient graph is displayed below:

	Poorest Third	Middle Third	Richest Third
Treatment	-0.425	-0.395	-0.701
	(0.152)	(0.129)	(0.097)
COVID cases	0.003	0.000	0.001
	(0.002)	(0.001)	(0.001)
Average Temperature	0.107	0.087	0.096
	(0.023)	(0.020)	(0.015)
Precipitation	-0.707	-0.614	-0.593
	(0.165)	(0.147)	(0.111)
Num.Obs.	19587	18149	21848
RMSE	6.47	5.23	4.37



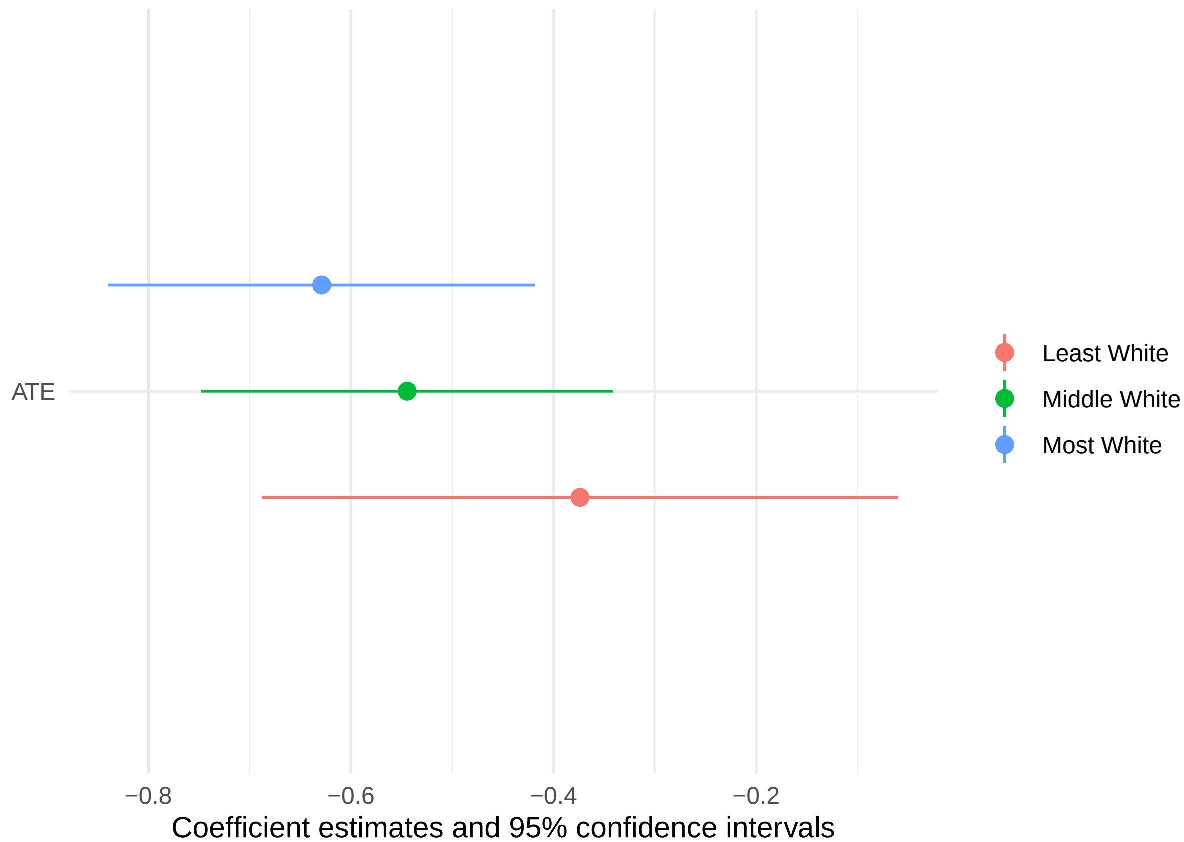
The fare reinstatement resulted in a 0.425 decrease in boardings for the poorest third of Census tracts, a 0.395 decrease for the middle third, and a 0.701 decrease in the richest third. The largest decrease in ridership thus seemed to occur in the wealthiest neighborhoods. However, the 95% confidence intervals for the coefficient estimates of the three groups overlap, so we cannot conclude that the reduction in boardings differs significantly.

This result might be explained by the comparatively higher degree of transit mode substitution wealthier residents are able to engage in. Having greater access to substitutes like cars or telecommuting, these residents display a greater elasticity with regards to the price of using the King County Metro.

Race

Next, we repeat the analysis, but this time using racial demographics. We again divide the data into three groups, this time corresponding to the percent composition of white residents. The results of this analysis is displayed below:

	Least White	Middle White	Most White
Treatment	-0.374	-0.544	-0.629
	(0.160)	(0.104)	(0.108)
COVID cases	0.002	0.002	0.000
	(0.002)	(0.001)	(0.001)
Average Temperature	0.109	0.098	0.084
	(0.024)	(0.016)	(0.016)
Precipitation	-0.801	-0.468	-0.641
	(0.177)	(0.116)	(0.121)
Num.Obs.	19673	19439	20472
RMSE	6.87	4.37	4.60

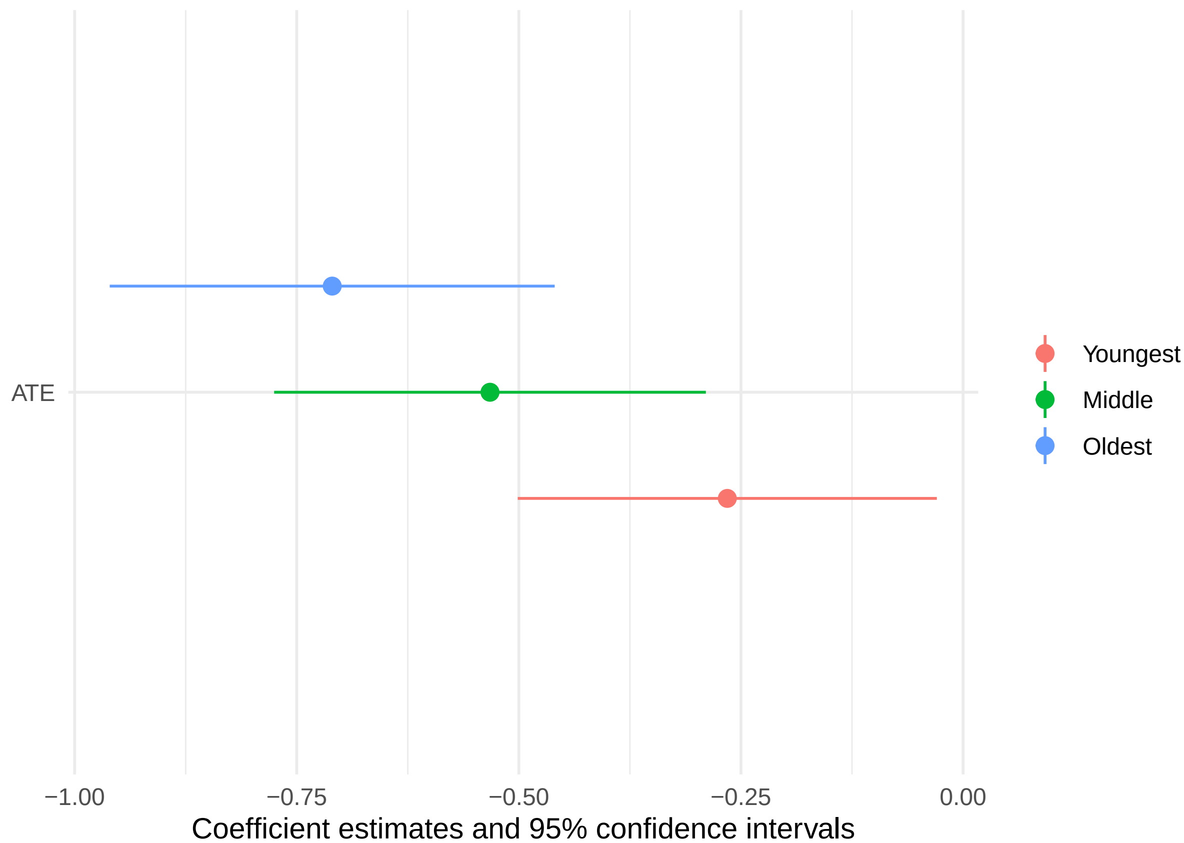


This analysis indicates that the least white Census tracts saw a 0.374 reduction in boardings, the middle third white Census tracts saw a 0.544 reduction in boardings, and the most white Census tracts saw a 0.629 reduction in boardings. The whitest Census tracts thus encountered the greatest average reduction in boardings. However, again, the 95% confidence intervals overlap, meaning we cannot conclude there was a significant difference in boardings based on the racial demographics of Census tracts.

Age

Lastly, we repeat the analysis based on median age.

	Youngest	Middle	Oldest
Treatment	-0.265	-0.532	-0.710
	(0.120)	(0.124)	(0.128)
COVID cases	-0.002	0.005	0.002
	(0.001)	(0.002)	(0.002)
Average Temperature	0.090	0.098	0.101
	(0.018)	(0.019)	(0.019)
Precipitation	-0.640	-0.411	-0.785
	(0.141)	(0.143)	(0.136)
Num.Obs.	19524	16087	23973
RMSE	5.21	4.77	5.90



The youngest Census tracts witnessed a 0.265 decrease in boardings, the middle aged Census tracts witnessed a 0.532 reduction in boardings, and the oldest Census tracts witnessed a 0.71 reduction in boardings. The oldest Census tracts encountered the greatest reduction in boardings. Younger residents might have less alternatives, like cars, to substitute for in the event of fare reinstatement.

In accordance with previous research, we find that some residents display a greater sensitivity to price changes in public transport, likely because of their ability to switch to alternative modes of transport or work from home. The wealthiest, oldest, and whitest Census tracts displayed the largest reduction in boardings. Further analysis is necessary, however, to determine whether these differences are statistically significant.

Limitations

One limitation of our approach is that we can only estimate the short-run elasticity of ridership upon price. Because we are only looking a month out, we cannot say, for instance, whether a fare reinstatement would increase ridership in the long-run by helping to improve service quality. The hetereogenous effects we found between socioeconomic groups might also differ, or even disappear, in the long-run. Previous research, for instance, found that high-income residents shifted disproportionately to cars in the beginning of the pandemic, but converged with lower-income residents in their public transit use later on. Perhaps ridership following fare reintroduction displays a similar pattern. Future research could repeat our analysis, but with a longer timeframe of ridership data, to understand the longer-term effects of fare reinstatement.

Additionally, while we included the most obvious and measurable time-varying confounders of precipitation, temperature, and COVID cases, there is no guarantee we eliminated all confounds. There might be alternative explanations of the decrease in ridership, which we do not include in the models. The fare reinstatement also occured during a global pandemic. This is a unique situation, meaning our results may not necessarily generalize to other time periods. Research on fare suspension and reinstatement outside of the pandemic context could help shed light on the broader dynamics of the policy.

Conclusion

Fare reinstatement in King County, Washington had a significant effect on ridership. Both OLS and Poisson estimates indicate around a 4.5% decrease in ridership in the weeks following the change. These estimates control for potential confounders like COVID-19 cases, precipitation, temperature, and day of the week.

We find the largest decreases in ridership in the wealthiest, whitest, and oldest terciles of the data. However, these differences might not necessarily be statistically significant. Our results correspond with previous research indicating that wealthier and more educated residents engage in comparatively more public transit mode substitution, such as during the start of the pandemic.

There are a number of limitations in our analysis. We advise caution in generalizing our results to broader public transit policy, as the fare reinstatement occurred in the midst of a pandemic, which is a highly unique situation. Additionally, we only estimate short-run effects of reinstatement. In the long-run, the policy could have different effects, including the heterogeneous treatment effects. Nevertheless, we believe our study presents a relevant and useful result for policymakers and researchers interested in understanding the price elasticity of public transit ridership across groups.

Works Cited

Brough, R., Freedman, M., & Phillips, D. C. (2021). Understanding socioeconomic disparities in travel behavior during the COVID-19 pandemic. *Journal of Regional Science*, 61, pp. 753-774. https://doi.org/10.1111/jors.12527.

Liu, L., Miller, H.J., & Scheff, J. (2020). The impacts of COVID-19 pandemic on public transit demand in the United States. *PLoS ONE*, 15(11). https://doi.org/10.1371/journal.pone.0242476