TRR
JOURNAL OF THE TRANSPORTATION RESEARCH BOARD

Research Article

Transportation Research Record
1–16

DOI: 10.1177/03611981221119190
journals.sagepub.com/home/trr
SAGE

Estimating Intermodal Transfer Barriers to Light Rail using Smartcard Data in Seattle, WA

James Eager[1], Chang-Hee Christine Bae[1], and Edward D. McCormack[2]

Abstract

Transit transfers are a necessary inconvenience to riders. They support strong hierarchical networks by connecting various local, regional, and express lines through a variety of modes. This is true in Seattle, where many lines were redrawn to feed into the Link Light Rail network. Previous transfer studies, using surveys, found that perceived safety, distance, and personal health were significant predictors of transfers. This study aims to use smartcard data and generalized linear modeling to estimate which elements of transfers are commonly overcome—and which are not—among riders boarding the Link Light Rail in Seattle and its suburbs. The aims of this research are twofold: (1) critical analysis of attributes of transfer barriers so that the future station area could serve improved riders' accessibility; (2) equity of transfer barriers among the users by analyzing the user breakdown of the origin lines and the destination. We use Seattle's One Regional Card for All smartcard data among the Link Light Rail riders in the Seattle metropolitan area in 2019, and applied a negative binomial generalized linear model. The model suggests that walking distance and walking grade have significant effects on transfers. For the users' equity analysis, the disabled population tends to transfer less, while the low-income and youth riders populations tend to transfer more often. Future research could incorporate a more mixed-methods approach to confirm some of these findings or include station amenities, such as live schedule updates for common transfer lines.

Keywords

data and data science, general, public transportation, transfers, smart card data

Effective public transportation systems should provide a safe, easy, and dependable travel experience regardless of the mode of travel. If a trip involves multiple modes of transit—for example, a bus to light rail connection—transit planning should minimize the transfer barriers where possible. Doing so supports the network's utility by offering clear and accessible paths that reduce the disruption of transfers as much as possible. Transfers, in any case, are a necessary aspect of any public transit network. The necessity of transfers is especially true in cities with various multimodal options and a clear hierarchy to the network (e.g., regional, rapid, and local transit options). Since 2003, planners in the Seattle Metropolitan Area have launched ambitious multimodal transit system plans, including the Link Light Rail (LLR) system and subsequent expansions. The current LLR stretches nearly 22 mi, connecting Seattle and its immediate surroundings areas. Some 27.6% of daily riders transfer from other modes of transit. Despite the importance of transfers and the necessity of their existence for the success of transit services, they can be an overlooked facet of transit planning's broader goals. A group of transit operators surveyed by the U.S. Federal Transit Administration considered "reducing rider difficulty" the second-highest priority in their transfer system objectives (*1*). However, only three of the 31 respondents offered concrete goals or objectives related to passenger convenience, revenue generation, or other factors. The rest provided transfer rules—such as time between rides

[1]Department of Urban Design and Planning, University of Washington, Seattle, WA
[2]Department of Civil and Environmental Engineering, University of Washington, Seattle, WA

Corresponding Author:
James Eager, jmc.eager@gmail.com

and cost of transfers—but lacked coordinated transfer policies to improve the system's user experience.

Transfer Experience and Penalties

The user experience of making connections within systems and between them is crucial in supporting positive public perceptions of transit networks and their success (*2*, *3*). Previous studies have found that the disruption to travel caused by a transfer plays a key role in whether people take transit at all (*4*). As a result, transfers can be a major barrier to the user experience of transit in both perception and relative cost (*5*). While studies have begun to look more closely at transfer penalties in travel demand modeling (*6–8*), explaining major barriers to transfers based on observed data is a limited field of literature.

Transfer penalties applied in the cost functions frequently used for transit optimization and modeling often appear as a standalone variable. However, whether this penalty includes both measurable costs (e.g., waiting and walking times) and perceptual costs (e.g., feelings of discomfort, safety) or exists as a singular penalty is debated in the literature (*4*, *9*). These costs can differ among groups, which also potentially causes various equity implications not thoroughly covered by existing research (*2*, *8*). Furthermore, the existing research tends to rely on survey and focus group analyses instead of observed transfers. This persists despite questions about the accuracy of stated preferences (*5*) and the ability to collect numerous travel diaries (*10*).

Smartcards, such as Sound Transit's (ST's) ORCA (One Regional Card for All), are contactless cards that store value for transit services in the Puget Sound region (*11*). Smartcard data offers a greater quantity of observations, improves replicability, offers a greater range of temporal comparisons, and is easier to collect once the infrastructure is in place (*12*). Smartcards also improve the general transit experience by reducing uncertainties about cost and intermodal operability that concern some would-be transfer passengers (*4*, *5*, *13*). The effect of these penalties is crucial to understanding how transfers might affect the mode choice decisions of travelers. The clear negative impact of transfers is a well-discussed point in the literature (*5*). The physical and environmental influences of transferring can include concerns about walking distances, elevation changes, or the speed of cars on the adjacent road. The perceptual and mental influences covered by focus groups and surveys include safety, anxiety, and mental disruption (*2*). This study focuses on attempting to measure the former, where observed transfers are used to determine how far people are willing to walk for a transfer to travel to their destination among smartcard users in the LLR network.

Seattle's Transit Network

The Seattle LLR system currently serves 16 stations across nearly 22 mi, running from Seattle's University District to Angle Lake in Seatac (*14*) (Figure 1). The Link is managed and operated by ST, which also operates the express bus services and Sounder commuter rail in the region. King County Metro (KCM) operates local bus service and bus rapid transit (BRT) in most of the Seattle Metropolitan Area, while Community Transit (CT), of Snohomish County, WA, serves some of the northern suburbs. In 2019, Link trains were boarded 79,674 times per week, on average. Of these boardings, about 21,990 are riders transferring from a different mode. Given the current state of the Link, with 13 of 16 stops in Seattle and a predominantly north–south orientation, multimodal transfers are a common and necessary part of the service. As the Link expands to 37 new stations, the need for convenient transfers will likely increase. This is not necessarily unique to the Link, as riders in cities across the U.S.A. predominantly transfer from bus to rail (*1*).

The planning and design process of the LLR system and its expansion takes transferring into account. In many cases, local transit authorities reroute existing bus lines to serve as more effective "feeder lines" for the LLR. In others, these agencies choose to forgo redundant bus lines altogether (*16*). With further expansion on the horizon, ST and other local and regional planning agencies continue to follow this process (*14*). As bus routes—local and express—rearrange to serve the LLR, not all potential transfer stops land near a Link station entrance. These non-adjacent feeder stations create possible walking transfers, where reaching the LLR system involves longer walks that sometimes stretch beyond ST's definition of a "station area" (0.5-mi radius).

Research Goals

The primary purpose of this paper is to use observed smartcard transfers to estimate the effects of physical and built environment characteristics of observed transfers to Seattle's LLR network using ORCA smartcard data. The first goal of this work is to better understand which barriers could be overcome for transfers, and which might not be. Barriers might include physical environment concerns, such as Seattle's hilly topography, or built environment concerns, such as the connectivity of the walking network in half-mile station areas.

This can support the broader literature about transfer penalties by offering an analysis of what features may create barriers to transferring riders via objective data. Further, this can guide station planning efforts to support more convenient high-information transfers where possible. In some cases, this might include adjusting routes to

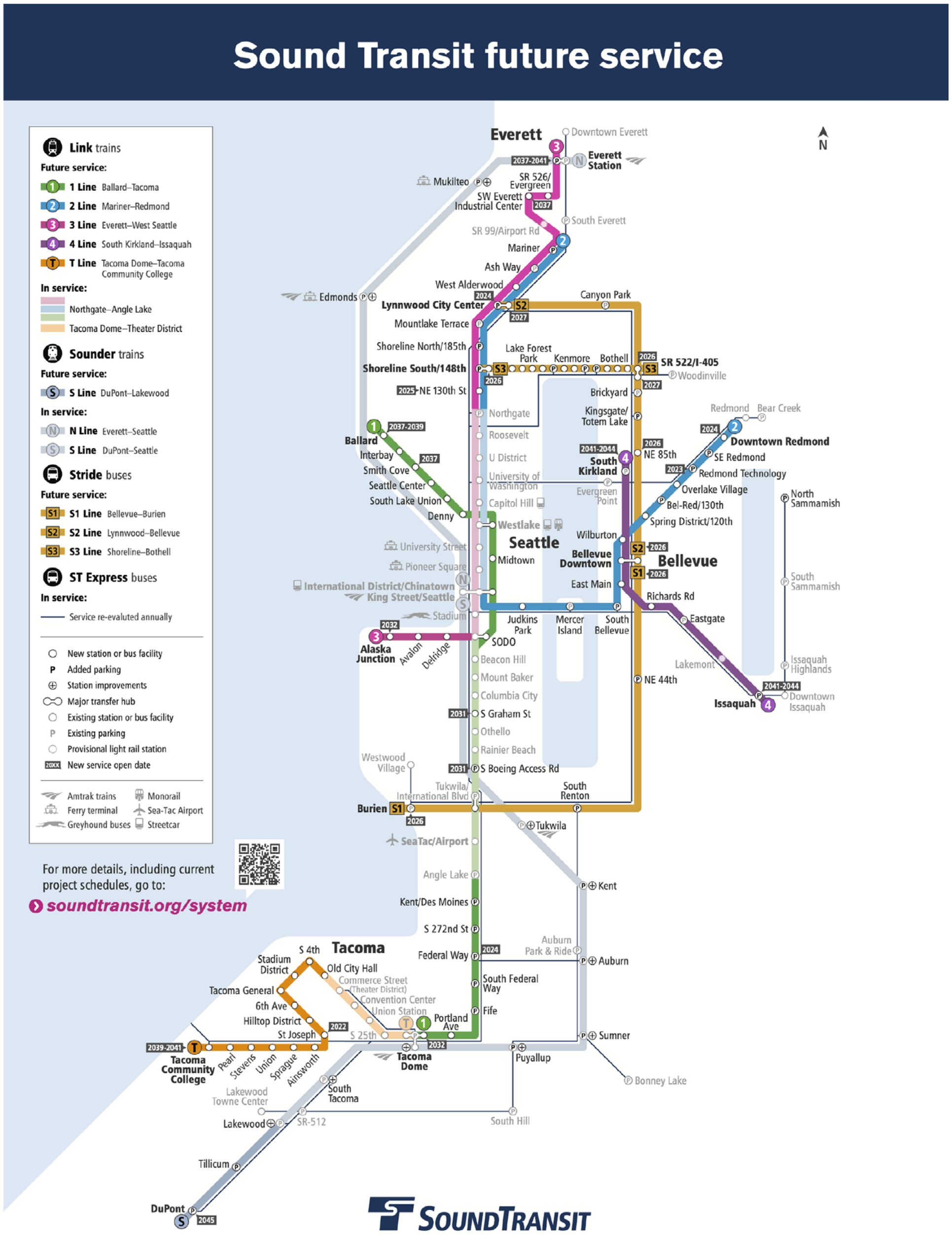

Figure 1. Current and planned Link, bus rapid transit, and Sounder networks (*15*).

provide clearer and more accessible routes between stations. In others, it may involve adjustments of road design and speed limits, offering improved perceptions of safety for transferring riders.

The second goal is to infer the interactions between transfer barriers and various equity populations. By analyzing the user breakdown of the origin lines and destination stations, this study identified where certain transfer pairs may need reexamination and adjustments to ensure that all populations have equitable access to high-quality transit (*17–19*). Among users of different ages, abilities, and backgrounds, the needs for certain elements in the transit network can differ (*2*, *5*, *19*), and taking a vertical equity approach to this analysis can help address the needs of transit-captive populations (*17*). While the data in this study does not include granular demographic data, studies have found that likely transit-captive populations often rely on transfers and interact with them in distinct ways compared to others (*5*). Using the ORCA data's specific user types to parse this information out can offer some exploratory findings on different age and income groups' transfer habits, as a proxy for more granular demographic data.

Methods

Three major sources of data underpin this research: Open Street Map (OSM), KCM's Google Transit Feed Specification (GTFS), and ST's ORCA transactions. R and Python programming languages support the access, analysis, and presentation of this data. Data collection is also supported by Google and OSM application programming interfaces (APIs).

Data

ORCA. ST supplied the ORCA smartcard data used in this study. The Washington State Transportation Center (TRAC) has organized the data into a matrix of origin bus lines to destination light rail stations. In this format, each LLR station has a column for each bus line it received transfers from, and the accompanying count of those transfers during each study period. TRAC has also organized the data to offer individual LLR station-level use for each study period. This station-level data includes breakdowns by type of user (Table 1) and type of card (such as those storing dollar values or those supplied by employers). This study only analyzes the user type as this elucidates how different groups of people might experience transfers differently. Data for this breakdown only considers the transit agencies that the transfers engage with: KCM, ST, and CT. A full table of definitions and user composition can be found in Table 1.

TRAC's ORCA data covers two time periods in 2019: 7 January to 23 March (winter) and 1 July to 31 August (summer). These time periods are slightly unequal, so transfers and ridership numbers were converted to weekly transfer and ridership averages. A "transfer" defined by the ORCA card system is any boarding that occurs within 2 h of a rider's initial boarding. As a result, there are a considerable number of transfers that occurred within the 2-h time-frame that were not direct connections between modes of public transit, which are referred to in this study as "financial transfers." There are also transfers from LLR "to" LLR, which fall under this same categorization. These could be continuations of trips with an interruption, or a round-trip ride where fewer than 2 h were spent at the destination. As a result, only lines with at least one weekly transfer and a station within a 1.5-mi walk were considered "reasonable transfers."

Google Transit Feed Specification. GTFS offers a common standard for transit agencies to publish spatial and scheduling data of their networks (*20*). Agencies offer static GTFS feeds made up of timetables, routes, and stops organized in several text files of relational data (Figure 2 outlines each piece of the GTFS data structure and how they connect) and real-time GTFS feeds, which are constantly updating with information about arrival and departure times. The static feed is offered as free and public information by KCM and offers the stations, timetables, and routes of both KCM and ST. This data offers the spatial framework for identifying nearest stops on each line and the line characteristics.

Open Street Map. OSM is an open-source platform for user-generated maps (*22*). These maps include detailed data on street, transit, and walking networks around the world, supported and updated by an active community of users. For this project, OSM data is accessed through the OSMnx package in Python, created by Geoff Boeing (*23*). OSMnx couples the complex network analysis of the NetworkX package with the data and routing from OSM's API. As a result, the package offers the ability to easily calculate network and routing metrics along with topological measures using Python. Unfortunately, OSM does not provide native topographic data; however, OSMnx can access the Google Elevation API to add elevation data to the underlying transportation networks.

Table 1. One Regional Card for All User Definitions

User	Qualification	Share (%)
Adult	None	81.6
Senior	At least 65 years of age	3.7
Youth	At least 6, but no more than 18 years old	5.0
LIFT	Household income less than 200% of federal poverty level	5.4
Disabled	Disability or Medicare recipient	4.3

Note: Share is ridership as a percentage of all trips recorded in the 2019 TRAC data. LIFT = Low Income Fare program.
Source: Definitions are adopted from Sound Transit (*14*).

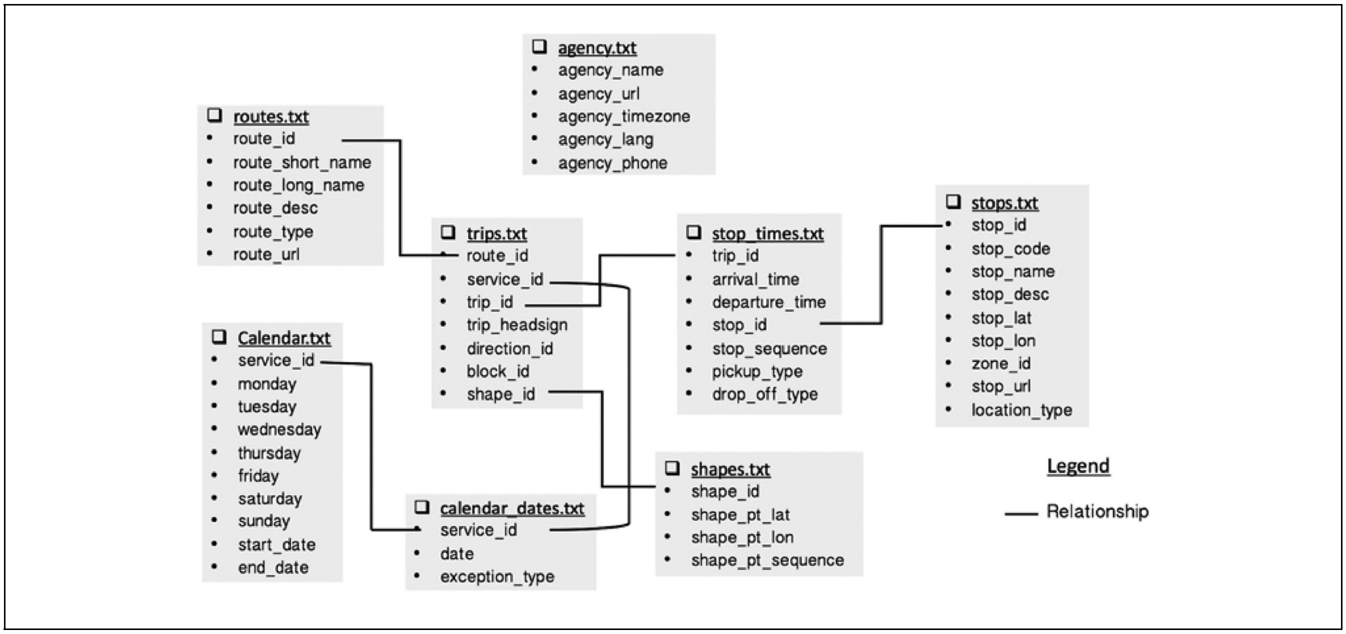

Figure 2. Static Google Transit Feed Specification feed file connections and organization (*21*).

Some of the station area statistics from OSMnx are nontraditional, such as the entropy and circuity of a street network. Per Boeing (*23*), the entropy of a street network is a method for evaluating how "grid-like" or ordered a given street network is. Entropy calculations from OSMnx are normalized on a scale of 0–1 (*24*), where 1 is a perfect grid-like network and 0 indicates streets open to all directions. According to Boeing's evaluations, Chicago, Miami, and Minneapolis are the most ordered cities (0.9), while Charlotte is the most disordered city (0.002) in the U.S.A. (*24*). Circuity is a measure of how circuitous, or rounded, the network is. In theory, a more ordered and straight network supports a more coherent and successful transfer walking route, as the network is easier to follow. Using a combination of nontraditional and traditional statistics (such as intersection density) could compensate for any limitations in any individual metric and offer a more complete picture of station area street networks.

Linking Sources. Using Python, the process of connecting these databases involves applying a few different packages, such as pandas, NumPy, and geopandas, for the manipulation and processing of both spatial and relational data. Using these packages, the ORCA transfer data was filtered to reasonable transfers and the nearest station for each origin line identified for routing. The route with the shortest walking distance was taken and metrics covering the distance, elevation change, and vehicular speed along the route were collected. For each Link station, OSMnx calculates and collects measurements of the order (or gridedness), circuity, and density of the half-mile station area. Finally, TRAC station-level ridership data was aggregated for each complete origin line and each individual Link station to produce proportions of each ridership type (adult, youth, senior, disabled, low-income) in these contexts. The list of these variables and their groups can be found in the data dictionary (Table 2), while the process is detailed by the flowchart in Figure 3. An example of what the final routing looks like is shown in Figure 4. In this graphic, nearby Link stations are pointed out to contextualize the density of stations in Downtown Seattle, and their proximity to some origin stations suggests not all riders take the shortest routes to their destinations. We will discuss the detailed process in the *Modeling* section.

The main predictors tested in this study are those that make up the transfer walking route. These variables, such as the mean and max grade of the route, offer the clearest parallel to the transfer preference surveys that identified what riders' stated transfer preferences are (*2*, *5*). The station area characteristics control for how connected, dense, and ordered the street network is in the half-mile surrounding each light rail station. These characteristics control for the possibility of other walking routes, as density and grid-like street networks offer more coherent walkable paths than those that are less grid-like (*24*). Finally, the ridership metrics of both the origin line and the destination Link station offer a lens for how willing to transfer different groups of riders are.

Table 2. Data Dictionary

Code	Description	Source
Transfer characteristics		
Weekly	Weekly transfers from origin line	TRAC
Origin	Most recent line ridden before the transfer to Link	TRAC
szn	Season of data collection period (winter or summer)	TRAC
Transfer walking route		
wkLen	Shortest possible walking distance	OSM
wkMnSpd	Mean vehicular speed limit along route	OSM
wkMxSpd	Maximum vehicular speed limit along route	OSM
wkGrade	Mean grade of shortest walking path	GEA
wkMxGrd	Max grade of shortest walking path	GEA
wkRise	Total elevation gain of shortest walking path	GEA
Link station area		
stEnt	Relative order (entropy) of transfer station area	OSM
stCirc	Total elevation gain of shortest walking path	OSM
stStLen	Sum of length of streets in station area	OSM
stStDen	Sum of street lengths divided by station area	OSM
stIntDen	Number of intersections divided by station area	OSM
Link station ridership		
lrLift	Percentage of LIFT (low-income) users at Link station	TRAC
lrSenior	Percentage of senior users at Link station	TRAC
lrYouth	Percentage of youth users at Link station	TRAC
lrDisable	Percentage of disabled users at Link station	TRAC
lrAdult	Percentage of adult users at stations on link route	TRAC
Origin line ridership		
ogLift	Percentage of LIFT users on origin line	TRAC
ogSenior	Percentage of senior users on origin line	TRAC
ogYouth	Percentage of youth users on origin line	TRAC
ogDisable	Percentage of disabled users on origin line	TRAC
ogAdult	Percentage of adult users on origin line	TRAC

Note: TRAC = Washington State Transportation Center; OSM = Open Street Map; GEA = Google elevation application programming interface; LIFT = Low Income Fare program.

Modeling

Model Selection. We conducted a series of negative binomial models to find the best fit predicting the number of transfers from each likely transfer station to each paired Link station. The negative binomial generalized linear model (GLM) should provide an improvement over other common models of count data, such as the Poisson and Gaussian (linear) GLMs. This is a result of the negative binomial not falling to the same assumptions made by these other types of modeling families. The negative binomial model allows for unbounded counts and correlation between events, and assumes an overdispersed outcome (*25*).

Model Building and Interpretation. To identify the preferred model—that is, the model that most effectively explains the data—several models were tested and run. This study took a stepwise approach to model building, beginning with the transfer route characteristic variables. These are the main predictor variables this study aims to use to explain transfer differences, so their inclusion is necessary. Gradually, variables from each following group (Table 2) were added to explain further variance. At least one variable from each of the link station area,

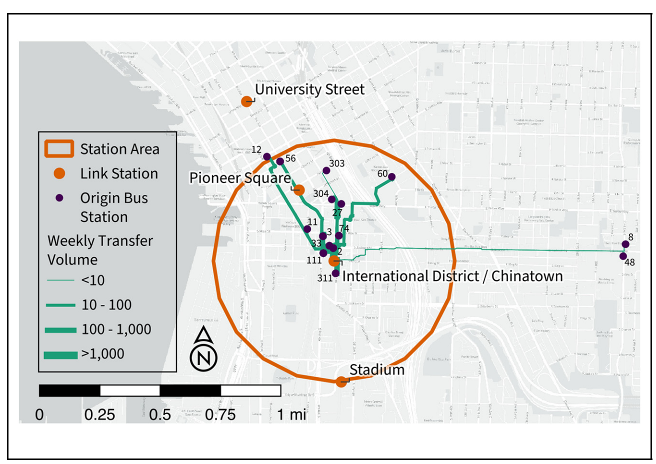

Figure 3. Flowchart of the data, modeling, and analysis methods.
Note: ORCA = One Regional Card for All; GTFS = Google Transit Feed Specification; API = application programming interface.

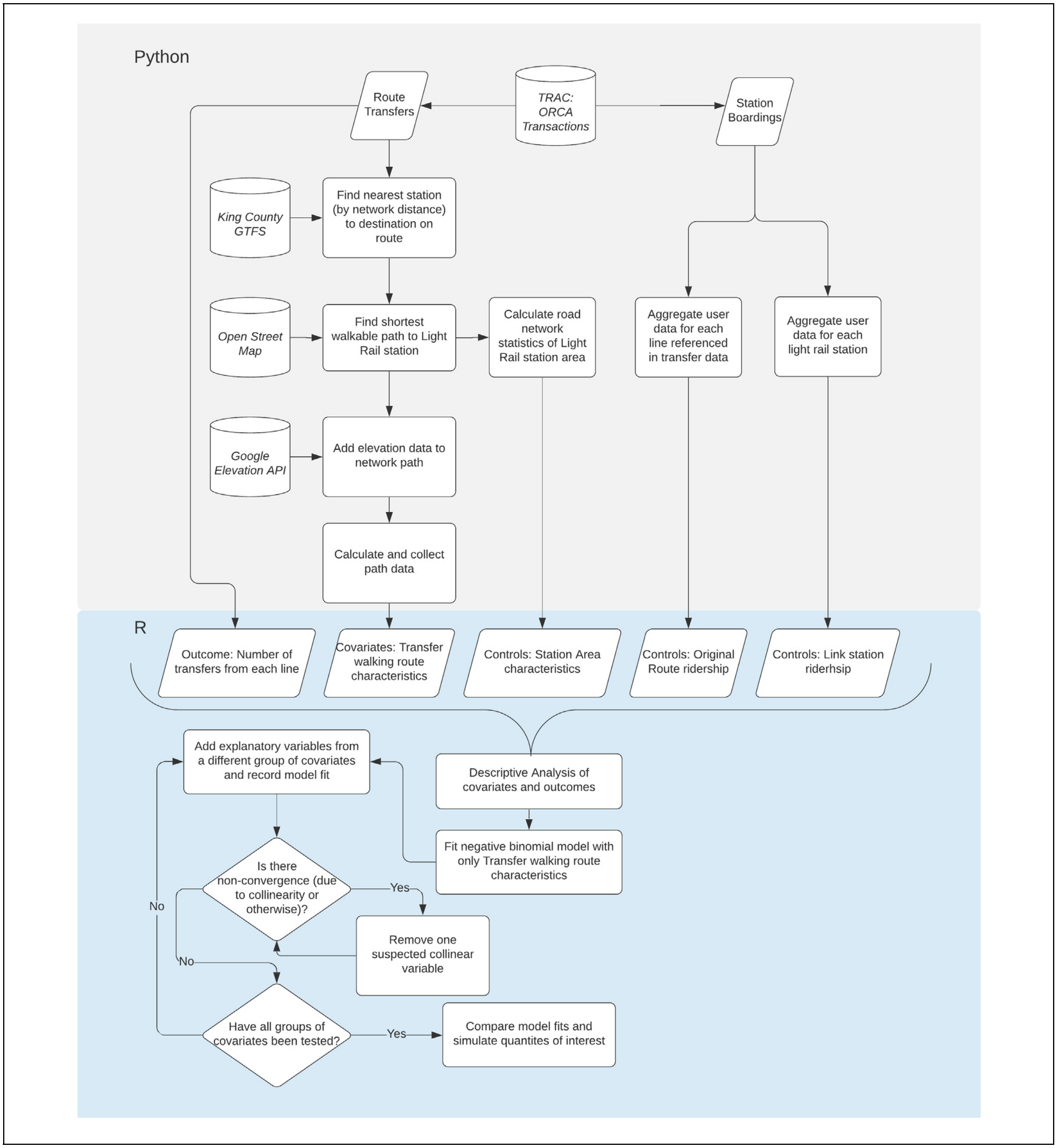

Figure 4. International district—Chinatown Station transfer walking routes.

link station ridership, and origin line ridership was present in the final model, as each of these groups had some degree of influence on the transfers between stations. Any variables that were collinear or prevented convergence of the model were removed. A flowchart describing this—along with the data collection process—can be found in Figure 3.

Each model was tested for the Akaike information criterion (AIC) and Bayesian information criterion (BIC). These metrics offer a method of comparing the fit and explanatory power of each model. The AIC differs slightly from the BIC, as it tends to pick models with more covariates, while the BIC penalizes additional covariates. It is possible that the AIC and the BIC will find the same model fits best. If they do not, it will offer an opportunity to explore why that might be.

Discussion about model parameters will be informed by quantities of interest and counterfactual scenarios rather than traditional frequentist discussions (*25*). The best-fitting model can then be used to simulate these scenarios, which offers a clearer picture of the model's implications and uncertainty (*25*). While coefficients and p-values offer some explanations for how one might interpret the model, these simulations offer more interpretable results. Given that the negative binomial results in log-link coefficients, reading these values off a table can be misguided at worst and imperfect at best (*25*). As such, simulation offers clearer interpretations and allows for greater detail in discussing the model's results.

Limitations

This study's approach to understanding transfers is limited in a couple of ways. Firstly, by focusing on using smartcard data, the collection of transfer information is limited to those who use smartcards and those who transfer to another mode of public transit. Those who do not use smartcards and tap in or off at LLR stations from other modes were not captured with the same detail. As a result, this study can only truly begin to address the questions of transfers to the LLR from the bus network in Seattle. While the bus network is the most used and most expansive mode in Seattle, there is literature that suggests people are willing to walk different lengths for different modes of transit. It would be interesting to understand if that relationship carries over to transfers.

This study does not account for people who simply do not take transit because of transfer limitations. Instead, it can serve as a proxy for which transfer characteristics are most conducive to transfers and support a clearer framework of understanding intermodal transit transfers going forward. Exploring how different barriers interact with the walking distance still informs the field. Furthermore, some of the more granular analyses using smartcard types or users is limited by the extent of the different cards' ability to get into different populations' hands. For example, it should not be assumed that 100% of those eligible for low-income fares (ORCA Low Income Fare program [LIFT]) are using the program, which also applies to the youth, senior, and disabled cards applied in this study.

On the subject of demographics, some may suggest inclusion of demographic data in the modeling process for this study. However, smartcards do not capture specific demographic information. Furthermore, there is no explicit data about smartcard user demographics in Seattle. The most common way for frequent riders to pay for Link trips is with a smartcard, although ticket machines also sell one-way and day passes, which are not captured here. Transfers from bus to LLR in Seattle require a smartcard unless the rider pays twice (for example, with cash on the bus and a day pass on the LLR). For these reasons, the limited discussion about demographics and equity is entirely based on users of the equity-focused ORCA card types. Despite these limitations, applying explicit demographic indicators could reveal more direct implications about transit-dependent populations, rather than the proxy indicators we derived through smartcard types.

There is also the possibility—especially in denser more walkable neighborhoods—that people take different paths to transfer with minor differences in distance. Particularly in station area transfers, it is unlikely that the speeds, slopes, and distances of a transfer walking path would differ significantly, but it is possible that certain routes may be avoided because of reasons not captured by this study. The station area covariates are expected to capture some of this potential variance by accounting for the ease by which a rider might take a different walking route.

A difficulty with data collection also relates to the nature of individuals' smartcard use. The model built assumes that the majority of riders diligently tap the card readers for boarding, alighting, and transfer. However, based on the ORCA data, it appears that riders may not tap when transferring from the bus to the Link. The issue appeared more frequent at terminus stations like Angle Lake, where riders who likely transferred from the bus to a different station did not tap until disembarking. As a result, these transfers were categorized as Angle Lake transfers, despite their lack of feasibility. Figure 5 shows all the bus lines that registered a transfer to the Angle Lake Link station. The approved class of bus lines was determined to be feasible, while the discarded class was not.

Results

Modeling Results

Model Building. The first step in the modeling process is checking the distributional characteristics of the data. The histogram of the transfers (Figure 6) shows that the outcome is overdispersed or strongly skewed. A more rigorous way to test for overdispersion is by implementing a test that checks the Poisson assumptions outlined earlier (see the *Model Selection* section), specifically comparing the null hypothesis of $\mathrm{var}(y) = \lambda$ to an alternative of $\mathrm{var}(y) = \lambda + c * f(\lambda)$, where the function $f(.)$ is a

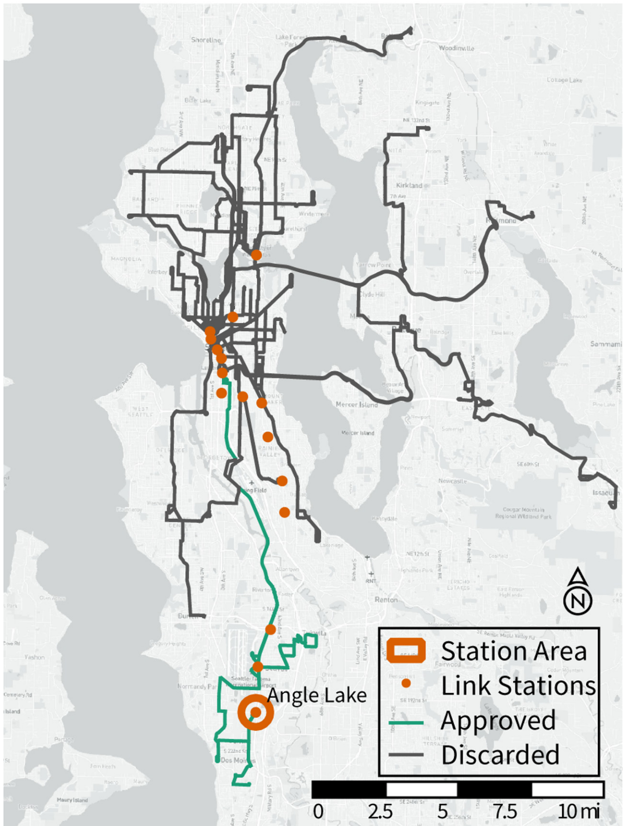

Figure 5. Angle Lake transfer validation result.

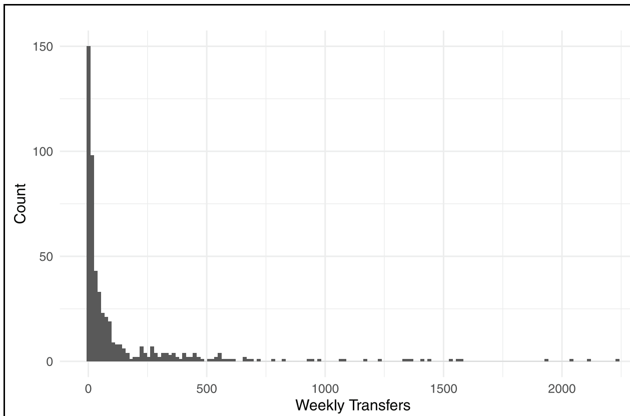

Figure 6. Histogram of weekly transfer volumes to all Link stations.

linear monotone function (*26*). The resulting coefficient c is a measure of the dispersion of the outcome, where $c>0$ implies overdispersion and $c<0$ implies underdispersion. The result of this test was a c value of 647 with a z -score of 5.13, clearly implying significant overdispersion.

We applied a form of stepwise model building. By adding variables in groups, the study could gauge how well each group of variables adds to the descriptive power of the model. However, during this process some models did not converge. The likely reason is that some covariates are moderately or strongly collinear. As a result, a correlation matrix of the variables collected was created and referenced during model building. During this process, non-convergence led to the removal of variables that were collinear and did not contribute significantly to the descriptive power of the model.

In the case of variables that described the origin line ridership and the LLR station ridership, each of the variables were collinear within these groups. In both groups, greater adult ridership was heavily linked to lower ridership of youth, senior, disabled, and low-income riders. Similarly, increases in each of these four equity rider groups (youth, senior, disabled, and low-income riders) was correlated to increases in the other three. In other words, as youth ridership increased, senior, disabled, and low-income riders also tended to increase. ORCA youth riders were the strongest predictors and preserved in the models, while the other groups were discarded.

Among the station area variables, similarities between the covariates led to strong collinearity between intersection density, street density, street length, and entropy. Entropy was the strongest predictor among these and was preserved in the models alongside circuity. These two variables present a clear picture of the station area's walking network characteristics, where high entropy and low circuity suggest a dense grid-like network with minimal curvature.

The results of each added group of covariates to the model are shown in Table 3. In the table, model goodness-of-fit, coefficient estimates and their 95% confidence intervals, and frequentist significance is arranged from the least complex model—with only the transfer walking route variables—to the most complex—with all variable groups represented. Given the different scales of the covariates, it is difficult to comment on the comparative strength of each of them, but what is clear is that many of the coefficient estimates line up with the theory behind this model.

For example, higher street network entropy (*stEnt*) —indicating a more grid-like network—is consistently a positive correlation (2.5–3.8; CIs 0.5–5.8). Similarly, *wkGrade*—measuring the mean grade of the walking route—is consistently negative through all models (−0.1 to −0.3; CIs −0.1 to −0.4). Both of these match the theory underlining this model, where the study expected to find walks along steeper streets would be less conducive to transfers and more grid-like street networks would support or indicate more transfers. Given that the 95%

Table 3. Regression Table: Covariate Estimates, 95% Confidence Intervals, and Significance from the Model Building Process

	Dependent variable:			
	Weekly			
	Walking route	+ Station area	+ Origin line	+ Link station
szn	0.455*** (0.159, 0.751)	1.042*** (0.745, 1.339)	0.877*** (0.587, 1.167)	1.075*** (0.787, 1.363)
log(*wkLen*)	−0.261*** (−0.378, −0.143)	−0.391*** (−0.512, −0.270)	−0.316*** (−0.433, −0.198)	−0.372*** (−0.502, −0.242)
wkMnSpd	0.011 (−0.052, 0.074)	−0.033 (−0.096, 0.029)	−0.051 (−0.112, 0.010)	−0.022 (−0.083, 0.040)
wkMxSpd	−0.014 (−0.043, 0.014)	−0.001 (−0.029, 0.027)	0.006 (−0.022, 0.033)	−0.002 (−0.030, 0.025)
wkGrade	−0.135*** (−0.204, −0.066)	−0.245*** (−0.314, −0.176)	−0.239*** (−0.307, −0.172)	−0.303*** (−0.381, −0.225)
wkMxGrd	0.007** (0.00002, 0.015)	0.022*** (0.015, 0.030)	0.020*** (0.013, 0.027)	−0.007 (−0.032, 0.017)
stEnt	na	2.497** (0.497, 4.498)	3.560*** (1.608, 5.512)	3.773*** (1.737, 5.808)
stCirc	na	18.287*** (12.026, 24.548)	13.406*** (7.165, 19.647)	19.158*** (12.882, 25.434)
ogYouth	na	na	0.106*** (0.072, 0.139)	0.124*** (0.087, 0.160)
ogSum	na	na	0.001*** (0.001, 0.002)	0.001*** (0.001, 0.002)
lrYouth	na	na	na	0.0005** (0.0001, 0.001)
lrSum	na	na	na	0.019*** (0.012, 0.027)
Grd:MxGrd	na	na	na	0.005*** (0.002, 0.009)
Constant	5.492*** (4.080, 6.904)	−14.544*** (−20.928, −8.160)	−11.358*** (−17.770, −4.946)	−18.602*** (−25.094, −12.110)
Observations	524	524	524	524
Akaike Information Criterion	5814.004	5783.977	5747.850	5710.561
Bayesian Information Criterion	5848.096	5826.592	5798.988	5774.483

$^{**}p < 0.05$; $^{***}p < 0.01$.
Note: na = not applicable.

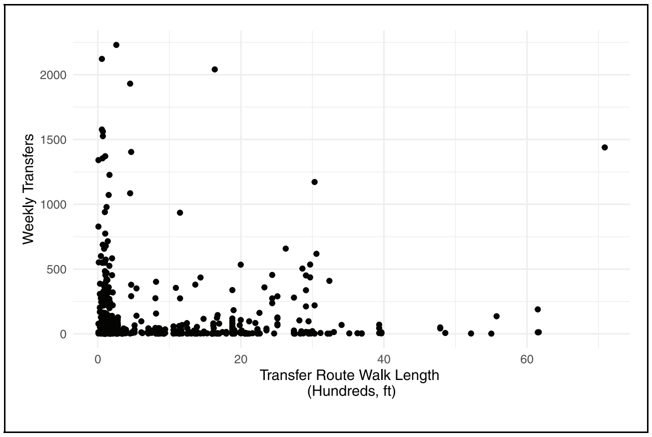

Figure 7. Walk length and weekly transfers plot.

confidence intervals for these variables also never overlap 0, we can be assured of the implications of these coefficients.

The stepwise process involved several additional changes. Firstly, adjusting the length of the transfer walk distance (*wkLen*) to a logarithmic relationship from a linear one. The plot of walk length and weekly transfer volumes (Figure 7) suggests that the relationship has a slight curve, with few outliers. This prompted the change to a logarithmic relationship between walking distance and transfers, which will be discussed further in the following sections.

Secondly, we decided to include an interaction term in the full model to better capture the effects of the walk grade when accounting for all other variables. This interaction term is between the mean grade (*wkGrade*) and maximum grade (*wkMxGrd*) variables. What the interaction term asserts is that the effect of having a higher mean walking grade is different when the max grade is higher or lower; a steep slope in a given route increases the effect that the mean grade has because the coefficient for this term is positive in the final model. Both the p -value (<0.01) and estimate ($0.005 \pm .003$) of this variable in Table 3 support this relationship between covariates.

Thirdly, a couple of changes for the sake of easier interpretation were also included in each of these models. These had no impact on the results, and only served to support a more coherent discussion about their effects. These involved changing the scales of a few variables. The first was adjusting *wkLen* to hundreds of feet instead of individual feet. The other two variables to change in scale were the sums: *ogSum* and *lrSum*. In the case of origin line ridership, it was adjusted to hundreds of weekly riders. Link weekly ridership was adjusted to thousands of weekly riders.

It is important to note that this model intends to explain current transfer patterns, while offering some thoughts on how changes to these variables could support station planning efforts. Some variables were left in the model despite appearing insignificant because of their p -values and confidence intervals. For example, the coefficients describing vehicular speeds along the walking route (*wkMsSpd* and *wkMxSpd*) are insignificant at the 95% confidence level but still contribute to the model goodness-of-fit. So, while they might not have a significant effect on whether riders transfer between certain station pairs, they do contribute to the explanatory power of the model; more importantly, the walking route variables answer questions about their impact on transfers. Specifically, it does not appear that the vehicular speeds along walking routes had a statistically significant impact on transfer volumes between stations in Seattle's LLR network.

Through the model building process one can also see where additional variables help explain variance that was picked up by other variables previously. For example, the maximum grade of the walk between stations (*wkMxGrd*) was highly significant until the variables for light rail ridership were included (Table 3). Statistical significance (or lack thereof) for a given parameter may not imply a better model for out-of-sample results (*25*). So, if this same model were applied to a larger sample of transfer data in the LLR network, the inclusion of these variables theoretically could support improved estimates of transfer volumes.

Model Interpretation. This research focuses on attributes of the main variables influencing the transfer walking route. These variables are the main glimpses as to which transfer barriers people are willing to overcome regularly or not. The clearest among these is the length of the walk between stations. The other variables tested describe the elevation changes and vehicular speeds along the walking route (see Tables 2 and 3); however, these variables and their effects are not as straightforward as walking distance.

Starting with the effect of walking distance, the logarithmic relationship identified in Figure 7 is further examined in Figure 8. In this graphic, each point represents the estimated change in weekly transfer volumes (and each line the 95% confidence interval) when a station is moved one block closer from three different starting points. On the y-axis is the model as each group of control variables is added, with the top group consisting of solely the walking route characteristics, and the bottom group the full model.

A street block is taken as 300 ft, per OSM, and the starting distances are the near-minimum (two blocks), near-mean (five blocks), and edge of the station area (10 blocks). The plot clearly shows that for stations already further from the LLR, changes to the walking distance

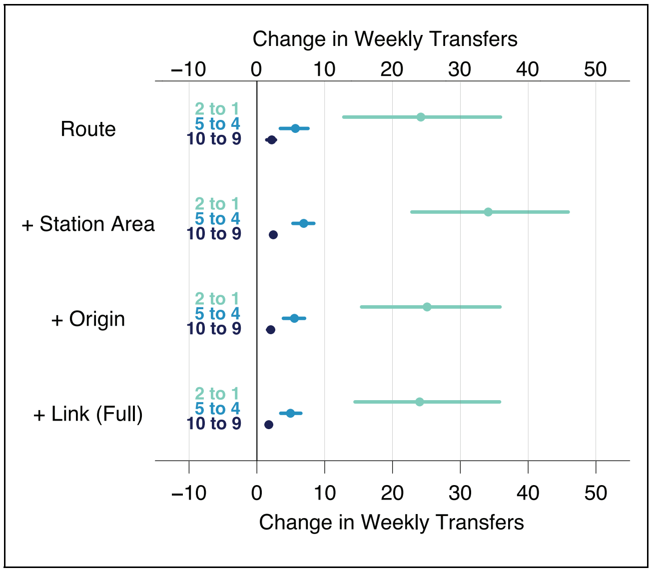

Figure 8. First differences in transfer ridership after moving a station one block closer than the given starting position for each model (with 95% CIs).

have diminished effects. Changes among nearer stations have significantly stronger effects, but can be considerably more varied. This plot also shows that regardless of controls included, the simulated differences based on distance are not significantly different from one another. In other words, regardless of which controls are in the model, logarithmic distance is a very strong explanatory variable for weekly transfers.

Looking more closely at distance along with grade and vehicular speed, Figure 9 shows the marginal effects for *wkLen*, *wkGrade*, and *wkMaxSpd* on weekly transfers for each of the three models. One "simple" model contains just the transfer walking route variables and log distance (see column *Walking route* in Table 3). The other two models are the final "full" model with log distance and one full model with linear distance. For reference and context, the mean weekly transfer ridership is 140.

First of all, the log transformation clearly improves the model significantly (see previous discussion about Figure 7). There is clearly not a linear relationship between distance and weekly transfer pairs. The confidence intervals of both log distance models are tighter, and based on the effects they are significantly greater than 0 (0.14–0.27), while the linear distance effect is effectively 0, since one could draw a straight flat line within its confidence interval ("Lin" model in Figure 9). Furthermore, the full model shows a slightly stronger drop off with walking distance compared to the simple model. This suggests that the other variables included helped explain the increased ridership implied by the simple model more effectively.

Next, the walking grade plot suggests strong agreement between all models. Effectively, all three models exhibit the same effect of walking grade on transfers. For reference, Seattle's mean walking network grade is 3.9% and the median is 2.3%. Among the data, walks tend to average around 3.3% with a median of 2.7%—not far off from the rest of the city. From the second plot in Figure 9, it is apparent that transfers drop off pretty quickly around a 3% or 4% grade before leveling off. Finally, the maximum vehicular speed plot of Figure 9 also suggests a lack of significant difference in this variable among the three models. However, only two of these—the log distance models—are effectively 0. The linear distance model shows a possible slightly negative slope. Regardless, maximum vehicular speed (*wkMxSpd*) is clearly the weakest of the three variables, despite the thought that it may have some deterrent effects on transfer volumes because of a perceived reduction in walking safety. By comparison, the walking distance (*wkLen*) demonstrates the strongest effect on transfer volumes.

The full and final model fit only included the percentage of ORCA youth riders at both the Link station and on the originating line. As covered in the *Model Building* section, only youth cards were selected because of multicollinearity. Youth ridership can serve as a proxy for healthier and younger riders—who tend transfer more willingly (*2*). Similar studies refer to health not only as lacking a disability, but also as a greater likelihood of engaging in physical activity (*2*). To more closely examine the effects each ORCA card type might have on the model, a separate model was fit for each of the other non-adult cards: disabled, senior, and LIFT. The results were used to simulate expected transfers given a change within the ranges of each user type found in the data. In these plots, results are only focused on interpolation, not extrapolation. It is unlikely that ridership would diverge considerably from existing observations, so extrapolating beyond these ranges was deemed unnecessary. This was done for each card type at both the destination Link station and the originating line, and is presented in Figure 10.

Based on Figure 10, some user groups appear to have limited impact on transfer volumes. For example, disabled ORCA ridership at both the LLR station and origin line levels appear to has a limited effect on transfer volumes. At the origin line level, the effect of disabled riders on transfer volumes was significant at a 0.01 level, but was weaker than that of youth ridership's effect (.09 ± .06 compared to .12 ± .04). Similarly, senior ridership at the origin line level was significant at the 0.01 level, while it was also very similar in effect compared to youth ridership (.16 ± .08). However, the range of senior ridership proportion on origin bus lines was smaller than that of youth riders (1.4%–8.8% compared

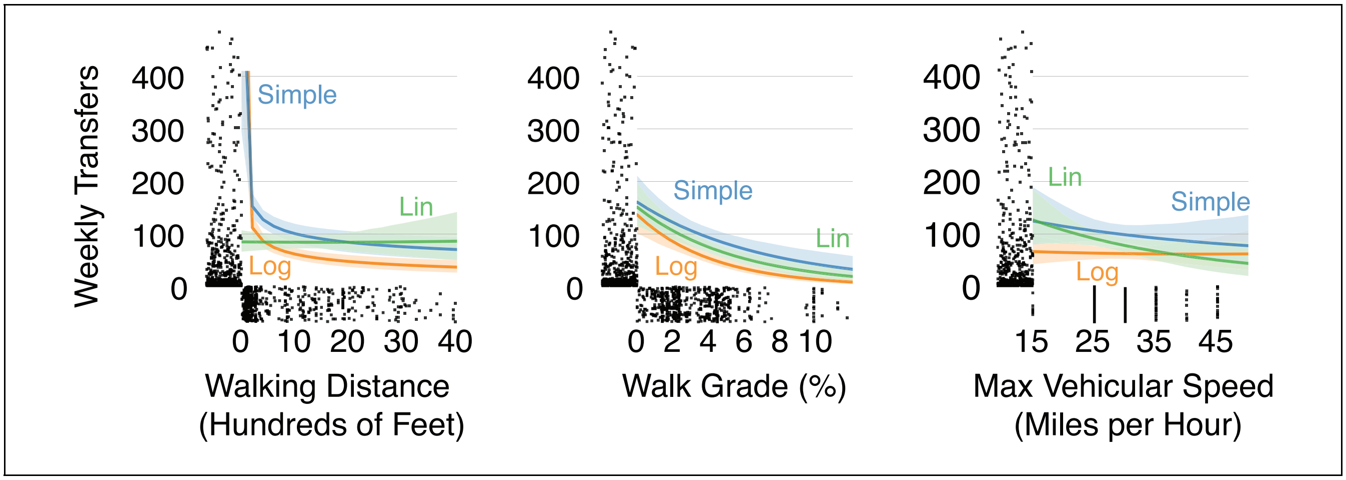

Figure 9. Comparison of expected transfer volumes given changes in walking route characteristics for the simple model (only walking route variables) and the full models with either a logarithmic walking distance (log) or a linear walking distance (lin).

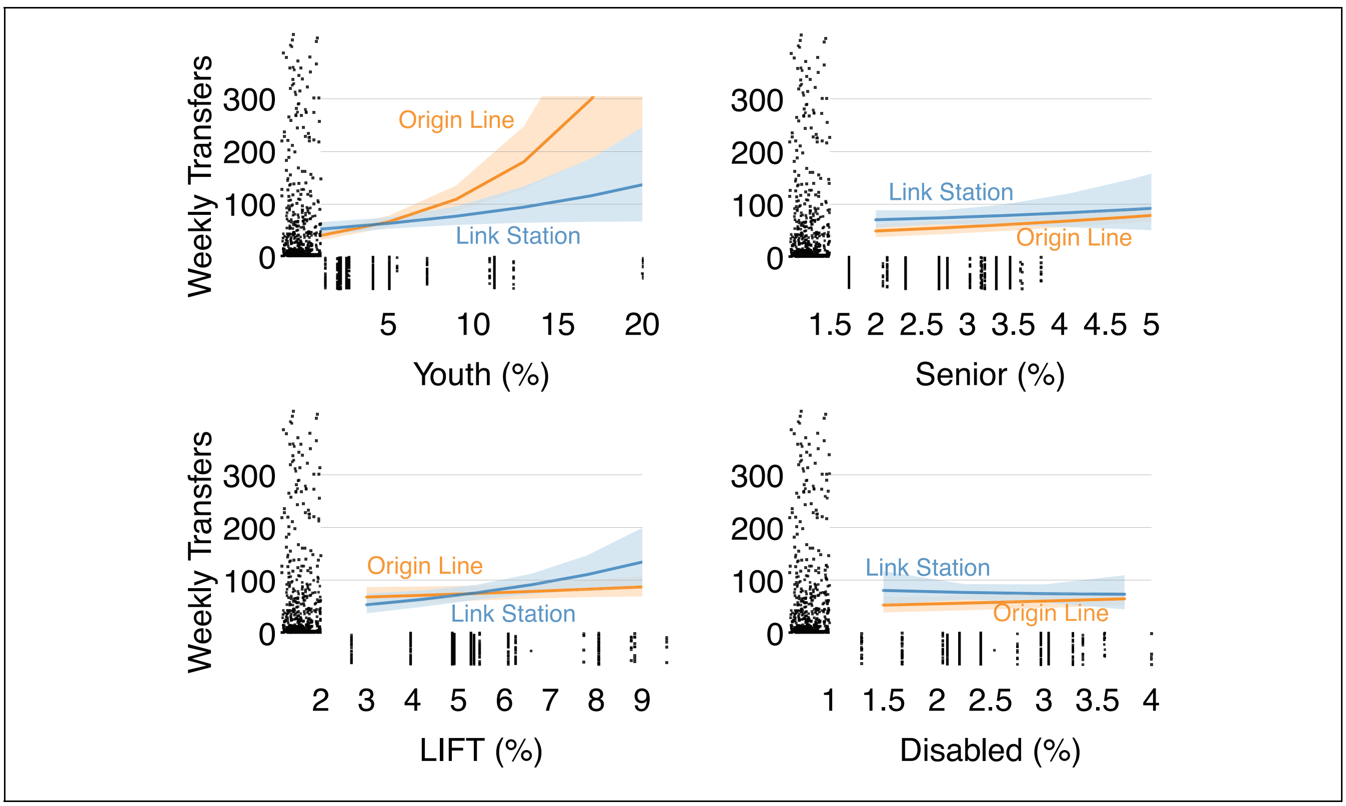

Figure 10. Comparison of expected transfer volumes given changes in rider composition at the Link Light Rail station and on the origin bus line.

to 0.8%–22.7%). In all cases, the impact was not significantly different between the origin line and Link station, even if one was considered significant.

The group of riders most similar to youth appears to be LIFT (Figure 10). Higher proportions of LIFT riders also seem to explain higher transfer ridership, especially at the LLR station level (.16 ± .12). This effect was both statistically significant at the .01 level and possibly stronger than the effect of youth ridership at the LLR station (.05 ± .04), although the confidence intervals overlap suggesting this difference may not be significant at the .05 level. It is possible that LIFT users are transit-captive at higher rates than other user types. This would support the idea that Link stations with a greater proportion of LIFT users are more likely to have higher transfer volumes, potentially connecting to transit-captivity as low-income users are forced into making transfers regardless of convenience because of more limited travel options compared to middle- and higher-income riders. In any case, using youth ridership improved the model goodness-of-fit more than any of the other rider groups alone. Both the AIC and the BIC of the youth models were lower than any of the other three models tested.

Conclusions

Findings

Based on this study and the referenced work, walking distance between stations is a notable barrier to effective and convenient transfers. Along with distance, both the mean and maximum slope of a walk can influence would-be transfers negatively. However, using vehicular speeds to gauge perceived safety did not offer as clear an effect in this study. While smartcards are designed to at least make the transferring experience more clear and efficient, these physical barriers continue to affect perceptions and actions related to transferring.

With regard to equity, it is clear that users with limited mobility or disability are not transferring at the same rates as other users. This could partially be because of a general aversion to transfers or high physical barriers caused by Seattle's innate hilly topography, but is directly tied to these users' limited mobility. Conversely, at stations with higher proportions of low-income and youth users, transfers are more common according to the model. These stations in practice include Mount Baker, Othello, and Rainier Beach, which all have above average youth and LIFT ridership. This suggests some level of potential transit-captivity for LIFT users. The considerable increase of transfers among stations with high youth ridership is an expected confirmation of survey-based research that considered youthfulness and health as explanatory variables for transfers.

Planning Implications

In response to these findings, there are a few potential planning implications for practitioners in the region to consider. Clearly, station proximity plays a large role, but cannot always be adjusted to the extent necessary because of other factors. However, given the effect of vehicular speed in some models, it is worth considering identifying transfer routes where vehicular speed may pose a perceived barrier. Finally, in planning future stations, planners should consider the street networks in the surrounding areas. Placing stations as close to grid-like, dense, and non-circuitous networks helps support increased transfers, as routes between stations are likely easier to follow in these networks because of their order. In less dense regions, mobility hubs could provide support, and these hubs could offer amenities such as real-time updates on arrivals and departures.

When considering different user groups, supporting low-mobility users may simply be a case of finding crucial transfers and reducing the proximity for their connections. Other solutions might include more demand-responsive measures to eliminate the need for inconvenient and difficult transfers, where possible. In the case of low-income users, ensuring that transit-captive populations have access to frequent and affordable transit that mirrors their schedule is important. Seattle's current network is focused considerably on first-shift commutes. Making sure that lines serving lower-income neighborhoods are responsive to their potentially different needs in span (the hours per day of service), frequency, and location are crucial. Span, in particular, becomes relevant when considering many of these workers are second- and third-shift workers, a time period when the frequency and convenience of Seattle's transit network is comparatively limited.

Future Research

Future studies on this topic could incorporate ridership surveys in the region to validate or expand on the topics covered by this study. A combination of both observed data and stated preferences could help elucidate the reasons for certain barriers and whether they line up with perceptions. Potentially, using more precise data on user groups' transfer habits could expand on the briefly discussed equity questions brought up in this study. Unfortunately, the limitations of our data prevent us from making much more than exploratory inferences about the behaviors of potentially transit-captive populations; however, one could apply more granular demographic data to better assess the behaviors of transit-captive populations where available, as this was a clear limitation of the data used in this study. Other possibilities include looking at the stations themselves and their amenities to describe whether these have any effect on transfers. An examination of transfers from the Link or including other modes is also a possibility for future studies.

Acknowledgments

The authors thank TRAC for their data and other assistance during this project. The authors thank Ryan Avery and Mark Hallenbeck of that organization for their work in support of this research.

Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: J. Eager, C.H. Bae; data collection: J. Eager; analysis and interpretation of results: J. Eager, C.H. Bae, E. McCormack; draft manuscript preparation: J. Eager, C.H. Bae, E. McCormack. All authors reviewed the results and approved the final version of the manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

James Eager https://orcid.org/0000-0003-1926-2388
Chang-Hee Christine Bae https://orcid.org/0000-0003-0579-0913
Edward D. McCormack https://orcid.org/0000-0002-2437-9604

References

1. Administration, U. F. T. *Transit Cooperative Research Program Synthesis 19: Passenger Transfer System Review*. Technical Report. U.S. Federal Transit Administration, 1996. http://onlinepubs.trb.org/onlinepubs/tcrp/tsyn19.pdf.
2. Cheng, Y.-H., and S.-Y. Chen. Perceived Accessibility, Mobility, and Connectivity of Public Transportation Systems. *Transportation Research Part A: Policy and Practice*, Vol. 77, 2015, pp. 386–403. https://doi.org/10.1016/j.tra.2015.05.003. https://www.sciencedirect.com/science/article/pii/S0965856415001238.
3. Currie, G. The Demand Performance of Bus Rapid Transit. *Journal of Public Transportation*, Vol. 8, No. 1, 2005, pp. 41–55. https://doi.org/10.5038/2375-0901.8.1.3. https://scholarcommons.usf.edu/jpt/vol8/iss1/3.
4. Cascajo, R., E. Lopez, F. Herrero, and A. Monzon. User Perception of Transfers in Multimodal Urban Trips: A Qualitative Study. *International Journal of Sustainable Transportation*, Vol. 13, No. 6, 2019, pp. 393–406. https://doi.org/10.1080/15568318.2018.1476632.
5. Garcia-Martinez, A., R. Cascajo, S. Jara-Díaz, S. Chowdhury, and A. Monzón. Transfer Penalties in Multimodal Public Transport Networks. *Transportation Research Part A General*, Vol. 114, 2018, pp. 52–66.
6. Yu, B., Z.-Z. Yang, P.-H. Jin, S.-H. Wu, and B.-Z. Yao. Transit Route Network Design-Maximizing Direct and Transfer Demand Density. *Transportation Research Part C: Emerging Technologies*, Vol. 22, 2012, pp. 58–75. https://doi.org/10.1016/j.trc.2011.12.003. https://www.sciencedirect.com/science/article/pii/S0968090X11001689.
7. Iseki, H., and B. Taylor. Not all Transfers are Created Equal: Towards a Framework Relating Transfer Connectivity to Travel Behaviour. *Transport Reviews*, Vol. 29, No. 6, 2009, pp. 777–800. https://doi.org/10.1080/01441640902811304. http://offcampus.lib.washington.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=44746301&site=ehost-live.
8. Guo, Z., and N. H. M. Wilson. Assessment of the Transfer Penalty for Transit Trips Geographic Information System-Based Disaggregate Modeling Approach. *Transportation Research Record: Journal of the Transportation Research Board*, 2004. 1872: 10–18.
9. Guo, Z., and N. H. M. Wilson. Assessing the Cost of Transfer Inconvenience in Public Transport Systems: A Case Study of the London Underground. *Transportation Research Part A: Policy and Practice*, Vol. 45, No. 2, 2011, pp. 91–104. https://doi.org/10.1016/j.tra.2010.11.002. https://www.sciencedirect.com/science/article/pii/S0965856410001564.
10. Stopher, P., and S. Greaves. Household Travel Surveys: Where are we Going? *Transportation Research Part A: Policy and Practice*, Vol. 41, 2007, pp. 367–381. https://doi.org/10.1016/j.tra.2006.09.005.
11. ORCA. ORCA: Get a Card, 2021. https://orcacard.com/ERG-Seattle/getACard.do. Accessed May 2021.
12. Pelletier, M.-P., M. Trépanier, and C. Morency. Smart Card Data Use in Public Transit: A Literature Review. *Transportation Research Part C: Emerging Technologies*, Vol. 19, No. 4, 2011, pp. 557–568. https://doi.org/10.1016/j.trc.2010.12.003http://www.sciencedirect.com/science/article/pii/S0968090X1000166X.
13. Blythe, P. Improving Public Transport Ticketing Through Smart Cards. *Proceedings of the ICE -Municipal Engineer*, Vol. 157, 2004, pp. 47–54. https://doi.org/10.1680/muen.2004.157.1.47.
14. Sound Transit. *Sound Transit 3: The Regional Transit System Plan for Central Puget Sound*. Technical Report, Sound Transit, 2016. https://st32.blob.core.windows.net/media/Default/Document%20Library%20Featured/8-22-16/ST3_System-Plan_2016_web.pdf.
15. Sound Transit. *System Expansion Implementation Plan*. Technical Report. Sound Transit, 2018. https://www.soundtransit.org/sites/default/files/documents/system-expansion-implementation-plan-june-2018.pdf.
16. Lindblom, M. It Took 11 Years, but Sound Transit Officially Breaks Ground for Lynnwood Light-Rail Line. *The Seattle Times*. https://www.seattletimes.com/seattle-news/transportation/it-took-11-years-but-sound-transit-officially-breaks-Section:TrafficLab https://www.seattletimes.com/seattle-news/transportation/it-took-11-years-but-sound-transit-officially-breaks-ground-for-lynnwood-light-rail-line.
17. Welch, T. F. Equity in Transport: The Distribution of Transit Access and Connectivity Among Affordable Housing Units. *Transport Policy*, Vol. 30, 2013, pp. 283–293. https://doi.org/10.1016/j.tranpol.2013.09.020. https://www.sciencedirect.com/science/article/pii/S0967070X13001534.
18. Welch, T. F., and S. Mishra. A Measure of Equity for Public Transit Connectivity. *Journal of Transport Geography*, Vol. 33, 2013, pp. 29–41. https://doi.org/10.1016/j.jtrangeo.

2013.09.007. https://www.sciencedirect.com/science/article/pii/S0966692313001762.

19. Song, Y., H. Kim, K. Lee, and K. Ahn. Subway Network Expansion and Transit Equity: A Case Study of Gwangju Metropolitan Area, South Korea. *Transport Policy*, Vol. 72, 2018, pp. 148–158. https://doi.org/10.1016/j.tranpol.2018.08.007. https://www.sciencedirect.com/science/article/pii/S0967070X17302664.
20. Google Developers. GTFS Static Overview, 2021. https://developers.google.com/transit/gtfs. Accessed May 2021.
21. Prommaharaj, P., S. Phithakkitnukoon, M. G. Demissie, L. Kattan, and C. Ratti. Visualizing Public Transit System Operation With GTFS Data: A Case Study of Calgary, Canada. *Heliyon*, Vol. 6, No. 4, 2020, p. e03729. https://doi.org/10.1016/j.heliyon.2020.e03729. https://www.sciencedirect.com/science/article/pii/S2405844020305740.
22. Haklay, M., and P. Weber. OpenStreetMap: User-Generated Street Maps. *IEEE Pervasive Computing*, Vol. 7, No. 4, 2008, pp. 12–18. https://doi.org/10.1109/MPRV.2008.80.
23. Boeing, G. OSMnx: New Methods for Acquiring, Constructing, Analyzing, and Visualizing Complex Street Networks. *Computers, Environment and Urban Systems*, Vol. 65, 2017, pp. 126–139. https://doi.org/10.1016/j.compenvurbsys.2017.05.004.
24. Boeing, G. Urban Spatial Order: Street Network Orientation, Configuration, and Entropy. *Applied Network Science*, Vol. 4, No. 1, 2019, p. 67. https://doi.org/10.1007/s41109-019-0189-1.
25. Ward, M. D., and J. S. Ahlquist. *Maximum Likelihood for Social Science: Strategies for Analysis*. Cambridge University Press, Cambridge, England, 2018.
26. Cameron, A., and P. K. Trivedi. Regression-Based Tests for Overdispersion in the Poisson Model. *Journal of Econometrics*, Vol. 46, No. 3, 1990, pp. 347–364. https://doi.org/10.1016/0304-4076(90)90014-K. https://www.sciencedirect.com/science/article/pii/030440769090014K.