Does the Squeaky Wheel Get the Complaint? Linking Bus performance, Sociodemographic Characteristics, and Customer Comments

For Citation Please use: Hawa, L., DeWeese, J., Diab, E., Crumley, M., & El-Geneidy A. (2023). Does the squeaky wheel get the complaint? Linking bus performance, sociodemographic characteristics, and customer comments. *Travel Behaviour and Society,* 32, 100594.

ABSTRACT

Public transit agencies rely on service operations performance measures to guide service improvement efforts. They also collect customer feedback on performance to measure levels of satisfaction among users. Connecting these two performance measures can help public transport agencies increase ridership satisfaction and loyalty. In this study, we link service performance measures obtained from automatic vehicle location (AVL) and automatic passenger count (APC) with customer complaints. This is done while controlling for the amount of ridership to understand if routes with worse service performance, as identified by operations measures, also have more service complaints. We also investigated if an area's level of affluence affects the number of service complaints per rider that the route receives while controlling for route service performance and ridership. The AVL/APC and customer-feedback data were provided by Portland, Oregon's TriMet transport agency for the period between August 2018 and January 2019. Descriptive statistics at the route level and a series of mixed-effect multilevel logistic regression models were used to quantify the relationship between route service performance, service complaints, and a service-area vulnerability index, at the route-day level. The likelihood of receiving a service complaint for a route in a day was found to increase based on service performance and the vulnerability of the neighborhood being served by that route, all else held equal. Findings from this research unmask the relationship between service complaints, bus operations and socioeconomic characteristics of the neighborhood a route is serving, offering insights to transport planners and researchers in the psychology behind bus performance complaints.

Key words: Automatic vehicle location (AVL), Automatic passenger count (APC), Customer complaints, Bus operations, Vulnerability index

1. INTRODUCTION

Bus transport agencies generate performance measures regularly to evaluate their performance and apply for funding from various levels of government (Berkow, El-Geneidy, Bertini, & Crout, 2009; Hartman, Kurtz, & Winn, 1994). At the same time, they collect customer feedback to understand how their services are perceived by the public (Harreman-Fernandes et al., 2021; van Lierop, Badami, & El-Geneidy, 2018; van Lierop & El-Geneidy, 2016). Customer feedback comes in several formats, either through customer satisfaction surveys—conducted by the agency or a subcontractor—or from comments and complaints the agency receives directly through various channels such as phone calls, web forms, or social media posts.

Conducting customer satisfaction surveys to understand the perceived service quality is expensive. Additionally, it can face many challenges from bus operators' union rules, sampling errors, and questionnaire design defects. These questionnaires can be simply not asking the right questions. They could be collected in specific time periods failing to discover experienced issues by transit users in a timely fashion. In contrast, customer comments and complaints can provide better value by understanding service quality in a timely manner as they promptly reflect defects in the service quality. Therefore, transit agencies of different sizes tend to report the number of annual complaints as an indicator in their annual reports to understand the perceived quality of service (Diab, Badami, & El-Geneidy, 2015; Saskatoon Transit, 2021; STM, 2021). Accordingly, some researchers argue that customer comments and complaints are a better indicator to understand the perceived service deficiencies (Harreman-Fernandes et al., 2021).

On the other hand, customer feedback from different resources (i.e., surveys or complaints) is generally handled by a dedicated department in public transport agencies while bus operations and planning are handled by a different department. These two entities within a public transport agency usually interact at higher management levels and the data shared between them are mostly at the aggregate level. The data are rarely linked. Linking customer feedback to operations at the disaggregate level has become a subject of interest over the past years in part because doing so has the potential to generate more targeted operations strategies that are more directly responsive to public needs (Carrel, Mishalani, Sengupta, & Walker, 2016; Harreman-Fernandes et al., 2021).

More specifically, linking customer feedback to operations data collected from automatic vehicle location (AVL) and automatic passenger counters (APC) has shown to be of value and has been the focus of a few previous studies (Carrel et al., 2016; Harreman-Fernandes et al., 2021; van Lierop & El-Geneidy, 2017). For example, van Lierop and El-Geneidy (2017) linked customer satisfaction questionnaires to AVL/APC data collocated for an express bus route in Vancouver, British Columbia, Canada to explore the main factors influencing customer satisfaction. Similarly, Carrel et al. (2016) demonstrated the value of connecting customer satisfaction questionnaires to AVL and smartphone tracking data to understand customer satisfaction. In contrast, Harreman-Fernandes et al. (2021) focused on exploring qualitatively the list of most common complaints. They also used summary statistics to understand the distribution and frequency of user complaints and to explore the link between AVL/APC data and customer complaints for four different routes in Portland, Oregon. This is to demonstrate the potential use of customer complaints and transit operations data to help identify and validate perceived service deficiencies.

Despite these previous efforts, none of them explored the relationship between transit service complaints, service quality, and equity issues by modeling customer complaints as a function of service quality and socioeconomic issues. This will help in exploring the drivers that increase the odds of receiving a complaint from a certain group transit user. Such effort will help transit agencies to adjust the service to minimize users' stress and criticisms, improving the overall image of the transit system. In other words, the nature of the relationship between customer comments, bus operation, and the surrounding operating context and environment requires more careful examination.

To address this gap in the literature, the aim of this research paper is twofold, the first is to model the relationship between transit service operation issues, identified as on-time performance (OTP), and actually received negative customer comments (complaints) related to operations, while controlling for the number of passengers. The second is to understand the relationship between negative customer comments (complaints) and the level of affluence in the neighborhoods the bus routes serve, while controlling for OTP and passenger demand. To achieve these goals, we analyzed customer comment and suggestion data and archived AVL/APC data provided by the Tri-County Metropolitan Transportation District of Oregon (TriMet) agency of the Portland, Oregon. In this study, we use data collected from August 2018 through January 2019 for the 82 regular bus routes. Buses that only run at night were excluded from the study. Findings from this study can be of value to public transport operators, planners and researchers as it models the link between bus operations, customer feedback, and demographics of the areas served, enabling agencies to develop better bus operations policies that respond to customer concerns.

2. LITERATURE REVIEW

2.1 Bus performance measures

Stop-level bus data are generated by bus dispatch systems that include AVL/APC technology, which grew in popularity from 1995 to 2004. During this time period, the number of public transit agencies that deployed AVL/APC systems in the United States expanded from 86 to 257 (Radin, 2005), with comparable levels of implementation across the world (Schweiger, 2003). A good example is TriMet, which relies on these systems (Berkow et al., 2009). Data from AVL/APC systems are used in several application that includes generating bus service performance indicators, developing bus schedules, and managing service quality for real-time applications (Barabino, Francesco, & Mozzoni, 2017; Hounsell, Shrestha, & Wong, 2012; Ma, Ferreira, & Mesbah, 2014; Mandelzys & Hellinga, 2010; Moreira-Matias, Mendes-Moreira, de Sousa, & Gama, 2015; Verbich, Diab, & El-Geneidy, 2016).

General transit performance measures from archived stop-level AVL/APC data, including the percentage of on-time performance, are widely used for transport planning and operations (Bertini & El-Geneidy, 2003). Bus on-time performance, a measure of schedule adherence, is particularly pertinent when measuring transit service reliability (Strathman et al., 1999). Bus service reliability is a unique service performance measure because it impacts the decisions of both the service operators and users (Abkowitz, Slavin, Waksman, Englisher, & Wilson, 1978). According to a comprehensive review of service reliability indicators for 15 cities across North America, OTP was found the most common measure used by transit agencies (Diab et al., 2015). Several researchers used OTP to perform social equity analysis (Javanmard, Lee, Kim, Liu, & Diab, 2023; Palm, Shalaby, & Farber, 2020), understand bus service quality before and after service improvement (Surprenant-Legault & El-Geneidy, 2011), or propose other indicators to capture bus service quality (Barabino, Lai, Casari, Demontis, & Mozzoni, 2017).

2.2 Customer comments and complaints

Public transport agencies often rely on technical service measures to guide service improvements, for example decreasing travel times and improving on-time performance, in an effort to increase ridership. However, riders' perceptions of public transport service quality and their level of satisfaction is what actually steers their travel behavior choices (Abou-Zeid, Witter, Bierlaire, Kaufmann, & Ben-Akiva, 2012). In fact, customer satisfaction has been linked to customer loyalty, including whether existing public transport users will continue to use and refer it to other people (Diab, van Lierop, & El-Geneidy, 2017; Zhao, Webb, & Shah, 2014). Public transport customer satisfaction depends on both the personal attributes and attitudes of the rider as well as the bus service attributes associated with their on-board and off-board experience (Eboli & Mazzulla, 2009; Harreman-Fernandes et al., 2021; Hensher, Stopher, & Bullock, 2003; Mouwen, 2015; Tyrinopoulos & Antoniou, 2008). To demonstrate this with an example, overall express bus route passenger satisfaction was found to depend on their satisfaction with both service measures and other aspects of the user experience: crowding, service frequency, safety while on the bus, and cleanliness (Harreman-Fernandes et al., 2021; van Lierop & El-Geneidy, 2017).

To analyze customer satisfaction and loyalty, a considerable number of studies used customer satisfaction surveys collected by researchers or transit agencies (van Lierop et al., 2018). These surveys provide a snapshot of the satisfaction of transit users at a point in time, which can change over time or can be missing important events impacts users' perception. Therefore, transit agencies report and use the number of customer complaints in addition to customer satisfaction surveys to better understand users' overall perception (Diab et al., 2015; Saskatoon Transit, 2021; STM, 2021). These negative complaints affect the transit system's overall image and may influence users' willingness to the use the system in the future. Nevertheless, a very limited number of studies focused on understanding, categorizing, and analyzing customer complaints (Harreman-Fernandes et al., 2021).

2.3 Linking operations data to customer feedback

Connecting customer satisfaction with service performance measures based on operations data is an important step in investigating where to invest funding to improve service (Allen, Muñoz, & Ortúzar, 2019; Carrel et al., 2016; Friman & Fellesson, 2009; Harreman-Fernandes et al., 2021). Linking operations data to customer feedback is also important for teasing out the differences between perceived and actual public transport service quality, and for understanding how customers might respond to service improvements or a decline in service (Diab et al., 2015; Harreman-Fernandes et al., 2021; van Lierop & El-Geneidy, 2017). Although bus rider satisfaction can depend on service and non-service specific bus attributes, operations data have been rarely linked to customer feedback specifically about operations. Harreman-Fernandes et al. (Harreman-Fernandes et al., 2021) successfully linked passenger feedback about TriMet bus service to bus operation issues obtained from TriMet's archived AVL/APC data. Using summary statistics, they indicated that linking the two datasets can help suggest which service improvements can address complaints (Harreman-Fernandes et al., 2021).

2.4 Vulnerability index / equity in transport

Research on vertical transport equity, the notion that services should be allocated in accordance with need, typically employs social indicators, or vulnerability indexes, to define potentially disadvantaged groups (Carleton & Porter, 2018). Vulnerability indexes combine multiple aspects of disadvantage into one multi-faceted metric (Carleton & Porter, 2018). Selecting attributes to include in a vulnerability index should be context specific and reflect the characteristics that the research aims to discuss (Foth, Manaugh, & El-Geneidy, 2013). Common socioeconomic factors of transport equity are also frequently discussed in social equity literature as well as in regulations and laws, including race and ethnicity, income, and employment status (Carleton & Porter, 2018). Factors that are included in a vulnerability index are chosen to represent populations that could be considered disadvantaged according to their geographic context (Carleton & Porter, 2018). We thus chose median household income, percent of residents that are non-white, and proportion of residents who do not have a bachelors' degree to generate our vulnerability index based on the American context and their incorporation in previous research (Carleton & Porter, 2018; Grengs, 2014; Kramer & Goldstein, 2015).

3. METHODS AND DATA

This paper investigates the relationship between passenger complaints, bus operations, and ridership sociodemographic characteristics. To achieve this goal, we used three data sets, complaints and AVL/APC data obtained from TriMet and socioeconomic data calculated at the route level and obtained from the US Census. Complaint data were linked to the operations (AVL/APC) data using route and date identifications. In other words, we linked each complaint record using the route and incident date to the bus operation records during the same day that the complaint is associated with. It should be noted that some discrepancies might occur as incident reports might be wrongly coded or dated. These complaints were excluded when possible.

For the analysis, first we conduct a descriptive analysis by ranking the TriMet bus routes by service measures (the percent of late bus trips, late OTP), complaints (number of service complaints per average daily ridership per route), and service area sociodemographic attributes. We concentrate on the top 20 routes based on three sorting orders. This ranking analysis enables us to identify if bus routes performing poorly compared to others are receiving a high number of complaints per average daily ridership per route. The analysis helps determine if the number of complaints obtained from routes serving affluent neighborhoods are higher compared to those serving low income areas. As the ranking analysis is more of an exploratory nature, we generated a series of logit regression models to quantify the relationship between service operation measures and route sociodemographic attributes with bus rider complaints. The dependent variable in the first model is the odds of receiving a complaint related to operations that day, while the independent variables include OTP, passenger demand, passenger load, and socioeconomic characteristics of the area around each route. The other three models concentrate on certain kinds of complaints in a day and their relationship to operations and demographic characteristics.

3.1 Complaints

Withing the context of transit planning and operations, a complaint can refer to a negative comment by a transit user in which they express their dissatisfaction or a problem that they are experiencing and would like the transit agency to resolve. Transit agencies usually receive these complaints directly from the users and used as an indicator to understand the experienced issues (Diab et al., 2015; Saskatoon Transit, 2021; STM, 2021). For the purpose of this study, complaints were classified into standardized categories based on Harreman-Fernandes et al.'s methodology (Harreman-Fernandes et al., 2021). Before cleaning the data, there were 3,887 unique complaint records which were categorized into 18 groups (Harreman-Fernandes et al., 2021). The dataset was cleaned to limit complaint incident report dates to weekdays between August 20, 2018, and January 31, 2019. Holidays were also removed from the dataset, including Labor Day (September

3, 2018), Thanksgiving Day (November 22, 2018), Christmas Day (December 25, 2018), New Year's Eve (December 31, 2018), and New Year's Day (January 1, 2019). Additionally, complaints about routes that were deadheads, strictly night buses, or otherwise behaved oddly were removed. Routes 7, 13, 42, 272, 291, 921, 922, and 992, were removed from the complaints data, which left 82 bus routes in the dataset. Complaints that were not associated with a route were also left out of the dataset. After cleaning the complaints data based on route and date, 2,756 remained. Complaints that could not be logically related to service operations, such as those regarding vandalism or cleanliness, were excluded from this analysis, as we retained only complaints pertaining to bus service operations. The retained categories were: bus deviation from schedule (541 complaints), pass-up incidents (727 complaints), and reckless driving behavior (262 complaints), for a total of 1,530 service complaints with a confirmed date and route information.

3.2 AVL/APC

For the operations dataset, TriMet provided detailed AVL/APC information. TriMet is one of the leading transit agencies in North America in employing AVL/APC systems. These systems provide stop-level information with a considerable level of accuracy (Furth, 2000; Kimpel, Strathman, & Callas, 2008; Strathman et al., 2001), which has been used extensively in academic research to understand and model transit service quality since the 1990s (El-Geneidy, Strathman, Kimpel, & Crout, 2006; Kimpel, Strathman, Bertini, & Callas, 2005; Strathman et al., 1999). The initial dataset contained 56.2 million stop-level observations. The AVL/APC data was prepared to be matched with the complaints. Dates were filtered to match the retained complaints data, while removing recording errors and duplicate records. More specifically, to match the complaints dataset, dates were filtered by removing weekends, holidays, any observation before August 20, 2018, or after January 31, 2019, as were observations associated with a route that behaved oddly (7, 13, 42, 272, 291, 921, 922, and 992) due to service changes. Using a similar approach discussed by Verbich et al. (2016) recording errors and duplicate records were removed from the dataset. After this process, there were 40.1 million stop-level observations remaining in the AVL/APC dataset, which were used in the analysis.

The AVL/APC data were used to generate service performance measures for each of the 82 bus routes over the entire study period, which are mapped in **Figure 1**. Performance measures include average late and early OTP in a day, total boardings per trip in a day, average bus load of a bus in a day, average max speed in day, and the variation of the previously mentioned variables. A dummy for each stop being late or early was created and used to build the averages for OTP late and early in a day. For our study, we considered a bus late if it departed a stop more than 4 minutes behind schedule and early if it left more than 2 minutes before based on the difference between scheduled and actual leave times for each stop in the AVL/APC data. We selected these temporal thresholds as a compromise between widely varying industry standards. Internally, TriMet uses a different metric to classify earliness or lateness (1 minute for early and 5 for late). For the purposes of this study, the discrepancy is unlikely to have an impact as users are more likely to formulate their complaints based on subjective experiences rather than by reference to specific agency standards. In the future, however, research such as this could be used to calibrate and validate measurements of OTP by exploring which thresholds most accurately predict user complaints in statistical models. It should be noted that OTP is one of the most common measure used by transit agencies to understand service reliability (Diab et al., 2015). In fact, transit agencies of different scales including TriMet, STM in Montreal, and TTC in Toronto use and report OTP as a measure for transit service reliability for both frequent and infrequent transit bus service (STM, 2021; TriMet, 2023; TTC, 2023).

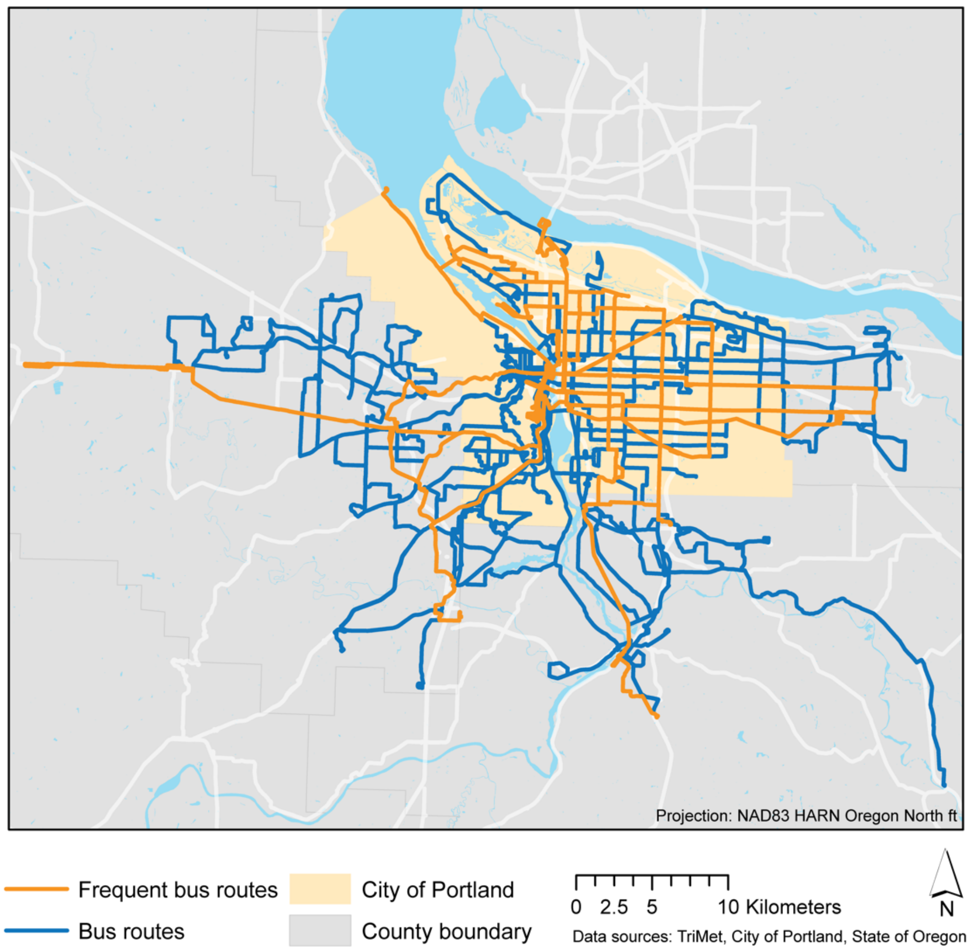

Figure 1 Study area

To prepare the AVL/APC data for regression modeling, the service characteristics were summarized at the route-day level. Thus, performance measures were calculated for every route for every day that it ran in the dataset, which resulted in 9,338 observations. The service date (which is at midnight for each day the bus operated) associated with the AVL/APC observation was used as the day that the observation of the bus was made. In some cases, stop times (in seconds after midnight) occur more than 24 hours after midnight of the service date, thus spilling into the following day. These more-than-24-hour-later stops are a continuation of trips from the previous day and they were counted with the previous day.

3.3 Sociodemographic characteristics

The tidycensus R package was used to retrieve census data for the Portland census metropolitan area. The census data that was collected is from the 2014-2018 five-year estimate by block group that is sourced from the American Community Survey (ACS5). The vulnerability index (VI) of each census block was calculated by summing three z-scores: inverse of median household income, proportion of population without a bachelor's degree, and the percent of residents that are non-white, divided by three (DeWeese et al., 2020). A larger VI indicates that the population in the census block is more vulnerable and a smaller VI indicates that the population in the census block is less vulnerable, for this purpose income was multiplied by negative 1.

The average median income and VI were generated for each route based on the areas surrounding the bus routes, which can indicate who the bus route serves. To calculate the route average sociodemographic characteristics, a 400-meter buffer around one direction of each route was mapped and intersected with the block groups. The 400-meter distance has been commonly used by researchers and in practice to understand bus transit catchment areas (Bree, Fuller, & Diab, 2020). The weighted sum of the proportion of buffer area that intersected each census block was used to calculate the route median income and VI. Thus, the route average median income and VI could be linked to the descriptive analysis of each route and the regression models at the route-day level.

4. RESULTS

4.1 Descriptive comparison

The 82 bus routes were summarized by mean late and early OTP (the percent of observations where a bus along that route was late and early), complaints, ridership, and sociodemographic attributes over the course of the study period. They were then separately ranked by late OTP, service complaints per average daily ridership per 100 people (total number of service complaints per route divided by the average daily ridership of the route, multiplied by 100), and riders per trip, and the top 20 routes from each sorting method were retained.

Table 1 shows the top 20 routes sorted by late OTP. In a transit agency's language, a route like 6 is a route that is facing major OTP issues and is expected to have high rider's complaints. Also, it is a route that a public transport agency will be directing resources towards fixing. Surprisingly, looking at the number of complaints per average daily ridership, it is clear that this route is not receiving many complaints despite its performance. The same applies to many of the routes in the list with high percentage of late OTP and low number of complaints per average daily ridership. Another observation related to **Table 1** is the average median income is around $75,000 for the top 20 routes with OTP lateness issues. There is a mix between the income and the rate of complaints that are submitted per average daily ridership, showing no clear relation between the three variables of interest.

Table 1 Top 20 routes ranked by late OTP

Route number	Mean OTP late*	Mean OTP early**	Total service complaints	Service complaints per average daily ridership (per 100 people)	Riders per trip on average	Route average median income (in dollars)	Route average VI
6	0.36	0.02	46	0.79	77.21	61,907	0.40
16	0.30	0.02	36	4.71	29.11	87,862	-0.45
23	0.28	0.00	6	4.42	10.70	65,450	0.64
24	0.27	0.01	3	0.63	19.12	78,708	-0.19
55	0.26	0.01	2	2.79	37.37	91,072	-0.55
36	0.26	0.04	5	2.29	36.20	105,313	-0.73
14	0.25	0.02	36	0.57	77.57	62,913	0.07
74	0.25	0.01	7	1.69	17.84	62,364	0.69
71	0.24	0.02	24	0.68	63.04	66,374	0.13
76	0.23	0.03	24	1.00	75.15	66,865	0.07
51	0.23	0.02	7	2.46	19.54	114,867	-0.76
10	0.23	0.02	21	0.87	57.05	63,596	0.37
66	0.23	0.07	5	0.99	53.20	81,019	-0.44
25	0.22	0.03	1	0.62	12.96	46,317	1.06
19	0.21	0.02	43	0.82	91.15	71,093	-0.07
99	0.21	0.15	17	2.05	54.27	62,228	0.06
47	0.21	0.04	14	1.44	28.15	91,776	0.27
18	0.21	0.01	2	6.61	6.27	107,466	-0.91
8	0.20	0.02	32	0.50	78.98	66,783	-0.08
77	0.20	0.04	48	0.87	105.40	59,744	0.26
Average	0.24	0.03		1.84		75,686	

** Mean OTP late refers to the percentage of times a bus departed late at stops along a route*

*** Mean OTP early refers to the percentage of times a bus departed early at stops along a route.*

Ranking the bus routes by service complaints or ridership and averaging the values of the top 20 routes reveals a different story. **Table 2** shows that the mean late OTP for the top 20 bus routes sorted by service complaints per average daily ridership is equal to the mean late OTP for the top 20 bus routes sorted by ridership per trip. However, the mean early OTP is higher for the top 20 bus routes sorted by service complaints than by ridership. As seen in the table, bus routes with worse late OTP are also linked to more service complaints per riders per day than just bus routes with more ridership per trip. Interestingly, bus routes with worse late OTP serve lower median income areas than bus routes with the highest number of service-related complaints per daily riders. Notably, the top 20 bus routes with the most service complaints per average daily ridership also had the highest percent of mean early OTP, compared with the top 20 bus routes sorted by late OTP and ridership. The bus routes with the most ridership served areas with lower average median income compared to the bus routes sorted by late OTP and least service complaints per person per day.

Table 2 Summary of top 20 routes by late OTP, complaints, and median income

	Mean OTP late*	Mean OTP early**	Complaints per average daily ridership (per 100 people)	Route average median income
Averages sorted by OTP	0.24	0.03	1.84	75,686
Averages sorted by service complaints	0.18	0.06	3.00	83,254
Averages sorted by ridership	0.18	0.03	0.70	64,764

** Mean OTP late refers to the percentage of times a bus departed late at stops along a route*
*** Mean OTP early refers to the percentage of times a bus departed early at stops along a route.*

4.2 Statistical Models

Bus route service measures, service complaints, and sociodemographic characteristics were intersected and analyzed at the route-day level of analysis using mixed-effect multilevel logistic regression models. In this study, we used four different dependent variables in 4 statistical models using the same statistical technique. The first dependent variable is related to all service complaints, the value of this variable is one if the agency received any complaints related to service performance in the study day for this particular route (Model 1). If no complaints were received in that day for the particular route, a value of zero is recorded. The second dependent variable concentrates on deviation from schedule complaints (Model 2), while the third dependent variable is concerned with complaints about a pass-up (Model 3). Finally, the fourth dependent variable concentrates on receiving reckless driving complaints (Model 4). The statistical models included 9,338 total observations, which were made up of 82 routes and 114 days (route 2 only ran 104 days). Due to the longitudinal nature of the data in terms of the repetition of the data for the same route, multilevel modeling was utilized in this study.

Table 3 includes the summary statistics for all the variables included in the statistical model. In addition to regular routes, TriMet operates special frequent bus services. We controlled for these in the model using the “frequent service routes” variable (17 out of the 82 routes). Early OTP is the percentage of times a bus departed early (2 minutes early or more) in a day along one route, while late OTP is the percentage of time the bus departed late at a stop in a day along a route (4 minutes late or more). The mean load, mean load squared, and boardings per trip were calculated at the route-day level from the AVL/APC data and used in the models. The mean maximum speed was generated from the AVL/APC data at the route-day level, but the variable was left out of the models because we could not control for bus routes that were meant to travel at faster speeds. For example, if a portion of their route was on the highway as this was out of the scope of the study. The average median income of the area surrounding the bus routes was tested in the models instead of the average VI of the areas surrounding the bus routes. Although the route median income was significant, we chose to use the route average VI to account for additional sociodemographic characteristics (education level and percent of the population that is non-white) in addition to average median income.

Table 3 Summary statistics (for data at the route-day level)

Categorical Variables	**Percent of observations**			
Frequent service route	20.65			
Dependent variable = any service complaint	12.81			
Dependent variable = deviation from schedule complaint	5.22			
Dependent variable = pass up complaint	6.70			
Dependent variable = reckless driving complaint	2.53			
Continuous Variables	**Mean**	**Std. Dev**	**Min.**	**Max.**
Mean OTP late	0.16	0.10	0.00	0.76
Mean OTP early	0.04	0.05	0.00	0.46
Mean load	6.54	3.52	0.00	20.50
Mean load squared	55.14	52.60	0.00	420.43
Boardings per trip per day	46.85	40.29	0.40	706.54
Route average VI	-0.0068	0.44	-0.91	1.06

Table 4 shows the models' output reported as the odds ratio. The regression results for Model 1 suggest that a 1% increase in the percentage of time a bus departs late along a route would increase the odds of a service complaint by 5.69 times on that day, while keeping all other variables constant at their means. This variable was positive and statistically significant, yet it is important note that a 1% increase in the delay along bus route in a day is a substantial amount in bus operations. Being early did not show statistical significance in the model, yet the coefficient and the confidence interval for this variable requires careful interpretation. A 1% increase in the proportion of bus observations that departs from the stop early increases the odds of service complaint by 5.90 times. This suggests that bus riders are more sensitive to early buses than late ones and are more likely to complain if a bus is early than if it is late.

Model 1 shows that if the mean load of the bus is one person more in a day along all trips serving that route, the odds of a service complaint increase by 1.65 times, although the mean load squared odds ratio indicates that this relationship plateaus. Both variables showed statistical significance, while keeping all variables constant at their means. These two variables can be interpreted as service complaints will increase with more riders are on a bus until a certain threshold where adding more passengers in a day won't trigger such complaints. The number of boardings per route per day has a statically significant positive impact on the odds of a service complaint, while holding other variables constant at their mean. If the studied bus route is a frequent service route it increases the odds of a service complaint compared to non-frequent service routes. A 1% increase in the vulnerability index of the population surrounding a bus route increases the odds of a service complaint 1.65 times, showed a low level of statistical significance (95%). In other words, routes serving more deprived areas are receiving more complaints. A low or no statistical significance in this variable indicates that routes serving vulnerable areas are receiving as much if not more complaints related to bus service operations.

Model 2 examines the impact of service performance measures and route vulnerability index on the likelihood of a complaint about a bus's deviation from schedule. Model 2 further underscores Model 1's findings that an early bus is more likely to be associated with a complaint than a late bus. A 1% increase in early bus performance increases the odds of a complaint about deviation from schedule by 47.56 times, whereas a 1% increase in late bus performance increases the odds of a complaint about deviation from schedule by 28.39 times. Both variables showed statistical significance while keeping all variables constant at their mean. This makes sense given that bus riders could have missed the bus if it is early, and thereby, they are more likely to contact the agency and submit a complaint. Similar to Model 1, if the mean load of the bus is one person more, the odds of a deviation of service complaint by 1.74 times, and the mean load square also indicate that this relationship plateaus. Boardings per trip per day, if a bus route is a frequent service route, and the route average vulnerability index did not show a statistically significant relation to the odds of a deviation from schedule complaint occurring, although their odds ratios are still in line with their odds ratios in Model 1.

Model 3 explores the impact of service performance measures and route vulnerability index on the odds of a complaint about a pass-up occurring. Early and late OTP are not statistically significant indicators if a pass-up complaint will occur. An increase in the mean load along a route over the course of the day by one person, increases the odds of a pass up complaint by 1.79 times. This could be because a bus that has a higher load would be more likely to make a pass-up, or skip a stop. If a rider does not request a stop and there is not enough space for additional passengers, the driver might skip a stop. Like Models 1 and 2, the average load has a limited positive impact on the odds of a complaint occurring, and the odds of a pass-up complaint happening is independent of the boardings per trip per day. Being a frequent bus route increase the odds of a pass-up complaint occurring by 2.03 times compared with regular bus routes, while keeping all other variables constant at their mean. A 1% increase in the vulnerability index of the area surrounding bus routes that are more vulnerable increases the odds of a complaint about a pass up 1.83 times.

Model 4 investigates the impact of bus service performance measures and the route average vulnerability index on the odds of a complaint about reckless driving occurring. We tested this model as a multi-level and found it was not significant, such that running it as a multi-level did not change the regression results much compared to a non multi-level. However, we retained the multi-level model to enable a direct comparison across the models. An increase in the mean load by one person increased the odds of a reckless driving complaint by 1.63 times, although the odds ratio of the mean load squared once again indicates that this relationship also levels off at a certain point. An increase in the boardings per trip per day by one person increases the odds of a reckless driving complaint by 1.01 times. Whether a bus route is a frequent service route compared to if it is a regular route does not significantly change the odds of a reckless driving complaint. An increase in the route average vulnerability index increases the odds of a reckless driving complaint by 2.04 times, which could indicate that areas with more vulnerable populations are more sensitive to reckless driving behavior than other areas.

Table 4 Regression results

	Model 1 (any service complaint)				Model 2 (deviation from schedule)				Model 3 (pass up)				Model 4 (reckless driving)*			
	O.R.		95% CI		O.R.		95% CI		O.R.		95% CI		O.R.		95% CI	
Mean late OTP	5.69	**	2.17	14.91	28.39	**	8.49	94.96	0.90		0.24	3.39	1.84		0.37	9.30
Mean early OTP	5.90		0.43	81.72	47.56	*	2.53	894.89	0.17		0.00	8.00	8.22		0.28	238.98
Mean load	1.65	**	1.42	1.92	1.74	**	1.44	2.10	1.79	**	1.46	2.19	1.63	**	1.30	2.05
Mean load sq.	0.98	**	0.97	0.99	0.98	**	0.97	0.99	0.97	**	0.96	0.98	0.98	**	0.96	0.99
Boardings per trip per day	1.00	*	1.00	1.01	1.00		1.00	1.00	1.00	*	1.00	1.01	1.01	**	1.00	1.01
Frequent service route	1.81	**	1.18	2.77	1.23		0.81	1.85	2.03	**	1.28	3.23	1.25		0.87	1.81
Route average VI	1.65	*	1.10	2.47	1.18		0.78	1.79	1.83	*	1.15	2.92	2.04	**	1.35	3.08
Constant	0.01		0.00	0.01	0.00		0.00	0.00	0.00		0.00	0.01	0.00		0.00	0.00
Log likelihood	-3098.85				-1743.40				-2003.09				-1020.20			
Wald chi-squared	155.07				121.58				116.13				117.68			
Interclass correlation	0.11				0.08				0.11				0.02			
Akaike's information criterion	6225.18				3504.81				4024.17				2058.39			
Bayesian information criterion	6215.71				3569.08				4088.45				2122.67			

*p<0.05 **p<0.01

*multilevel chi-bar2 was low

5. CONCLUSION

Transit agencies understand customer feedback by using several approaches that include collecting customer satisfaction surveys or tracking complaints received directly through various channels such as phone calls, web forms, or social media posts. Results from these approaches are commonly tracked and reported annually by transit agencies to understand users' perceived quality of service. Nevertheless, while a considerable number of studies focused on analyzing customer satisfaction surveys, less attention has been given to understanding the drivers behind people's complaints. These complaints can be a better indicator to understand the perceived service deficiencies. To address this gap in the current literature, our study analyzed bus service operations data in conjunction with reported customer complaints to investigate if bus routes with worse relative performance are also those that attract more service complaints per rider, showing the perceived service deficiencies in the system. We also linked census data of the areas surrounding bus routes to investigate if the number of service complaints is related to how advantaged or disadvantaged surrounding populations are while controlling for operations and passenger demand. Modeling and quantifying the relationship between customer complaints and service quality and socioeconomic aspects, not only provide transit planners and practitioners with a meaningful understanding of the impact of service deficiencies, but also equity issues related to who is reporting more deficiencies than others. This will aid them in minimizing users' criticism of the service, which will improve the overall image of the transit system.

The investigation found that the routes with the overall highest percent of late on time performance are not necessarily always the routes with the highest number of service-related complaints per average daily ridership. This suggests the importance of examining operations data with customer feedback data when seeking to invest in service improvements to avoid complaints and keep customers happy. In ranking the top 20 routes by late OTP, service complaints per rider per day and ridership, we found that the routes with the most service complaints per average daily ridership also had the highest average proportion of early OTP compared to the other two rankings. Interestingly, the top 20 routes ranked by number of service complaints per average daily ridership passed through areas that had the highest average median income compared to the top 20 routes ranked by late on time performance and ridership.

This study also employed mixed-effect multilevel logistic regression models to quantify the relationships between bus route performance, service-related complaints per person per day, and the vulnerability index of the areas surrounding the bus routes. These models reveal a number of interesting trends, particularly when contrasted with the ranking-table results. Indeed, the models appear to diverge in some ways from the descriptive analysis based on the ranking tables. Based on the ranking tables, the top 20 routes with the most service complaints per average daily ridership served areas with a significantly higher income than those areas served by the top 20 routes ranked based on daily ridership or poor OTP performance. The models, on the other hand, reveal that an increase in the vulnerability index of the area surrounding the routes increased the odds of a service-related complaint, all else constant. Assuming that the population living along the bus routes also constitutes the ridership for those routes, this could suggest that feedback from disadvantaged populations is not suppressed.

The models also highlighted that bus routes with a higher proportion of trips that depart stops early increase the odds of a service-related complaint even more than a route with a high proportion of trips that depart stops late, all other variables held at their mean. Transit operators should keep in mind that bus riders could conceivably be more dissatisfied with missing a bus that they expected to make based on the operator's schedule than with other service issues, such as bus load or if a bus departs late. The analysis also suggests that bus routes that pass through areas with more vulnerable populations have higher odds of a receiving a reckless driving complaint. This may suggest that more vulnerable populations might be more sensitive to reckless driving than less vulnerable populations, all else held equal. Vulnerabilities are often co-occurring and it is not inconceivable that economic disadvantage may exist alongside physical disabilities, which could, in turn, create more sensitivity to the types of operating conditions associated with reckless-driving complaints, such as rough rides. The finding could also be indicative of physical conditions, such as poorer road infrastructure, that might stem from under investment in disadvantaged areas, though more specific research and analysis would be required to confirm this. Lastly, whether a bus route was a frequent service route also increased the odds of a service-related complaint, especially for a pass-up complaint, potentially indicating that riders have higher expectations for frequent service bus routes. This underscores the importance of linking bus performance measures with customer feedback when considering how to improve customer satisfaction.

Limitations of using customer complaints include that there is a lack of standardization regarding how they are recorded, collected, and stored in and across different public transport agencies. This is a barrier to using complaints data analyses to compare public transport agencies. Complaints data also requires significant data cleaning in order to use, which is another barrier to more widespread complaints data use in studying customer satisfaction for public transport. Along a similar vein, the date of the complaints used in this study may not have been the date of the actual incident that caused the complaint because of the way the complaints are collected. This is a potential source of error and limitation in linking complaints to bus service. The sociodemographic characteristics of the area near bus routes were used to link sociodemographic characteristics to complaints, rather than examining the characteristics of the person who submitted the complaint directly, which is another limitation of this study. Future research could include controlling for land uses in the vulnerability index calculation. For example, there could be parts of Portland, Oregon that have a high vulnerability index based on residential census data but also a high number of high paying jobs which could alter the population of bus riders in the area. This research used OTP as a measure of transit service quality and reliability, which is commonly used by transit agencies to understand bus service quality for frequent and infrequent bus service (Diab et al., 2015). Nevertheless, future research can explore the relationship between customer complaints and other reliability measures such as service variation. More research that explores the possible ways of using complaints data is still needed, which will help transit agencies in improving their image, which could result in retaining existing riders while attracting new riders.

6. REFERENCES

Abkowitz, M., Slavin, H., Waksman, R., Englisher, L., & Wilson, N. (1978). *Transit Service Reliability (No. UMTA MA-06-0049-78-1)*. Retrieved from Cambridge, MA:

Abou-Zeid, M., Witter, R., Bierlaire, M., Kaufmann, V., & Ben-Akiva, M. (2012). Happiness and travel mode switching: Findings from a Swiss public transportation experiment. *Transport Policy, 19*(1), 93-104.

Allen, J., Muñoz, J. C., & Ortúzar, J. d. D. (2019). On the effect of operational service attributes on transit satisfaction. *Transportation*.

Barabino, B., Francesco, M., & Mozzoni, S. (2017). An offline framework for the diagnosis of time reliability by automatic vehicle location data. *IEEE Transactions on Intelligent Transportation Systems, 18*(3), 583-594. doi:10.1109/TITS.2016.2581024

Barabino, B., Lai, C., Casari, C., Demontis, R., & Mozzoni, S. (2017). Rethinking transit time reliability by integrating automated vehicle location data, passenger patterns, and web tools. *IEEE Transactions on Intelligent Transportation Systems, 18*(4), 756-766. doi:10.1109/TITS.2016.2585342

Berkow, M., El-Geneidy, A., Bertini, R. L., & Crout, D. (2009). Beyond Generating Transit Performance Measures: Visualizations and Statistical Analysis with Historical Data. *Transportation Research Record: Journal of the Transportation Research Board, 2111*(1), 158-168.

Bertini, R. L., & El-Geneidy, A. (2003). Generating Transit Performance Measures with Archived Data. *Transportation Research Record: Journal of the Transportation Research Board, 1841*, 109-119.

Bree, S., Fuller, D., & Diab, E. (2020). Access to transit? Validating local transit accessibility measures using transit ridership. *Transportation Research Part A: Policy and Practice, 141*, 430-442. doi:https://doi.org/10.1016/j.tra.2020.09.019

Carleton, P. R., & Porter, J. D. (2018). A comparative analysis of the challenges in measuring transit equity: definitions, interpretations, and limitations. *Journal of Transport Geography, 72*, 64-75.

Carrel, A., Mishalani, R. G., Sengupta, R., & Walker, J. L. (2016). In pursuit of the happy transit rider: Dissecting satisfaction using daily surveys and tracking data. *Journal of Intelligent Transportation Systems, 20*(4), 345-362.

DeWeese, J., Hawa, L., Demyk, H., Davey, Z., Belikow, A., & El-Geneidy, A. (2020). A Tale of 40 Cities: A Preliminary Analysis of Equity Impacts of COVID-19 Service Adjustments across North America. *Transport Findings*.

Diab, E., Badami, M., & El-Geneidy, A. (2015). Bus transit service reliability and improvement strategies: Integrating the perspectives of passengers and transit agencies in North America. *Transport Reviews, 23*(3), 292-328.

Diab, E., van Lierop, D., & El-Geneidy, A. (2017). Recommending transit: Disentangling users' willingness to recommend transit and their intended continued use. *Travel Behaviour and Society, 6*, 1-9. doi:https://doi.org/10.1016/j.tbs.2016.03.001

Eboli, L., & Mazzulla, G. (2009). A new customer satisfaction index for evaluating transit service quality. *Journal of Public Transportation, 12*(3), 21-37.

El-Geneidy, A., Strathman, J., Kimpel, T., & Crout, D. (2006). Effects of bus stop consolidation on passenger activity and transit operations. *Transportation Research Record, 1971*(1), 32-41. doi:10.1177/0361198106197100104

Foth, N., Manaugh, K., & El-Geneidy, A. (2013). Towards equitable transit: examining transit accessibility and social need in Toronto, Canada, 1996–2006. *Journal of Transport Geography, 29*, 1-10.

Friman, M., & Fellesson, M. (2009). Service supply and customer satisfaction in public transportation: The quality paradox. *Journal of Public Transportation, 12*(4), 57-69.

Furth, P. (2000). *TCRP Synthesis 34: Data analysis for bus planning and monitoring*. Retrieved from Washington , DC:

Grengs, J. (2014). Nonwork Accessibility as a Social Equity Indicator. *International Journal of Sustainable Transportation, 9*(1), 1-14.

Harreman-Fernandes, M., Diab, E., Cui, B., DeWeese, J., Crumley, M., & El-Geneidy, A. (2021). Chapter 4: Transit customer satisfaction research: is the customer always right? In G. Currie (Ed.), *Handbook of Public Transport Research* (pp. 57–69). Cheltenham, United Kingdom: Edward Elgar Publishing.

Hartman, R. J., Kurtz, E. M., & Winn, A. B. (1994). *The Role of Performance-Based Measures in Allocating Funding for Transit Operations* (TCRP Synthesis 6). Retrieved from Washington DC: http://onlinepubs.trb.org/onlinepubs/tcrp/tsyn06.pdf

Hensher, D. A., Stopher, P., & Bullock, P. (2003). Service quality—developing a service quality index in the provision of commercial bus contracts. *Transportation Research Part A: Policy and Practice, 37*(6), 499-517.

Hounsell, N., Shrestha, B., & Wong, A. (2012). Data management and applications in a world-leading bus fleet. *Transportation Research Part C: Emerging Technologies, 22*, 76-87. doi:https://doi.org/10.1016/j.trc.2011.12.005

Javanmard, R., Lee, J., Kim, J., Liu, L., & Diab, E. (2023). The impacts of the modifiable areal unit problem (MAUP) on social equity analysis of public transit reliability. *Journal of Transport Geography, 106*, 103500. doi:https://doi.org/10.1016/j.jtrangeo.2022.103500

Kimpel, T., Strathman, J., Bertini, R., & Callas, S. (2005). Analysis of transit signal priority using archived TriMet bus dispatch system data. *Transportation Research Record, 1925*(1), 156-166. doi:10.1177/0361198105192500116

Kimpel, T., Strathman, J., & Callas, S. (2008). Improving scheduling through performance monitoring. In Computer-aided systems in public transport (pp. 253-280). . In M. Hickman, P. Mirchandani, & S. Voß (Eds.), *Computer-aided systems in public transport*. Berlin Heidelberg: Springer.

Kramer, A., & Goldstein, A. (2015). Meeting the Public's Need for Transit Options: Characteristics of Socially Equitable Transit Networks. *Institute of Transportation Engineers. ITE Journal, 85*(9), 23-30.

Ma, Z., Ferreira, L., & Mesbah, M. (2014). Measuring service reliability using automatic vehicle location data. *Mathematical Problems in Engineering, 2014*, 468563. doi:10.1155/2014/468563

Mandelzys, M., & Hellinga, B. (2010). Identifying causes of performance issues in bus schedule adherence with automatic vehicle location and passenger count data. *Transportation Research Record, 2143*(1), 9-15. doi:10.3141/2143-02

Moreira-Matias, L., Mendes-Moreira, J., de Sousa, J., & Gama, J. (2015). Improving mass transit operations by using AVL-based systems: A survey. *IEEE Transactions on Intelligent Transportation Systems, 16*(4), 1636-1653. doi:10.1109/TITS.2014.2376772

Mouwen, A. (2015). Drivers of customer satisfaction with public transport services. *Transportation Research Part A: Policy and Practice, 78*, 1-20.

Palm, M., Shalaby, A., & Farber, S. (2020). Social equity and bus on-time performance in Canada's largest city. *Transportation Research Record, 2674*(11), 329-342. doi:10.1177/0361198120944923

Radin, S. (2005). *Advanced public transportation systems deployment in the United States: Year 2004 update*. Retrieved from Washington DC:

Saskatoon Transit. (2021). *Annual report 2021*. Retrieved from Saskatoon: https://transit.saskatoon.ca/sites/default/files/documents/TC-TR_AnnualReport2021_WEB3.pdf

Schweiger, C. (2003). *Real-time bus arrival information systems: A synthesis of transit practice*. Retrieved from Washington, DC: https://onlinepubs.trb.org/onlinepubs/tcrp/tcrp_syn_48.pdf

STM. (2021). STM 2021 Annual Report : Allying agility and mobilization. Retrieved from https://www.stm.info/en/press/press-releases/2022/stm-2021-annual-report---allying-agility-and-mobilization

Strathman, J., Dueker, K., Kimpel, T., Gerhart, R., Turner, K., Taylor, P., . . . Hopper, J. (1999). Automated bus dispatching, operations control, and service reliability: Baseline analysis. *Transportation Research Record: Journal of the Transportation Research Board, 1666*(1), 28-36.

Strathman, J., Kimpel, T., Dueker, K., Gerhart, R., Turner, K., Griffin, D., & Callas, S. (2001). Bus transit operations control: review and an experiment involving Tri-Met's automated bus dispatching system. *Journal of Public Transportation, 4*(1), 1-26.

Surprenant-Legault, J., & El-Geneidy, A. (2011). Introduction of reserved bus lane:Impact on bus running time and on-time performance. *Transportation Research Record, 2218*(1), 10-18. doi:10.3141/2218-02

TriMet. (2023). Ridership and performance statistics: Monthly performance data, ridership trends and rider characteristics. Retrieved from https://trimet.org/about/performance.htm

TTC. (2023). Daily customer service report. Retrieved from https://www.ttc.ca/customer-service/daily-customer-service-report

Tyrinopoulos, Y., & Antoniou, C. (2008). Public transit user satisfaction: Variability and policy implications. *Transport Policy, 15*(4), 260-272.

van Lierop, D., Badami, M., & El-Geneidy, A. (2018). What Influences Satisfaction and Loyalty in Public Transport? A Review of the Literature. *Transport Reviews, 38*(1), 52-72.

van Lierop, D., & El-Geneidy, A. (2016). Enjoying loyalty: The relationship between service quality, customer satisfaction, and behavioral intentions in public transit. *Research in Transportation Economics, 59*, 50-59.

van Lierop, D., & El-Geneidy, A. (2017). Perceived reality: Understanding the relationship between customer perceptions and operational characteristics. *Transportation Research Record: Journal of the Transportation Research Board, 2625*, 87-97.

Verbich, D., Diab, E., & El-Geneidy, A. (2016). Have they bunched yet? An exploratory study of the impacts of bus bunching on dwell and running times. *Public Transport, 8*(2), 225-242. doi:10.1007/s12469-016-0126-y

Zhao, J., Webb, V., & Shah, P. (2014). Customer loyalty differences between captive and choice transit riders. *Transportation Research Record, 2415*(1), 80-88. doi:10.3141/2415-09