# The Effect of Transit Subsidies on Employee Transit Utilization in Smartcard Data

Rohan Aras
TRAC-UW
Seattle, Washington
rohana@uw.edu

## ABSTRACT

This research aims to understand how Commute Trip Reduction (CTR) transit subsidy programs, when controlling for various built environment variables and the structure of the transit network, impact the number of trips individual employees of large employers in the Central Puget Sound region take commuting from their worksite. This transit utilization is measured using the ORCA fare card records over two nine week periods in 2015 and 2016. Manipulating monetary costs is a known method of transportation demand management. Earlier preliminary research has suggested that these transit subsidies do have a significant impact on transit utilization. However, results in the wider literature suggest that transit utilization was operationalized in a way—defined on the level of an individual card without accounting for the existence of people who never take transit—that may have altered the influence of control variables. Indeed, some of the results were counterintuitive. In this research I attempt to avoid this by focusing solely on trips of employees of large employers to and from their worksites. This allows me to deduce how many employees are not utilizing transit and add them to the dataset. I then create a regression tree model that predicts the number of trips taken to and from an employer worksite on an individual card. The features of the model include subsidy values associated with each card, the closeness centrality of the stops around worksites weighted for travel time and headways—a measure designed to reflect the quality of transit service to that employer site—and the existence of employer provided parking. I find that higher centrality of worksites and higher pass subsidies both increase transit utilization while the existence of parking provided by the employer, whether free or paid, depresses transit utilization.

## KEYWORDS

Transit, Electronic Fare Payment, Seattle, Transportation Demand Management

## 1 INTRODUCTION

### 1.1 Background

In 1991, the Washington State Legislature passed the Commute Trip Reduction Law. The legislature, realizing that employers have a vested interest and significant role in the success of the transportation network, sought to improve air quality, reduce fuel consumption, and reduce traffic congestion through employer based programs. Fifteen years later, in 2006, the legislature passed the CTR Efficiency act to encourage local governments to work with employers to expand responsibility for program success and to align CTR programs with local transportation and land use policies. The CTR law requires that employers encourage their employees to use transportation options other than driving alone when commuting to worksites. Employers are then required to submit a program report for each of their affected worksites.

In 2009, seven transportation agencies in the Puget Sound region adopted a common electronic fare payment system called ORCA (One Regional Card for All). One of the original reasons for moving to the ORCA system from the existing paper based system was that the electronic system would provide a large quantity of data about the travel behavior of users. The transportation agencies hoped that this data could be used to improve regional transit planning. However, until fairly recently that data had not been used for that purpose. This year the UW was granted access to two nine—week chunks of ORCA data from spring 2015 and 2016, each containing roughly 21,000,000 transit boardings. Each boarding is associated with a timestamp, information about the route/trip and stop, card number and card metadata. This metadata includes the type of pass and whether the card is associated with a business. Three additional datasets were provided: Automatic Vehicle Location (AVL), Automatic Passenger Counts (APC), and CTR plans by business. Over this last summer a Data Science for Social Good team did a lot of work cleaning the data and created an additional origin/destination table that approximates where bus riders disembarked from their busses.

### 1.2 Questions and Significance

The goal of this project is to understand how Commute Trip Reduction (CTR) subsidy programs, when controlling for various built environment variables and the structure of the transit network, impact "transit utilization."

This analysis will be performed on the ORCA card records of Sound Transit and King County Metro. In part due to the transportation infrastructure and land use patterns that have developed over the last century, the private automobile has become the dominant mode of transportation for most Americans. A resultant outcome of this has been ever increasing traffic congestion, a phenomenon that has had a number of externalities with a negative impact on society as a whole. In 2005, congestion caused the average commuter in the United States to spend an extra thirty-eight hours traveling at with the additional fuel consumed costing of $710 per commuter. This time and fuel consumed has additional productivity, health, and environmental costs [5].

A potential solution to this problem is the creation and encouragement of viable alternatives to private automobile use that do not produce the same externalities. Transit in particular has the potential to be a viable alternative since the infrastructure—trains, buses, and boats—already exists in most large cities. Thus, the more abstract goal of this paper is to understand how people evaluate the costs and benefits of using transit versus using a private car. These costs are both monetary and logistical. Inasmuch as adjusting the monetary cost of the service through subsidies is a potentially widely applicable tool in adjusting the appeal of transit, understanding how subsidies affect utilization in the context of other transit properties is important part of this more abstract goal.

## 2 LITERATURE REVIEW

### 2.1 Travel Demand And Subsidies

Given the context of the Commute Trip Reduction program, it is important to understand the role of transit utilization within travel demand as a whole. Thankfully, there is a large body of work surrounding this topic. In "The Measurement of Urban Travel Demand," Daniel McFadden built a theory of travel demand as the agglomeration of travel decisions made by many individuals based on their own personal needs. The consequence of this formulation is that it is necessary to study those decisions made by the individuals to understand travel demand as a whole. McFadden asserts that, if we treat these individuals as rational economic consumers, "we can postulate a utility function summarizing the sense of well-being of [each individual] as a (decreasing) function of the level of deprivation [they] experience" [7]. This deprivation is shaped by the time, cost, and comfort associated with individual travel decisions.

Beginning in the 1970s, transportation planning started discussing how the existing transportation system could be better managed such that increasing travel demand could be satisfied without building more capacity. The resulting study and tools were termed Transportation Demand Management (TDM). In "Demand management as an element of transportation policy," Michael Meyer describes TDM as any set of actions designed to influence people's travel behavior by presenting alternative options for mobility. Examples of qualifying actions include increasing the price of parking or charging tolls. These tools are flexible in their application: TDM can either focus on mitigating current congestion problems or be strategic in attempting to avoid future congestion problems.

Two constituencies are defined as being particularly vital for the success of TDM projects: environmental groups and the business community. Environmental groups tend to support these policies due to a desire to reduce automobile use and associated environmental impacts. The business community often decides to support these policies when it concludes that TDM is good for business by increasing employee job satisfaction and retention, thus decreasing costs and improving productivity.

Meyer also notes that perceived costs are often the most important factor that influences travel decisions. Economists have long argued that car users do not pay the full costs of their travel. However, most voters feel that taxes are already too high. Thus, the political willingness to implement TDM policies simply does not exist in most urban areas. Therefore, any pricing scheme, in order to survive politically, cannot make voters feel worse off. Meyer suggests using subsidies rather than pricing to solve this problem, assuming that the funds are available [8].

Badoe & Yendeti in "Impact of Transit-Pass Ownership on Daily Number of Trips Made by Urban Public Transit" attempt to analyze a form of this subsidy based TDM: transit passes. They attempt to understand the impact of transit pass ownership on the number of trips that a person takes by transit using transportation survey data from the greater Toronto area. They found that transit pass owners had a transit trip rate about four times that of non-pass owners. After building a Poisson—number of transit trips displayed a Poissonian distribution—regression model that controlled for age, gender, employment, education status, occupation, whether the workplace was in the CBD, free parking, vehicles available, and household size, Badoe & Yendeti found that this relationship persisted, with ownership of a transit pass being the most important predictor [1].

Williams & Petrait cover a specific TDM transit pass subsidy program in "Upass: A Model Transportation Management Program That Works." The University of Washington U-PASS program was implemented in 1991 in an attempt to mitigate potential impacts from planned campus development. This program was a precursor and model to the programs designed to comply with the then new CTR program in the years that followed. It included increases in transit service, shuttle services, carpools, vanpools, commuter tickets, merchant discounts, and increased the prices of on-campus parking to market rate in order to pay for pass subsidies. This paper attempted to quantify the impacts of the program. They found that traffic decreased by 15 percent in the morning peak and 9 percent in the evening peak immediately after the program was implemented. Traffic continued to decline by 16 percent in the morning peak and 10 percent in the evening peak the following year. In addition, the commute mode share of driving alone went from 33 percent before U-PASS to 23 percent afterwards, and transit from 21 percent to 33 percent [11]. However, it is hard to parse out whether these effects were due to the subsidizing TDM (transit passes) or penalizing drive alone travel (increased parking fees) from the data presented in this study.

The main inspiration for this project was a draft paper written by Eric Howard: "CTR Programs, Subsidies, Built Environment, and Transit Level-of-Service, an analysis of factors that influence the rate of transit utilization using electronic fare card transaction records." Like my project, this paper attempts to evaluate the impact of CTR on transit utilization—specifically individual frequency of transit use—while controlling for the built environment and transit level of service. Howard's analysis was conducted on a portion of the dataset that I have access to: a somewhat less clean version of data from March of 2015. The built environment variables measured were bicycle infrastructure density, street intersection density, employment density, residential density, land use mix index, a park-and-ride indicator variable, Hispanic population proportion, non-white population proportion, proportion of population living below the poverty level, median age of the population, median household income, proportion of the population reported taking transit as the primary mode of commuting. Built environment variables were measured around the morning boarding for each card, a proxy for "home," and the afternoon boarding, a proxy for work location. Transit level of service variables were similarly measured around morning and afternoon boarding locations. These measures included average headways of routes serving the stop, the average headway during just the peak period, the total number of trips that service a stop, and the total connected ridership for each stop.

These variables were used in a series of five different Poisson regression models to account for limitations in the size of the dataset. These five models included the association between pass type and transit use; the association between subsidy value and transit use; the association between CTR programs, pass subsidies and transit use; the association between built environment, level-of-service on transit use; and the association between all of the variable mentioned in the previous four. A Poisson regression tree was also fit to the final model for the same reason as in the Bodoe & Yendeti study.

This paper found a number of interesting intuitive results âĂŞ Puget Passes were used 2.05 more than normal e-purses, a $10 increase in subsidy value was associated with a 2% increase in transit usage frequency. However, there were also some interesting unintuitive results. The biggest of these was that the availability of free parking had statistically significant (p « 0.05) increase in average frequency of transit usage [4]. While this may be a valid result, it could also be due to the construction of "transit utilization" as "average frequency of transit usage." It's possible that commuters who *did not* have the option of free parking were more likely to try transit a few times compared to commuters who *did* who might not take any trips on transit at all. This would skew the distribution of transit utilization for those who did not have the option of parking to the left in comparison, as riders who do not take transit do not appear in the dataset in the first place. This would skew the average higher for those who have access to parking. In general, this construction of transit utilization suffers from the potential for subsidies to fundamentally alter the distribution of number of trips taken. Bodoe & Yendeti note that those purchasing a transit pass (ostensibly at full price) do so with the intention to reduce the costs of their many mandatory trips [1]. Providing subsidies would suggest that that relationship no longer holds for all users, especially when the subsidy covers the entire cost of the pass.

## 2.2 Level of service

There are also additional ways in which to measure the level of service that a transit network provides that may prove more comprehensive. One potential way is to model the transit network a graph of nodes (stops) and edges (some sort of connection between stops) and then measure the centrality of each node. Sybil Derrible, in "Network Centrality of Metro Systems," applies this type of methodology to 28 "metro" systems across the globe. Noting the potential application of this sort of analysis in forecasting increases in ridership linked with opening new lines, this paper specifically attempted to capture the importance of various nodes as points of transfers. In order to do so, the paper focused on using betweenness centrality. Betweenness centrality measures the importance of a node $P$ as a transit point between any other pair of nodes by counting the number paths between the pair that go through $P$ [3].

In "Performance indicators for public transit connectivity in multi-modal transportation networks," Sabyasachee Mishra et al. discuss various other measurements of centrality. Degree centrality counts the number of nodes that each stop is connected to directly within the network. This fails to account for the number of nodes that are potentially indirectly accessible. Eigenvector centrality, acknowledging that connections to some nodes may matter more than others, assigns relative scores to each node that weight for these more important connections. Closeness centrality is defined as

$$D_{cc}(n) = \frac{\sum_{n_1 \in N} L_{n,n_1}}{N-1}, \forall N > 2$$

where $L_{n,n_1}$ is the distance between $n$ and $n_1$. [9].

## 2.3 Built environment impacts on transit

Modifications to certain built environment variables can also serve as a TDM strategy—similar to the price of parking and existence of tolls discussed by Meyer—and would need to be controlled for. In "Travel Demand and the 3Ds: Density, Diversity, and Design," Cervero and Kockelman note that "A host of urban design philosophies—new urbanism, transit-oriented development, traditional town planning—have gained popularity in recent years as ways of shaping travel demand" [2]. The people who espouse these philosophies argue that three dimensions—diversity, density, and design—all influence travel demand. This paper went on to lend some credibility to these claims: density, diversity, and design all have positive associations with reductions in Vehicle Miles Traveled (VMT) and increases in non-motorized transportation.

Moudon & Stewart, in a meta-study of the variables associated with VMT and non-motorized transportation (NMT) usage, "Tools for Estimating VMT Reductions from Built Environment Changes," adds a few more "D's" to the list articulated by Cervero and Kockelman: destination accessibility, distance to transit, and demand management. They state that these variables are typically measured in the "neighborhood" surrounding an individual's residence or workplace. These are typically defined by using quarter mile to mile buffers, census definitions, or traffic analysis zones (TAZ).

Moudon & Stewart define density as a measure of the concentration of population, dwellings, or employment in a given land area. Increases in density were found to have been consistently correlated with greater walking and transit, and decreases in VMT. Density was also found to be correlated with other built environment variables that influence transportation choice: transit service, auto ownership, destination accessibility, distance to CBD, land use mix, gridded street patterns, and incomes. Studies that controlled for these covariates found that the strength of association between density and VMT or NMT usage was much weaker than when not controlling for these covariates. As such, Moudon & Stewart conclude that density, defined as people per unit area, might not influence travel decisions by itself.

In the same meta-study, diversity in the built environment context is defined as the variety of land uses in a given area. Some measures of diversity simply look at the percentage given to each type of land use while others attempt also measure the complimentary nature of certain land uses. Moudon & Stewart found that higher diversity in the Puget Sound region was found to be correlated with reduced VMT.

Moudon & Stewart define design as the spatial layout of streets and blocks. Typically this refers to the density of intersections as a measure of pedestrian oriented street network connectivity. Moudon & Stewart found that the effects of street network connectivity on utilitarian walking were mixed across the literature as a whole. They note that this could be due to potential land use differences between small blocks in urban vs suburban areas—small blocks before 1930 had diverse land uses, whereas small blocks in post WWII suburbs were largely low density sub divisions. Within the Puget Sound area, however, they found that street network connectivity measures were found to have been broadly positively associated with walking. Studies reviewed by Moudon & Stewart found that smaller block sizes were correlated with walking and that street network connectivity was negatively correlated with VMT and positively correlated with non-motorized miles traveled. However, they also noted that these measures are generally flawed—street network connectivity ignores other design elements of streets such as the existence of sidewalks and other pedestrian amenities that could also have positive or negative effects on NMT usage [10]. Some of these amenities are discussed in Maghelal's "Walking to Transit: Influence of Built Environment at Varying Distances." This study sought to understand how built environment variables influenced the percentage of people who walked to Dallas Area Rapid Transit stations. It found that road speed and road shoulders had negative associations with the percent of transit riders who walked to stations. Sidewalk density had a positive association [6].

In the Moudon & Stewart report destination accessibility is defined as a measure of the proximity from home to a number of specific places. They found that these measures consistently have strong relationships with both utilitarian walking and reductions in VMT in the Puget Sound region.

Moudon & Stewart found that distance to transit and the number of stops in a given area are strongly related to transit use. They also found that most transit riders walking to transit do so within a quarter mile of a stop. About half as many do so between a quarter mile and a half mile. While numerous studies that Moudon & Stewart reviewed found that people tend to walk longer distances to rail rather than busses, they noted that this could be due to differences in service attributes rather than the type of vehicle.

Moudon & Stewart define demand management as policies or programs designed to influence demand of different modes of transportation. With regards to the built environment, they found that one study on King County trip data found a negative association between parking costs at the end of trips and household VMT. A separate study on PSRC data found that that parking prices and free parking availability near homes did not have an effect on travel mode for trips between home and work or trips between home and other places.

Moudon & Stewart also noted that a number of composite measures of built environment variables have been defined. One interesting example mentioned was the use of k-means clustering on several built environment variables from PRSC data to identify eight built environment contexts. A similar study was conducted in France where they used the categories produced by the model to predict walking and cycling.

Moudon & Stewart also covered some survey based research on whether perceived safety and comfort affects transportation decision making process. Surveys in Portland and the San Francisco Bay Area found that transit users who walked to stations largely agreed that slower traffic speeds, sidewalks, and traffic devices affected their choices of routes to stations. Moudon & Stewart briefly cover how perceived safety from crime may negatively influence physical activity including NMT. They were quick to note, however, that much of the existing research is limited, since the surveys they rely on don't specify the sources of feeling unsafe. Moudon & Stewart also note that studies in Boston and Singapore found that elevation changes, including stairs, represented a disutility to walking.

# 3 METHODS

## 3.1 Data

Data were gathered from the ORCA transactions and King County Metro (KCM) Automatic Vehicle Location (AVL) databases over nine week period corresponding to KCM's 2016 spring service revision. The ORCA transactions dataset contains a pseudonymized record for each tap onto an ORCA card reader. Each record contained a hashed card id, the associated business (if relevant) stop, route, time of tap and ORCA product of the card. ORCA product refers to whether the card is an ePurse, Puget Pass or part of a Business Passport program. The AVL dataset contains a record for each time a vehicle on a specific trip in the KCM system reached a stop along its route. A trip uniquely identifies a route along a certain series of stops on a certain day of the week and at a certain time. The two datasets contain over 23 million and 31 million records respectively.

Data were also gathered from the 2014 Commute Trip Reduction mandatory survey. The survey collects information on each worksite owned by every employer who participates in the CTR program. This information includes the address of the main office at the worksite, the number of employees located there, the amount of parking available on and off campus owned/leased by the employer and their pricing, the existence of other free or paid parking off site, and some information on transportation subsidies. Parking data were condensed into binary variables $P_w$ and $F_w$ indicating the existence of paid parking and free parking owned/leased by the employer respectively for each worksite $w$. The existence of additional offsite parking not controlled by the employer was ignored This was done to account for the varying geographic expanses of different worksites in the data and potentially inaccurate reporting to the survey.

Additional data on the Sound Transit Link Light Rail schedule were collected from the Sound Transit website. Data on the cost of Puget Passes and cost of Business Passport programs by geographic location were collected from the ORCA and King County Metro websites respectively.

Subsidies in the CTR survey data are broken out into the dollar amount that an employee's transit pass is subsidized by the employer per month, the dollar amount of the maximum monthly bus subsidy per employee per month, the dollar amount of the maximum monthly non-specific transportation allowance per employee per month and the percentage subsidy that an employer pays toward an employee's transit pass where applicable. We ultimately calculated the percentage subsidy $S_{pct}$ on transit passes for each worksite given the other subsidy fields and the known costs of Puget Passes and Business Passport programs.

## 3.2 Calculating Transit Utilization

Transit utilization $T_i$ is defined as the number of trips taken by employee $i$ from their most used stop $s_i$ near their worksite $w$ on a weekday. In order to determine the worksite of each employee, we first define certain stops as serving certain worksites. To do so we merge all worksites that are located within a half mile of each other and their associated metadata. The new locations of the merged worksites are defined as the centroid of the component worksites. The closest stops to these worksites that account for 70% of employee boardings within a mile are defined as servicing the worksite. Employees are defined as working at the worksite served by the stops that they commute from the most.

## 3.3 Accounting for Non-Transit Riders

We define “Non-Transit Employees” as employees who do not show up in the ORCA transactions dataset as they have not used an employer subsidized ORCA card. These employees have a $T_i$ of 0. The number of non-transit employees $D_w$ is found by subtracting the number of employees that commute from a worksite on transit $R_w$ from the total number of employees recorded as working at that worksite in the CTR survey data $E_w$. In some cases, the number of employees recorded as commuting from a worksite is larger than the number of employees recorded as working at a worksite due to the CTR survey being two years older than the ORCA transactions data. In these cases we naively adjust the total number of employees working at worksite by pegging the number of employees who commute using transit to the maximum percentage of riders using transit less than 100% across all other worksites.

Each non-transit employee is are assigned an ORCA product in the same distribution as transit riding employees at their worksite. Similarly, each non-transit employee is assigned a "most used stop" in the same distribution as transit riding employees at the same worksite.

## 3.4 Defining Transit Network Structure

The transit network structure is defined in this work as the closeness centrality $C_s$ of each stop in the transit network, as discussed in Section 2.2, over five periods throughout the day: early morning (4:30 to 6:00), morning peak (6:00 to 9:00), midday (9:00 to 15:00), afternoon peak (15:00 to 18:00), and evening (18:00 to 22:00).

Specifically, we run the Closeness Centrality algorithm over a graph of $N$ nodes $s$, one for every stop in the KCM and KCM run ST system. Two nodes $s_1$ and $s_2$ are connected by an edge if they are within a quarter mile straight line distance from each other or a trip $j$ of any route connects the two stops. The Open Source Routing Machine (OSRM) is used to calculate the time it takes to walk ($t_{walk}$) between $s_1$ and $s_2$ if they are within a quarter mile of each other. The AVL data and Link Schedule are used to calculate the average route agnostic headway ($t_{head}$)—over the nine week period—preceding each trip $j$ and the average vehicle travel time ($t_{veh}$) on $j$ between $s_1$ and $s_2$ for each of the aforementioned periods. Since a transit rider can ideally ignore schedules and try to make a journey at any time we assume that transit riders will arrive at their initial stop $s_1$ in a uniformly random fashion. Thus we define the total travel time between $s_1$ and $s_2$ ($t_{total}$) as

$$t_{total} = \frac{t_{head}}{2} + t_{veh}.$$

The edge in the graph between $s_1$ and $s_2$ are weighted by the smaller of $t_{walk}$ and $t_{total}$.

## 3.5 Predicting Transit Utilization

In order to understand the relationships between transit utilization, CTR transit programs, parking availability, and transit network structure, we define a series of seven different linear regression models. For this analysis we limit the ORCA transactions data used to the month of April to avoid needing to handle changes in ORCA product type for individual employees. Differences between Business Passport programs and Puget Passes are tested for due to the large differences in how the programs are structured and benefits to the employee.

*3.5.1 Single Variable Models.* The association of transit utilization and pass subsidy is modeled using a linear regression of the form

$$T_i = \alpha + \beta S_i$$

where $S_i$ is the percentage amount that employee $i$'s transit pass is subsidized given their worksite and ORCA product. $T_i$ is the transit utilization for that employee as defined in Section 3.2.

The association of transit utilization and parking is modeled using a linear regression of the form

$$T_i = \alpha + \beta_1 F_w + \beta_2 P_w$$

where $F_w$ and $P_w$ are the worksite specific free/paid parking indicator variables defined in Section 3.1.

The association of transit utilization and transit network structure is modeled using a third linear regression model of the form

$$T_i = \alpha + \beta C_{s_i}$$

where $C_{s_i}$ is the centrality of the most used stop of employee $i$ as described in Sections 3.2 and 3.4, with $T_i$ binned into time periods corresponding to the centrality measure.

Lastly, the association between ORCA product type and transit-riding employee percentage was modeled as a linear regression of the form

$$\frac{R_w}{E_w} = \alpha + \beta Pr_w$$

where $R_w$ and and $E_w$ are the number of transit-riding employees and total employees respectively as defined in Section 3.3. $Pr_w$ is a binary indicator variable indicating that the worksite $w$ uses a Business Passport program rather than subsidizing Puget Passes.

*3.5.2 Combined Models.* We perform a full analysis of the association between transit utilization, pass subsidies, parking, and transit network structure to understand the association of subsidies and transit utilization when controlling for parking and the structure of the transit network. These associations are modeled using a linear regression of the form

$$T_i = \alpha + \beta_1 S_i + \beta_2 F_w + \beta_3 P_w + \beta_4 C_{s_i}$$

with $T_i$ binned into time periods corresponding to the centrality measure as in Section 3.5.1. This regression is also performed on the data of solely employees who have a Business Passport and solely those who have a Puget Pass.

# 4 RESULTS AND DISCUSSION

## 4.1 Transit Network Structure

The top ten most central and least central stops in the KCM/KCM run Sound Transit system are shown in Tables 8 and 9 respectively.

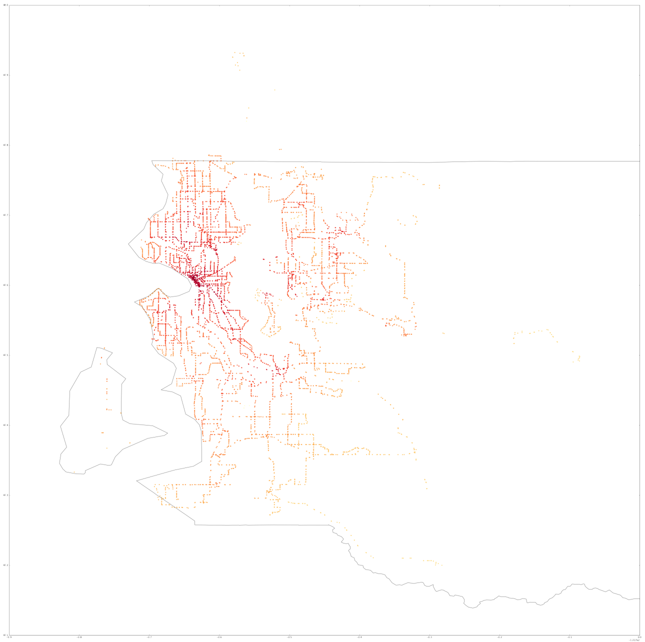

**Figure 1: Map of stops in King County colored by their centrality score during afternoon peak. Yellow stops are less central while red stops are more central.**

Figure 1 shows all stops mapped and labeled according to their centrality in the afternoon peak. All stops are mapped and labeled by centrality during all periods in Section A. The most central stops are located in Downtown Seattle throughout all five time periods. The top 10% most central stops in the system largely fall within the boundaries of Seattle as well—notably Downtown Seattle and the University District. Many of these stops also fall along frequent transit routes including the 7 (Rainier Avenue), 8 (crosstown through South Lake Union), 45 (crosstown through North Seattle), 48 (crosstown through the Central District), the E-Line (Bus Rapid Transit down Aurora Avenue), and several Link Light Rail stations. That said, Downtown Renton, the Evergreen Point Freeway Stations, and Mercer Island Park and Ride appear central throughout the day as well.

Stops also have higher centrality scores overall during morning and afternoon peak and the lowest scores in the evening period. This suggests that stops are closer during morning and afternoon peaks as would be expected.

## 4.2 Models

*4.2.1 Single Variable Models.* Table 1 presents the results of the Transit Utilization vs Pass Subsidies model. Subsidy percentage has a positive statistically significant effect ($p <<< .05$) on transit utilization. The $R^2$ for this model is 0.001. While the p Value for subsidy percentage is small suggesting the positive relationship is statistically significant, the small $R^2$ value indicates that subsidy percentage alone is not very predictive of an employee's transit utilization. This is not unexpected as this simple model does not account for the other factors being examined below.

**Table 1: Transit Utilization vs. Pass Subsidies Model Results**

| | Coefficient | p Value | 2.5% | 97.5% |
|---|---|---|---|---|
| $\alpha$ | 0.0888 | 0.000 | 0.071 | 0.107 |
| $\beta$ | 0.2441 | 0.000 | 0.225 | 0.263 |

**Table 2: Transit Utilization vs. Parking Model Results**

| | Coefficient | p Value | 2.5% | 97.5% |
|---|---|---|---|---|
| $\alpha$ | 0.4615 | 0.000 | 0.456 | 0.467 |
| $\beta_1$ | -0.3754 | 0.000 | -0.384 | -0.366 |
| $\beta_2$ | -0.2804 | 0.000 | -0.294 | -0.267 |

**Table 3: Transit Utilization vs. Centrality Model Results**

| | Coefficient | p Value | 2.5% | 97.5% |
|---|---|---|---|---|
| $\alpha$ | -0.7319 | 0.000 | -0.752 | -0.712 |
| $\beta$ | 4601.7855 | 0.000 | 4514.923 | 4688.648 |

**Table 4: Percent Transit Riders vs. ORCA Product**

| | Coefficient | p Value | 2.5% | 97.5% |
|---|---|---|---|---|
| $\alpha$ | 0.0337 | 0.053 | -0.000 | 0.068 |
| $\beta$ | 0.1928 | 0.000 | 0.148 | 0.238 |

Table 2 describes the results of the Transit Utilization vs Parking model. Both the availability of free parking and paid parking provided by the employer had a statistically significant negative effect ($p <<< .05$) on transit utilization. The $R^2$ for this model was 0.013.

Table 3 presents the results of the Transit Utilization vs Transit Network Structure model. Closeness centrality of the main workplace stop had a statistically significant positive effect ($p <<< .05$) on employee transit utilization. The $R^2$ for this model was 0.019.

Table 4 describes the results of the Percent Transit Riders vs. ORCA Product model. Having a Business Passport program has a significant ($p <<< .05$) positive effect on the percentage of employees who ride transit as opposed to simply subsidizing Puget Passes, suggesting that the incentives/disincentives or magnitude of incentives/disincentives for taking transit could be different under the two programs. This makes some intuitive sense since the cost of the Business Passport, whose cost is based on the average use of employees, should cost less to the employee than a comparable Puget Pass which is based on the value of the card.

*4.2.2 Combined Models.* The results presented by the combined transit utilization model in Table 5 expose the same effects as the individual single independent variable models. Employee subsidy percentage has a statistically significant ($p <<< .05$) positive impact on employee transit utilization. Both measures of parking—the availability of free parking provided by the employer or the availability of paid parking provided by the employer—have significant ($p <<< .05$) negative effects on employee transit utilization. Finally,

**Table 5: Combined Transit Utilization Model Results**

| | Coefficient | p Value | 2.5% | 97.5% |
|---|---|---|---|---|
| $\alpha$ | -0.6985 | 0.000 | -0.729 | -0.668 |
| $\beta_1$ | 0.1900 | 0.000 | 0.171 | 0.209 |
| $\beta_2$ | -0.1948 | 0.000 | -0.205 | -0.185 |
| $\beta_3$ | -0.3173 | 0.000 | -0.331 | -0.304 |
| $\beta_4$ | 4082.1480 | 0.000 | 3985.863 | 4178.433 |

**Table 6: Combined Transit Utilization Model Results: Business Passport**

| | Coefficient | p Value | 2.5% | 97.5% |
|---|---|---|---|---|
| $\alpha$ | -0.7456 | 0.000 | -0.783 | -0.708 |
| $\beta_1$ | 0.0799 | 0.000 | 0.056 | 0.104 |
| $\beta_2$ | -0.2183 | 0.000 | -0.230 | -0.207 |
| $\beta_3$ | -0.2925 | 0.000 | -0.311 | -0.274 |
| $\beta_4$ | 5008.8954 | 0.000 | 4892.020 | 5125.770 |

**Table 7: Combined Transit Utilization Model Results: Puget Pass**

| | Coefficient | p Value | 2.5% | 97.5% |
|---|---|---|---|---|
| $\alpha$ | -0.3897 | 0.000 | -0.433 | -0.346 |
| $\beta_1$ | 0.1208 | 0.000 | 0.094 | 0.147 |
| $\beta_2$ | -0.0698 | 0.000 | -0.088 | -0.051 |
| $\beta_3$ | -0.0696 | 0.000 | -0.087 | -0.052 |
| $\beta_4$ | 1775.4958 | 0.000 | 1634.580 | 1916.411 |

closeness centrality has a significant ($p <<< .05$) positive impact on employee transit utilization. The $R^2$ for this model is 0.025.

The direction of these effects hold when the data is split between Business Passport programs and Puget Pass only programs. However, the magnitude of these effects changes differently for each. In the Business Passport case, the effects of free parking availability and centrality both are larger, while the effects of subsidies and paid parking are smaller. That is, the availability of free parking decreases expected transit utilization even more for Business Passports than all transit pass types considered at once. With an $R^2$ of .031 the model explains more of the variance in transit utilization. For the Puget Pass case, the effects of all four variables are smaller. The $R^2$ of this model, at 0.008, is also significantly smaller than the $R^2$'s of the two other combined models. This may suggest that transit utilization by those using a Puget Pass may be influenced less by the employer's worksite, and more by factors external to the employer's location.

*4.2.3 Discussion.* These analyses indicate that subsidies, parking availability, and transit network structure all hold a significant influence on the number of trips an employee will take on transit from their worksite. The results on parking, in particular, come in direct contradiction to the results that Howard found in his work. Additionally, the differences in $R^2$ among the single independent variables models suggest that, while increasing the percentage that an employees transit pass is subsidized will increase the number of trips that they take from their worksite, the centrality of their worksite and the availability of parking—both free and paid—have a much larger effect on transit utilization.

Finally, these analyses show that these effects vary across the two main categories of ORCA products: Business Passports and Puget Passes. While these effects do hold for Puget Passes, the combined model used explains far less of the variation in transit utilization as compared to how much it explains for employees with Business Passports.

## 5 FUTURE DIRECTIONS

We believe that this work provides a number of avenues for further exploration. Given the number of assumptions that are made in the manipulation of the data in processing, future work would ideally perform some form of sensitivity analysis on both the apportioning of stops to worksites and the adjusting of employee counts. The construction of the properties assigned to each non-transit rider should also be more rigorously investigated.

More variables, especially built environment variables, should be used as predictors in the models to increase their explanatory and predictive power. This would also allow for the differences between the two ORCA products to be more thoroughly understood. It would be interesting to measure the variation in stop closeness centrality, given variation in on-time performance, to see how that combines with current static measure of centrality to predict transit utilization in general. This would build in reliability of transit services into the model.

Ideally, future work would be able to expand this work to defining transit utilization as all trips, not just trips from the workplace. This would allow for the inclusion of built environment variables and demographic variables associated with home locations that simply aren't relevant in a worksite context. However, this future work would have to identify other ways to include non-transit riders who may have similar circumstances but do not show up in the ORCA transactions data.

Additionally, since the models presented in this paper were trained solely on data from King County, it would be useful to expand the scope of analysis to include all agencies available in the ORCA data. Further out, it would be of great interest to generalize this work to additional cities and regions with similar wide spread subsidy programs.

## 6 CONCLUSIONS

In this work, we present an analysis of the effects of Commute Trip Reduction (CTR) employer provided transit pass subsidies on transit utilization from worksites while controlling for the structure of the transit network at those worksites and the availability of parking. This work is done through a large data set of ORCA transactions data, Automatic Vehicle Location data, and CTR survey data. The models built, while not predictive, show that each of the variables studied have significant impacts on transit utilization as defined.

Specifically, we find that subsidies have a significant positive effect on transit utilization, but that this effect is much smaller that the significant positive effect produced by transit centrality and the significant negative effects produced by the existence of both free and paid parking provided by the employer. Thus while the subsidies provided to employees do provide some encouragement to take transit from their worksite, the amount of transit service provided near the worksite and the providing of parking provide a more consistent encouragement or discouragement respectively.

Additionally, we find that the magnitude of these effects depend on the ORCA product that the employee is provided with. Business Passports see significantly higher rates of usage. It is unclear, however, as to how holding a Business Passport as opposed to a Puget Pass impacts an employee's perception of the structure of the transit network and parking availability.

Overall, this study provides support to the idea that subsidies can have a positive effect on transit utilization, even while controlling for transit service and parking availability. Our next steps for this work involve exploring the degree to which this holds in the presence of additional variables in the models and more rigorous assumptions in the data manipulation stage.

## A APPENDIX

Tables 8 and 9 show the ten most central and least central stops in the system respectively. Figures 2, 3, 4, 5, and 6 depict the centrality scores of all stops in the system. The coloring is consintent across all maps.

## ACKNOWLEDGMENTS

I would like to thank my friends and CEP cohort for providing me support throughout the process of this Senior Project and college more generally. I would especially like to thank my family for letting me attend Transition School and supporting me thoughout the last six years.

Lastly, thanks to my lab TRAC-UW for their camaraderie and to my mentor, Mark Hallenbeck, for giving me the opportunity to do this research in the first place and providing a generous amount of support the entire time.

## REFERENCES

[1] D. A. Badoe and M. K. Yendeti. 2007. Impact of Transit-Pass Ownership on Daily Number of Trips Made by Urban Public Transit. *Journal of Urban Planning and Development* 133, 4 (dec 2007), 242–249. https://doi.org/10.1061/(ASCE)0733-9488(2007)133:4(242)

[2] Robert Cervero and Kara Kockelman. 1997. Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research Part D: Transport and Environment* 2, 3 (1997), 199–219. https://doi.org/10.1016/S1361-9209(97)00009-6

[3] Sybil Derrible. 2012. Network centrality of metro systems. *PLoS ONE* 7, 7 (jul 2012), e40575. https://doi.org/10.1371/journal.pone.0040575

[4] Eric J Howard. 2016. CTR Programs, Subsidies, Built Environment, and Transit Level- of-Service, an analysis of factors that influence the rate of transit utilization using electronic fare card transaction records. (2016).

[5] David Lewis. 2008. America ' s Traffic Congestion Problem : Toward a Framework for Nationwide Reform. (2008), 10. http://www.brookings.edu/{~}/media/Research/F

[6] Praveen K Maghelal. 2011. Walking to Transit: Influence of Built Environment at Varying Distances. *ITE Journal* February (2011), 38–43.

[7] Daniel McFadden. 1974. The measurement of urban travel demand. *Journal of Public Economics* 3, 4 (1974), 303–328. https://doi.org/10.1016/0047-2727(74)90003-6

[8] Michael D Meyer. 1999. Demand management as an element of transportation policy: Using carrots and sticks to influence travel behavior. *Transportation Research Part A: Policy and Practice* 33, 7-8 (1999), 575–599. https://doi.org/10.1016/S0965-8564(99)00008-7

[9] Sabyasachee Mishra, Timothy F. Welch, and Manoj K. Jha. 2012. Performance indicators for public transit connectivity in multi-modal transportation networks. *Transportation Research Part A: Policy and Practice* 46, 7 (2012), 1066–1085. https://doi.org/10.1016/j.tra.2012.04.006

[10] Anne Vernez Moudon and Orion Stewart. 2013. *Tools for Estimating VMT Reductions from Built Environment Changes*. Technical Report. Washington State Transportation Center (TRAC), Seattle.

[11] Michaele Williams, Kathleen L Petrait, and M E Williams. 1993. Transportation System Management, Parking, and Travel Demand Management. *TRANSPORTATION RESEARCH RECORD* (1993). http://onlinepubs.trb.org/Onlinepubs/trr/1993/1404/1404-011.pdf

**Table 8: Most Central Stops During Afternoon Peak**

| Rank | Stop ID | Stop Name | Closeness Centrality |
|---|---|---|---|
| 1 | 450 | 3rd Ave & Union St: SE Bound | 0.00031449 |
| 2 | 565 | University St Station: Bay A | 0.00031446 |
| 3 | 456 | University St Station: Bay D | 0.00031403 |
| 4 | 4211094410 | University St Station: Link | 0.00031321 |
| 5 | 570 | 3rd Ave & Union St: NW Bound | 0.00031206 |
| 6 | 336 | University St Station | 0.00031200 |
| 7 | 455 | University St Station | 0.00031164 |
| 8 | 314 | 3rd Ave & Union St | 0.00031125 |
| 9 | 1121 | Westlake Station: Bay A | 0.00031044 |
| 10 | 1110 | Pine St & 5th Ave: SW Bound | 0.00031009 |

**Table 9: Least Central Stops During Afternoon Peak**

| Rank | Stop ID | Stop Name | Closeness Centrality |
|---|---|---|---|
| 1 | 58139 | S 194th Way & 58th Pl S | 0.00000086 |
| 2 | 58137 | 58th Pl S & Russell Rd | 0.00000167 |
| 3 | 8413 | South King County Activity Center | 0.00000171 |
| 4 | 70597 | NE 100th St & 116th Ave NE | 0.00005841 |
| 5 | 7309 | 116th Ave NE & NE 100th St | 0.00005841 |
| 6 | 99474 | Monroe Ave & Railroad St | 0.00007175 |
| 7 | 59041 | Cole St & Stevenson Ave | 0.00007290 |
| 8 | 59033 | Griffin Ave & Roosevelt Ave E | 0.00007311 |
| 9 | 59031 | Griffin Ave & 1st St | 0.00007358 |
| 10 | 59032 | Wells St & Griffin Ave | 0.00007375 |

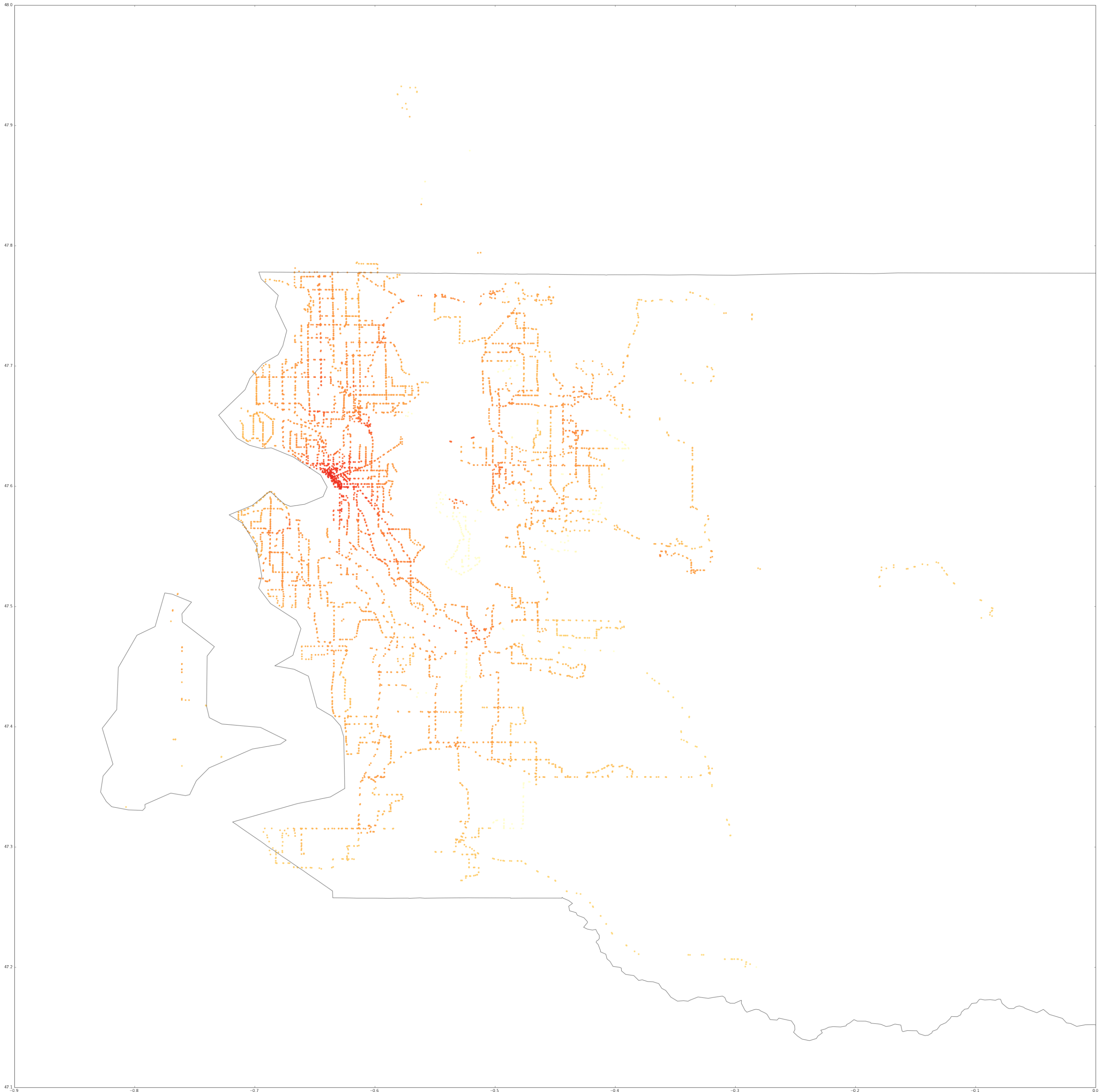

**Figure 2: Map of stops in King County colored by their centrality score during the early morning period. Yellow stops are less central while red stops are more central.**

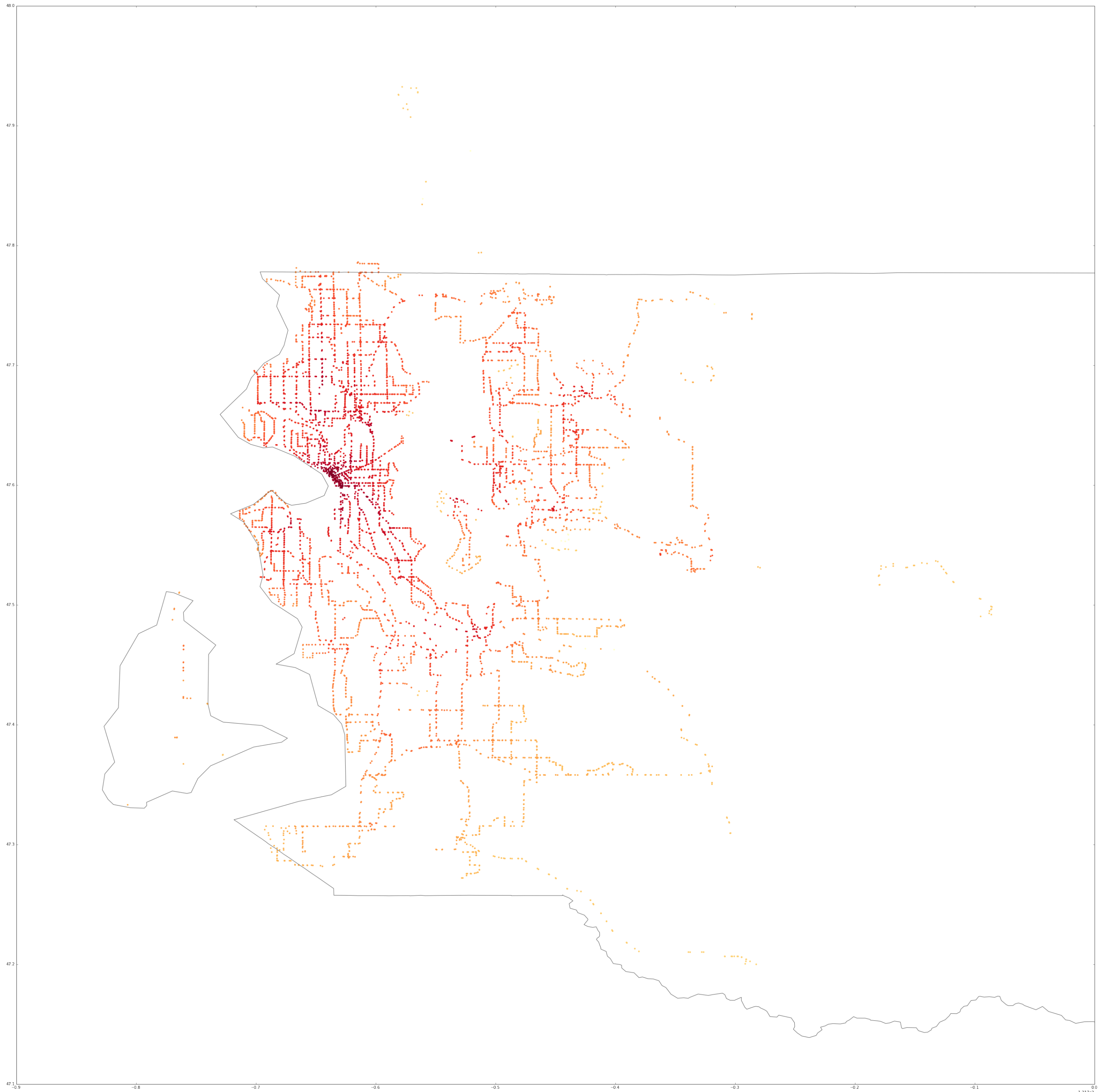

**Figure 3: Map of stops in King County colored by their centrality score during the morning peak period. Yellow stops are less central while red stops are more central.**

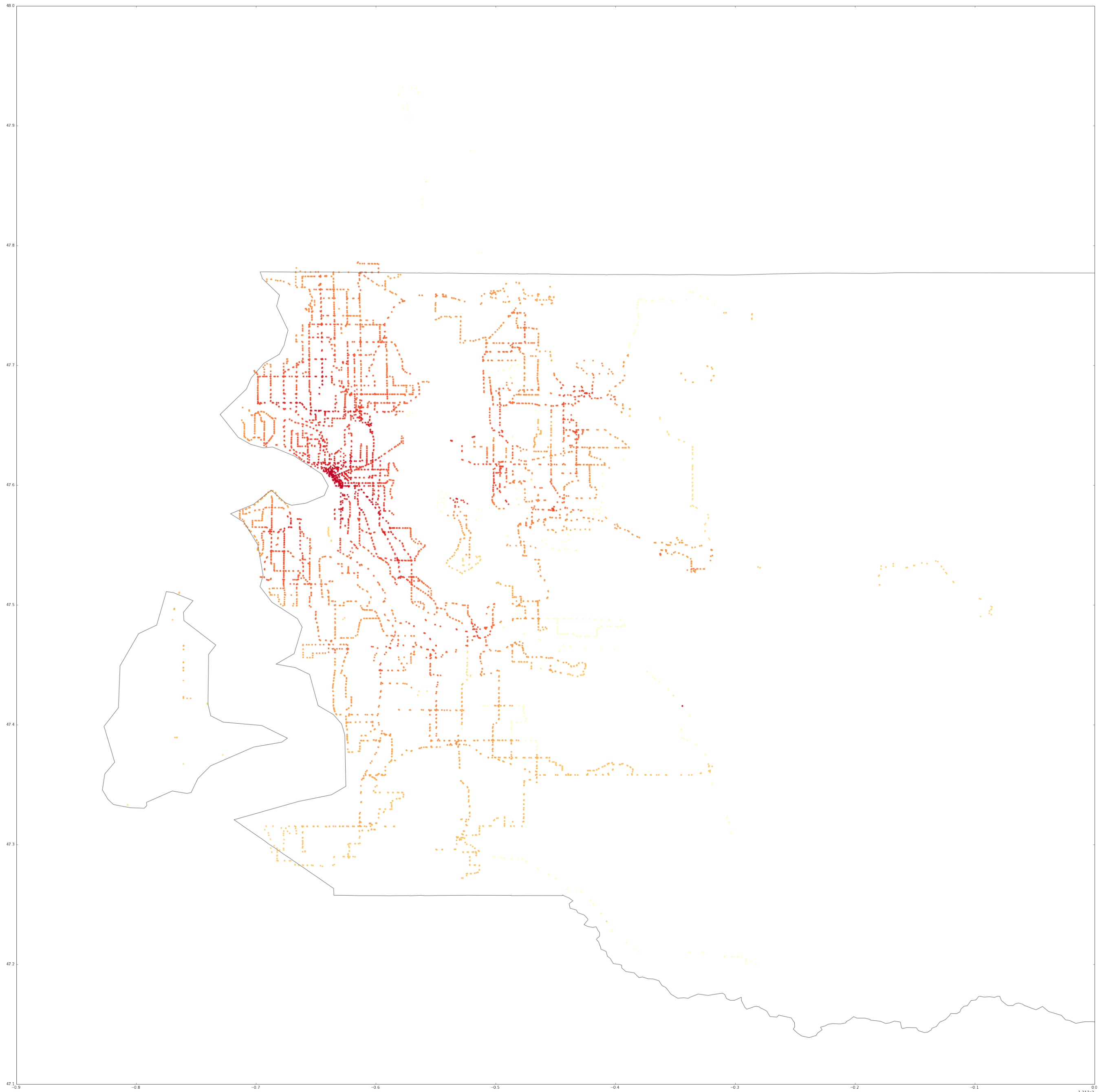


**Figure 4: Map of stops in King County colored by their centrality score during the midday period. Yellow stops are less central while red stops are more central.**

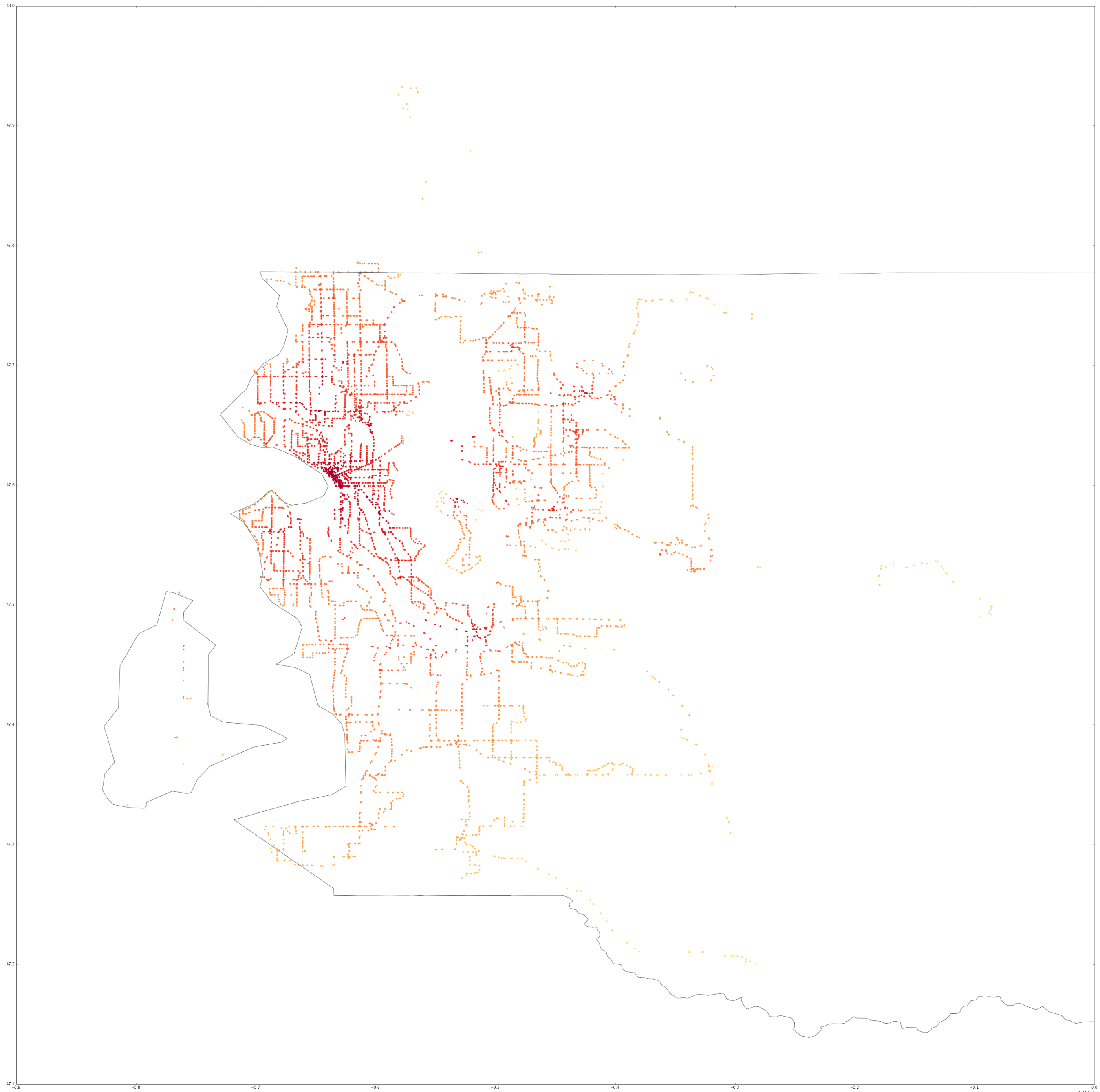

**Figure 5: Map of stops in King County colored by their centrality score during the afternoon peak period. Yellow stops are less central while red stops are more central.**

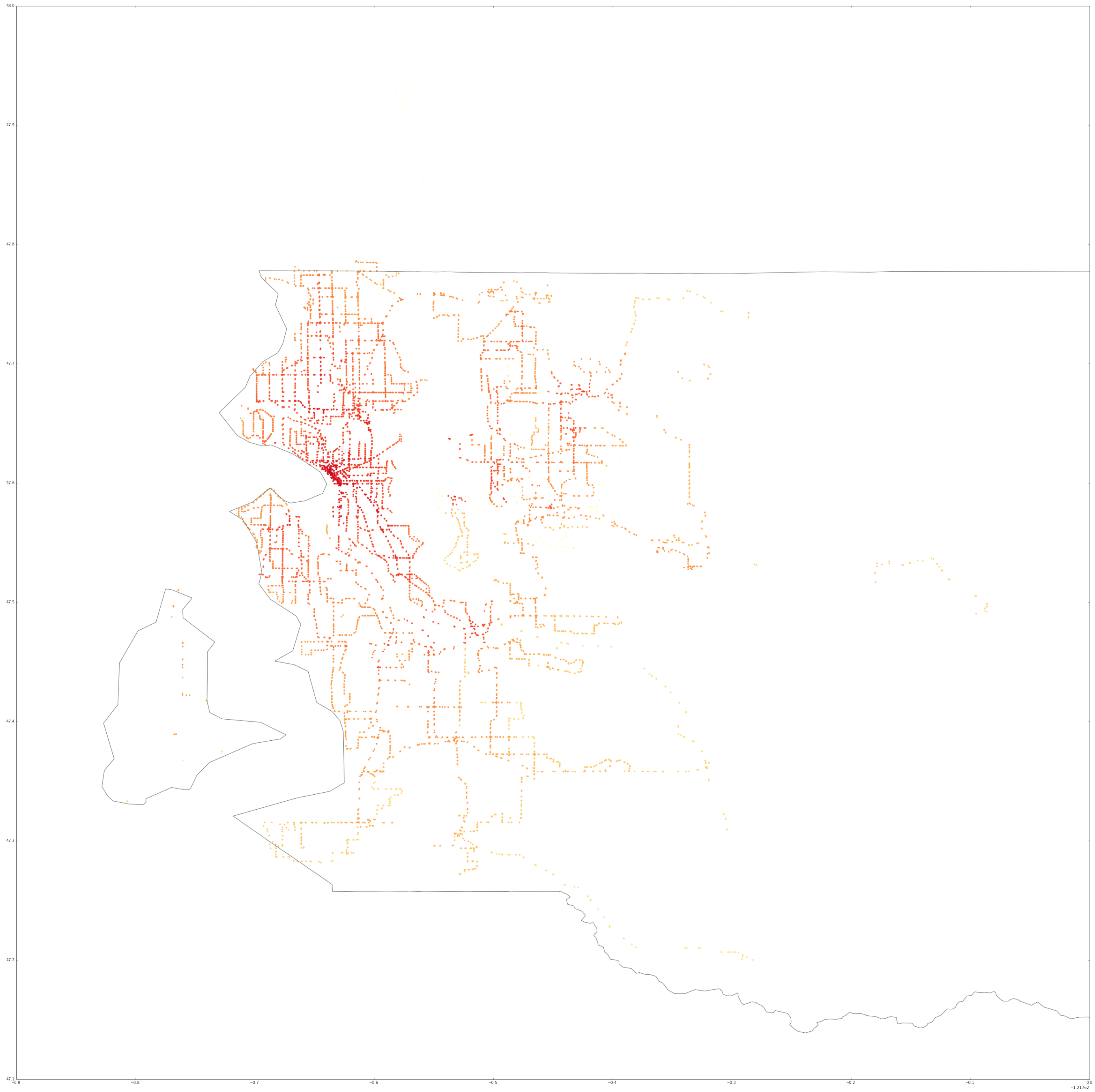

**Figure 6: Map of stops in King County colored by their centrality score during the evening period. Yellow stops are less central while red stops are more central.**