Volume 20, No. 1, 2017

David A. King
Juan Francisco Saldarriaga

Access to Taxicabs for Unbanked Households:
An Exploratory Analysis in New York City

Dea van Lierop
Ahmed El-Geneidy

A New Market Segmentation Approach:
Evidence from Two Canadian Cities

Seyed Amir Hosein Mortazavi
Meisam Akbarzadeh

A Framework for Measuring the Spatial Equity in the
Distribution of Public Transportation Benefits

Erkan Isikli
Nezir Aydin
Erkan Celik
Alev Taskin Gumus

Identifying Key Factors of Rail Transit Service Quality:
An Empirical Analysis for Istanbul

Orion T. Stewart
Anne Vernez Moudon
Brian E. Saelens

The Causal Effect of Bus Rapid Transit
on Changes in Transit Ridership

Xiaodong Liu
Yao Yang
Meng Meng
Andreas Rau

Impact of Different Bus Stop Designs
on Bus Operating Time Components

John Perry

Measuring the Accuracy of Bus Rapid Transit Forecasts

Inge Brechan

Effect of Price Reduction and Increased Service
Frequency on Public Transport Travel
The contents of this document reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation, University Research Institute Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.
CONTENTS

Access to Taxicabs for Unbanked Households: An Exploratory Analysis in New York City
David A. King, Juan Francisco Saldarriaga................................................................. 1

A New Market Segmentation Approach: Evidence from Two Canadian Cities
Dea van Lierop, Ahmed El-Geneidy.................................................................................20

A Framework for Measuring the Spatial Equity in the Distribution of Public Transportation Benefits
Seyed Amir Hosein Mortazavi, Meisam Akbarzadeh.......................................................44

Identifying Key Factors of Rail Transit Service Quality: An Empirical Analysis for Istanbul
Erkan Isikli, Nezir Aydin, Erkan Celik, Alev Taskin Gumus..................................................63

The Causal Effect of Bus Rapid Transit on Changes in Transit Ridership
Orion T. Stewart, Anne Vernez Moudon, Brian E. Saelens...................................................91

Impact of Different Bus Stop Designs on Bus Operating Time Components
Xiaodong Liu, Yao Yang, Meng Meng, Andreas Rau......................................................104

Measuring the Accuracy of Bus Rapid Transit Forecasts
John Perry..................................................................................................................119

Effect of Price Reduction and Increased Service Frequency on Public Transport Travel
Inge Brechan.............................................................................................................139
Access to Taxicabs for Unbanked Households: An Exploratory Analysis in New York City

David A. King
Arizona State University

Juan Francisco Saldarriaga
Columbia University

Abstract

Taxicabs are critical complements to public transit systems. In New York City, ubiquitous yellow cabs are as iconic as the city’s subway system, and the city recently added green taxicabs to improve taxi service in areas outside of the Central Business Districts and airports. In this paper, we used multiple datasets to explore taxicab fare payments by neighborhood and examine how paid taxicab fares are associated with use of conventional banking services. There are clear spatial dimensions of the propensity of riders to pay cash, and we found that both immigrant status and being “unbanked” are strong predictors of cash transactions. These results have implications for local regulations of the for-hire vehicle industry, particularly in the context of the rapid growth of services that require credit cards to use. At the very least, existing and new providers of transit services must consider access to mainstream financial products as part of their equity analyses.

Key words: Taxicabs, unbanked, immigrant populations, New York City

Introduction

Taxicabs represent an important transit service in urban areas, and the industry is undergoing rapid change. In recent years, new technologies developed by private firms have piqued substantial interest in growing the taxi industry from niche markets that complement transit systems to full-fledged alternatives auto ownership. Much of the current scholarly and popular interest in taxicabs focuses primarily on making taxicabs easier to use through smartphone-based e-hail applications and credit card payments. Although these technological innovations, have made taxi services—both
conventionally regulated taxicabs and upstart firms—easier to use for many travelers, these same innovations may make it harder for certain people to access taxis. In many US cities, large portions of low-income households do not have access to mainstream bank accounts or credit cards, which are required for smartphone apps. These unbanked or underbanked households are effectively excluded from new services, fare discounts for transit passes, and other transportation services that require access to credit cards.

A review of the literature suggests that transportation access is rarely addressed as an issue for underbanked households. Within the transportation literature, in contrast, ability to pay is generally considered a function of income or wages. Yet, in the case of the underbanked, ability to pay also includes access to fare payment media. In some studies, scholars have examined how the adoption of smart cards for transit fares may be affected by income, immigrant status, and other factors (Yoh et al. 2006). Within transit payments, low-income riders tend to not take advantage of volume discounts or unlimited fares, which likely is caused by their precarious financial straits. But all of these examinations assume that users are at least able to access transit services or other transportation facilities. In the case of private taxi and transit services, lack of a formal bank account and credit card (or branded pre-paid debit card) prohibit the use of these services, at least in the United States.

For this research, we use the definition of “unbanked” and “underbanked” of the US Federal Reserve (Gross et al. 2012). An unbanked household is one in which the head of household is without a checking, savings, or money market account, as is their spouse or partner. Underbanked people, in contrast, do have a checking, savings, or money market account but also use alternative financial services such as payday lending, title loans, or similar.

There are many reasons the unbanked and underbanked may opt out of or limit use of formal banking products. First, they might not have employment with regular paychecks. People who work odd jobs for cash may not need an account for savings. Second, they may have a regular job with a steady paycheck, but the fees charged for bank accounts with a debit card are too high for their income, or they receive most of their wages in cash tips. These people are likely to use check-cashing stores or pre-paid debit cards, and paying check cashing fees actually may be cheaper than using an ATM throughout the week. Third, immigrants—both legally in the country and illegally—are less likely to have formal bank accounts than native-born people. Together, low-income and immigrant status are associated with most of the unbanked and underbanked populations.

New York City has a higher share of unbanked households than the national average (Ratcliffe et al. 2015). In a 2012 report, the City’s Department of Consumer Affairs estimated that more than 10% of the adult population was without a bank account (New York City Department of Consumer Affairs 2012). The share of unbanked varies widely across the city, however, with nearly 30% of the population of the Bronx—the poorest borough—unbanked, whereas wealthier Staten Island has less than 2% unbanked. Moreover, nearly half of all unbanked live in just 10 neighborhoods, all
clustered in the poorest parts of the city, and happen to be places that traditionally have been underserved by taxicabs.

In the past few years, the City has launched a series of programs with the cooperation of financial institutions to increase access to mainstream services (New York City Department of Consumer Affairs 2008). These programs have had modest success for encouraging saving, even among very low-income people, and modest success moving people into mainstream accounts (New York City Department of Consumer Affairs 2013). Under Mayor de Blasio’s administration, the City has created a municipal identification card that does not require citizenship to acquire. This new ID card is hoped to assist at least some of the unbanked population to open new accounts. Yet, for all the City’s efforts, the evidence is mixed on the overall effectiveness of such “lifeline” services for promoting shifts into formal banking (Doyle et al. 1998).

Overall, the concern presented herein is that a particular aspect of poverty—whether or not a household has access to a formal bank account and, thus, potential access to credit cards—is critically important for access to a variety of transportation options. To the best of our knowledge, access to bank accounts and credit cards has been addressed in the literature only marginally, and not in the context of taxi services. This exploratory analysis used taxi data from New York City to identify spatial factors associated with the likelihood of being unbanked and cash fares for taxi trips.

Unbanked and Underbanked in New York City

Many households go between banked and unbanked depending on their circumstance. In general, there are a few factors strongly associated with being unbanked. The largest predictor of becoming unbanked is a steep decline in household income, followed by race and ethnicity factors, marital status, and housing characteristics (Rhine and Greene 2013). Most of these factors are found concentrated in particular neighborhoods, which suggests that households on the edge of poverty in certain communities will move in and out of the banking system as they can afford to.

The extent of underbanking recently has been recognized around the world, but both the diagnoses and remedies depend greatly on local context. A few generalized statements about underbanked households can be made. They are more likely to be poor, both by income and wealth, than households with bank accounts. In the literature, the primary concerns about the unbanked are usually about the high costs of being poor, especially as it relates to the cost of money. Check-cashing services can be more expensive than a savings account, for instance, as can be getting a money order to pay all bills. Recently, there has been some interest in pre-paid debit cards as a financial tool for low income families, but there are few examples of how these may work for transportation in the United States.

In New York City, studies show that physical proximity to conventional bank branches is unrelated to the likelihood of being unbanked (Ratcliffe et al. 2015). Throughout the city, retail bank branches are ubiquitous, although new bank branches are viewed as a sign of gentrification. Whereas not the focus of this paper, the neighborhoods with high
levels of unbanked households have mixed experience with gentrification, however it is defined.

Figure 1 shows the neighborhoods with the highest share of unbanked households along with subway stations throughout the city. This map is intended to show the spatial concentration of unbanked and underbanked households, which is why it is presented simply. There are approximately three main clusters of unbanked communities: northern Manhattan and the south Bronx, east-central Brooklyn, and Jamaica, Queens. The neighborhoods in Manhattan and the Bronx are the poorest neighborhoods in these boroughs, but they also have fairly good transit access by subway. Table 1 shows the actual share of unbanked households by neighborhood. Overall, the top 10 neighborhoods for share of unbanked and underbanked households represent about 450,000 unbanked people, or over half of all unbanked in the city.
The High Cost of Transportation

Poverty is a major urban policy concern. For much of the post-war period in the US, poverty was largely an inner-city phenomenon within metropolitan areas. One reason for concentrated poverty in the urban core was the availability of public transportation (Glaeser et al. 2008). Although poor, these households at least had access to transit networks that may allow for economic mobility, although our knowledge of how transportation affects poverty is limited (Sanchez et al. 2004, Sanchez 2008). In recent years, in part due to the Great Recession, poverty has suburbanized (Kneebone 2010), which has led to new concerns about the role of transit in suburban locations to prevent economic isolation for those who cannot afford to drive.

Poor households face a number of ways that reinforce how expensive it is to be poor. Inner-city neighborhoods pay higher retail prices (Talukdar 2008), for instance, or pay higher transit fares because they cannot take advantage of discounts. WNYC, a news radio station in New York, used data from the Metropolitan Transportation Authority (MTA) to demonstrate where riders purchase 7-day transit passes for $30 or unlimited transit passes for $112 per month (Schuerman 2015). The MTA data show that the 7-day passes are used more frequently than the unlimited passes, at 2.3 rides per day vs. 1.9. This means that the average fare paid is somewhat less for the typical 7-day pass holder—the higher usage means that these riders would receive substantial discounts simply by switching to a monthly unlimited pass. It is not known precisely why transit riders purchase 7-day passes when unlimited passes ultimately would save them money, but the most likely explanation is that the travelers simply do not have $112 to commit to transit trips at the beginning of each month. What these riders can do is buy a shorter pass when they are able and, if not, they do not travel or find other alternatives. This is a subtle example of how costs of living increase as income drops.

### TABLE 1.
Largest Share of Unbanked Households by Neighborhood

<table>
<thead>
<tr>
<th>Name</th>
<th>Borough</th>
<th>Unbanked</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mott Haven/Melrose</td>
<td>Bronx</td>
<td>56%</td>
</tr>
<tr>
<td>Morris Heights/University Heights</td>
<td>Bronx</td>
<td>53%</td>
</tr>
<tr>
<td>Highbridge/Concourse</td>
<td>Bronx</td>
<td>51%</td>
</tr>
<tr>
<td>Ocean Hill/Brownsville</td>
<td>Brooklyn</td>
<td>47%</td>
</tr>
<tr>
<td>Bushwick</td>
<td>Brooklyn</td>
<td>47%</td>
</tr>
<tr>
<td>Washington Heights/Inwood</td>
<td>Manhattan</td>
<td>46%</td>
</tr>
<tr>
<td>West Harlem</td>
<td>Manhattan</td>
<td>38%</td>
</tr>
<tr>
<td>East Harlem</td>
<td>Manhattan</td>
<td>37%</td>
</tr>
<tr>
<td>Central Harlem</td>
<td>Manhattan</td>
<td>36%</td>
</tr>
<tr>
<td>Jamaica</td>
<td>Queens</td>
<td>24%</td>
</tr>
</tbody>
</table>

Source: New York City Department of Consumer Affairs (2012)
Discrimination in Taxi Markets

Taxi markets are well-known for pockets of discrimination, especially with regard to people of color (Ambinder 1996; Anderson 2004). A key purpose of regulations is to ensure equitable access to taxicabs, which has led most cities to adopt a dispatch model of taxi service. In a dispatch model, all taxi rides are prearranged through a phone call or e-hail, in which the passenger calls a central dispatch hub and then waits for an assigned cab to arrive. The dispatch model is in contrast with the street hail model, in which all passengers are required to hail a cab on the street by raising their hand from the curb.

Dispatched taxicabs are favored because this model avoids two specific types of discrimination. First, by requiring all trips to be arranged through a central operator, taxi drivers must accept all trips within their licensed area. This prevents drivers from cruising only certain areas of a city, such as the Central Business District and airport, which offer more lucrative trips. It also provides a record of requests for rides made, regardless of their location, allowing regulators to ensure that no neighborhoods are systematically discriminated against. Second, dispatch services are nominally race-blind, such that the drivers are unable to target fares based on skin color. Although these protections do promote equal access to taxicabs across cities, in practice there are still major hurdles, and drivers do find ways to avoid certain neighborhoods.

In contrast to dispatch taxicabs, taxi drivers operating in a street-hail system often engage in more blatant discrimination where they simply will not stop their empty taxi for a person of color (Belcher and Brown 2012; Shuford 1999). This discrimination is very difficult to prove, however, as drivers claim they often did not see the person attempting to hail their services. Moreover, with street-hail services, drivers tend to avoid completely certain neighborhoods in which they feel unsafe or view as financially undesirable.

New e-hail technologies and services, such as Uber, Lyft, Juno and others, claim to avoid these known processes of discrimination. Being able to summon a taxi from your phone offers the convenience of a street hail with potentially race blind assurances that a taxi will accept the trip. However, as discussed in this paper, access to bank accounts is very much related to racial and immigration characteristics, which represents a new, albeit unintended, type of discrimination.

The Flexible Transit Market in New York

The New York City region is the nation’s largest transit market, with approximately one-third of all US transit riders, and the city is one of the few in the US that has seen consistent ridership growth over the past 15 years (APTA 2015). Beyond the well-known fixed-route services, New York also boasts many modes of flexible transit such as taxis, jitneys, and other types of for-hire services. These services are regulated by the Taxi and Limousine Commission (TLC), a City agency is run by a Commissioner appointed by the Mayor. The TLC’s primary responsibilities include licensing 50,000+ taxis, liveries, and commuter vans and about 100,000 drivers (TLC n.d.).
For-hire vehicles are a collection of distinct services; the most well-known, and part of the focus of this paper, are yellow medallion taxicabs. In the city are approximately 13,500 yellow taxis with medallions physically stamped onto their hoods\(^1\) that confer the exclusive right to pick up street-hail taxi fares in any part of the city. The service patterns of these taxis have been criticized for focusing on LaGuardia and JFK airports along with Manhattan’s central business districts rather than serving the city as a whole, a criticism not unwarranted, as airport travelers and short trips in the dense business districts have long been viewed as the most lucrative. Partially because of this, for years it was rare to see a yellow taxi on the streets of the outer boroughs (Brooklyn, Queens, Staten Island, the Bronx) or in communities of color. But this does not mean that these areas were not served; rather, they were served by a mix of informal and formal taxi services.

Neighborhoods outside of the Manhattan core have long relied on informal networks of community cars, livery vehicles, commuter vans, dollar vans, and other for-hire services. Each of these services tends to serve a particular niche, such as service between the city’s three distinct Chinatowns in Manhattan and Queens (Tsai 2010). Formalizing these services has been difficult for a number of reasons (King and Goldwyn 2014), one of which is that these services are used mostly by immigrants and low-income riders who always pay with cash.

In 2012, the City announced a program to increase the number of taxicabs outside the Manhattan core into traditionally-underserved neighborhoods. These new taxis, known as green cabs (because of their distinctive color) or borough cabs, cannot pick up passengers at the airports or in Manhattan south of 110th Street on the west side of Central Park or 96th Street on the east and are available as either a street hail or pre-arranged ride. The full effect of the green cab program is not yet known for overall taxi access or ridership as the program is still new, but preliminary data can be used to assess how trips made in green cabs differ from those made in yellow cabs.\(^2\)

The licensed taxicabs described above differ from livery licenses required for e-hailing services such as Uber, Lyft, and Via as well as longstanding livery companies that now have their own smart-phone applications, such as Carmel. These services are licensed by the TLC and must adhere to transparent regulations about insurance and safety but under livery guidelines that allow for more variation in vehicles and service standards. These licenses are unlimited in number and allow for trips anywhere within the city. However, these livery licenses require pre-arranging all trips (no street hails), which is satisfied through the smart-phone application.

\(^1\) An approximate number is used, as the true number of medallions has been affected by the growth of Uber and Lyft services. Some medallions have come out of service, but it is unknown exactly how many. For the purposes of this analysis, an approximate number of medallions is acceptable, as we are analyzing trips, for which we have all data.

\(^2\) Green cabs were introduced at the same time Uber, Lyft and other competitors entered the market. Since the growth of smartphone-enabled services, demand for green cabs has declined, and the City has not sold all available licenses.
Data and Methods

The data used in this analysis are from a recent TLC policy change. In 2004, the TLC initiated a program that required all taxicabs to use technology that allowed for credit card processing and also collected data about trip characteristics (King, Peters et al. 2012). This program, known as TPEP, was introduced in 2008. This research uses one month of geolocated trip data collected in October 2014 for all yellow and green taxicabs in New York City and provided by request from the TLC. The dataset includes trip origin, destination, time, number of passengers, fare paid, tolls paid, method of payment, tips (if paid by credit card) and other information. From the observed origins and destinations, distance traveled can be estimated but is not included in these analyses. These taxi data were combined with neighborhood level socio-economic data for analysis.

Table 2 shows the total trips by green and yellow taxis for the entire city during the study period. Yellow taxis make about 10 times the number of paid trips as green taxis. This is for many reasons, but primarily, the yellow taxis are used much more intensively and there are simply thousands more of them. Each yellow taxi typically is used for two 12-hour shifts daily, and medallion owners are eager to keep drivers in the cabs to make sure they collect rents of their assets. Green taxis, however, typically are owned by someone who drives part-time and leases the taxi for the balance of the week; thus, green taxis are used for more flexible shifts.

<table>
<thead>
<tr>
<th></th>
<th>Green Taxi Trips</th>
<th>Yellow Taxi Trips</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Trips</td>
<td>1,491,266</td>
<td>14,232,488</td>
</tr>
<tr>
<td>Cash Trips</td>
<td>820,747</td>
<td>5,684,248</td>
</tr>
<tr>
<td>Share of Trips Paid Cash</td>
<td>55%</td>
<td>40%</td>
</tr>
</tbody>
</table>

Source: New York City Taxi and Limousine Commission (n.d.)

The characteristics of trips by green and yellow are quite different. Most obviously, green taxis are prohibited from picking up passengers in many areas of the city, even though there may be high taxi demand relative to supply. More importantly, however, the data reveal that trip characteristics vary by location. A total of 55% of all green taxis trips—serving only outer boroughs by law—are cash fares. For yellow taxis, the likelihood of a cash fare is related to distance and whether the trip is an airport trip (these calculations are not shown).

Overall, there are observable differences for cash payments by taxi type, location, trip origin, and trip destination. It is impossible to know which characteristics differ between a typical yellow cab passenger and a typical green cab passenger, but something leads green cab passengers to use cash far more often than yellow cab passengers. The results shown on Figures 2 through 5 suggest that there is a spatial factor in play.

---

3 After the start of this research, the TLC has made all taxi GPS data available through its website; this was not the case when the research began.
FIGURE 2. Cash and credit payment types for green cabs by origin
FIGURE 3. Cash and credit payment types for yellow cabs by origin
FIGURE 4. Cash and credit payment types for green cabs by destination
FIGURE 5. Cash and credit payment types for yellow cabs by destination
In Figures 2 through 5, the relative frequency of payment types by origin and destination for yellow and green taxicabs. All maps show stark lines that demarcate where riders predominately use cash (shown in yellow) and where they use credit (shown in blue). With the exception of a credit card hotspot surrounding Columbia University in Morningside Heights (a blue area circled in Figure 3) Manhattan payment types divide cleanly along income lines, where wealthy neighborhoods flanking Central Park (the empty white rectangle in the middle of the map surrounded by blue to the south and yellow to the north) on the Upper West Side and Upper East Side pay for taxi trips mostly with credit cards, and poorer neighborhoods to the north in Spanish and Central Harlem are dominated by cash. One interesting aspect is that the socio-demographic characteristics of neighborhoods seemingly play a large role in determining payment type. It is likely that the cash or credit choice is a function of access to a bank account, for which these spatial data are a good proxy. Another takeaway is that much of the city still does not produce many taxi trips, and there is not enough data to present primary payment types.

**Statistical Analysis**

In this section, statistical analyses of the associations between socio-demographic characteristics and the share of cash fares for taxicabs are presented. The taxi trip data are limited in that they provide origins and destinations along with fare characteristics, but these data are associated with a known vehicle rather than a known passenger. As such, a series of assumptions can be made about a typical rider based on neighborhood factors. The regression analyses shown below used all trip data for the week of October 6–12, 2014, which is assumed to be a typical week in terms of good weather, no holidays, and no major school or employment breaks (n=3,217,092 for yellow taxi trips, n=330,024 for green taxi trips). These data are assumed to represent a close approximation of the average trip and, thus, the average trip-taker. Taxi trips were aggregated spatially to the neighborhood level, which were the smallest geographies with available demographic data, and then analyzed by trip origins and destinations for cash payments. Origins and destinations were treated separately, primarily because people leaving an area and returning to an area by taxicab may represent different groups of people.

Table 3 shows the summary statistics for dependent and independent variables considered for the regression models. These are shares of cash fares by origin (Ocash) and destination (Dcash) by neighborhood for all yellow and green taxi trips and are not mutually exclusive. Trips that begin and end in the same neighborhood were counted as both Ocash and Dcash. For most of the outer borough neighborhoods, the total number of intra-neighborhood trips was small and does not affect the overall results. For 2013, the percent of households in poverty headed by a foreign-born family member and unbanked are included. The unemployment rate in 2013 also was considered but ultimately was dropped from the analysis after post-test diagnostics.
Tables 4 and 5 show the regression results. The data are organized by neighborhood, and the dependent variable is either the share of cash trip by origin or cash trips by destination. Post-test diagnostics were used to evaluate multicollinearity, and the resulting models represent the best fit for the data. Ordinary least squares (OLS) was used along with generalized linear models (GLM), which accounts for the dependent variable not being normally distributed.

In all cases, the strongest predictors of cash fares are the share of foreign-born and the share of unbanked. These effects are largest for taxi trip destinations and are large and positive coefficients that are highly statistically significant; the share of households in poverty is not statistically significant. In both the OLS and GLM models, the direction of effects and approximate magnitudes are similar, suggesting that both models adequately represent the relationships among variables. The $R^2$ for the OLS models suggest that close to half of the variation of cash fares is explained by destination, which is a fairly high level of explanatory power for the model. It is likely that the reason poverty has an insignificant effect is that it is not a perfect predictor of banking status or immigrant status.
Discussion

Taxicabs and for-hire transportation services are premium services that complement fixed-route transit and supply critical accessibility to people who do not or cannot drive. Ensuring that these services are available to all who need them is a desirable policy goal. What the data in this research shows is that, in some cases, access to bank accounts and credit cards may affect access to certain types of taxi services. There are strong correlations between neighborhoods with high shares of unbanked households and taxi trips, especially green cabs, paid with cash.

These results underscore an important aspect of emerging taxicab technologies, which is that many supporters of expanding the taxicab supply base their support on the potential of new services to reach previously underserved markets. As potential can be refuted only through experience, existing firms in the taxi market look comparatively bad, as they have a history that can be checked. It is a common claim that smartphone-enabled taxi services will not employ the same geographic discrimination as conventional taxis because the drivers will respond to the service request. This may prove true at some point in the future, but many of the communities that need taxi services have high shares of unbanked households, who, by definition, cannot participate in a business that requires a credit card for access.

A scholarly example of this is a recent study by the BOTEC Analysis Corporation, in which researchers were sent into various neighborhoods to check response times and total trip costs for taxicabs and Uber drivers (Smart, Rowe et al. 2015). The study is methodologically sound, and the authors found quite conclusively that Uber cars arrive faster and cost quite a bit less, on average. But in the Los Angeles neighborhoods not well-served by taxis, households have very high rates of being unbanked (Khashadourian and Tom 2007). These households are in neighborhoods in which carpooling acts as taxi service and is far more prevalent than taxis (Liu and Painter 2011), and Uber cars are likely slower and more expensive than the taxi service actually used. It is possible that credit card-based taxi services simply are out of reach for many of these communities.

Washington Post writers collected data from Uber’s API and found that it offers faster service—measured by wait times after requests—to whiter and wealthier neighborhoods (Stark and Diakopoulos 2016). Such a claim is, by itself, not evidence of discrimination—and we want to be clear that is not part of our argument here—but taxicabs have long been subject to regulations, in part, to ensure access to service without regard to neighborhood, income, or race. Whereas a systematic review of tech-enabled taxi services is beyond the scope of this paper, the studies cited above are suggestive that there may be spatial differences in taxi access even with app-enabled hailing.

The green taxicabs in New York City also may have helped solve one problem—taxi access—but introduced another—decline of community cars. Community cars used to prowl the streets honking at prospective passengers, then a fare was negotiated for each trip. Although this practice was illegal, it was common. Through informal interviews with drivers and passengers of green cabs, some indicated they preferred the old system
of negotiated fares—the green taxis have the same fare schedule as the yellow taxis—because drivers would give breaks to certain people, while other paid higher fares. The poorest riders, who previously could have negotiated a trip for whatever cash they were willing to pay, now have to pay the meter fare, which often is higher. As these are not data collected systematically through interviews, the claims should be treated as speculation, but as anecdotes they are insightful observations about how at least a few very poor riders made use of taxi-type services with cash.

One shortcoming of the taxi GPS data used is there is no specific information about the passenger. We can only assume that high rates of unbanked households are related to high rates of cash payments. Although we feel this assumption is sound, the lack of passenger data limits its robustness and other analyses of taxi vehicle activities. It cannot be said for certain that a high share of unbanked households predict demand for cash payments for taxis, and this certainly requires additional surveys and passenger data. We also cannot evaluate these data for potential discrimination against passengers based on personal, locational, or payment characteristics. There may be unobserved discrimination that affects the results shown.

With the green cabs in New York, it is not clear that unbanked people are underserved by taxicabs. However, this does not mean that taxi regulations and transportation policy should not seek to protect vulnerable households. As the taxi industry goes through structural changes brought about by the rise of e-hailing applications, the City must consider ways to ensure access to all, not just those with a bank account.

Conclusions

This research presented an exploratory analysis of how taxi services in New York City exhibit market segmentation by fares payment methods. Overall, green cabs, which were designed to serve outer boroughs and underserved areas, disproportionately have cash fares. The yellow and green taxi markets exhibit some aspects of market segmentation in that yellow cab trips in unbanked areas are more like yellow cab trips elsewhere and green cab trips are more like community cars and likely serve different riders. The use of cash to pay for taxi trips is strongly associated with neighborhoods that have high shares of unbanked and immigrant households. Airports and central business district taxi trips are more likely to use credit cards, and these riders likely have different socio-economic characteristics than outer borough riders. Some potential implications from these findings are discussed above, but the key points are worth reiterating. Discrimination in the taxi market is a long-standing concern. Taxi drivers are infamous for avoiding certain types of people and certain neighborhoods, which is a key argument in favor of public regulation against discrimination. Such discrimination should not be tolerated. A concern based on the analysis in this paper is that limiting taxi services to those with a credit card also leaves many households unserved and may act as a new type of discrimination. Households on the edge of poverty go between having and not having bank accounts, and not having access to mainstream financial services may become a new type of discrimination without thoughtful policies.
References


New York City Department of Consumer Affairs. 2013. “Immigrant Financial Services Study.”


About the Authors

**David A. King** *(David.A.King@asu.edu)* is an Assistant Professor of Urban Planning in the School of Geographical Sciences and Urban Planning at Arizona State University.

**Juan Francisco Saldarriaga** *(juan.saldarriaga@columbia.edu)* is an Associate Research Scholar at the Center for Spatial Research at Columbia University.
A New Market Segmentation Approach: Evidence from Two Canadian Cities

Dea van Lierop and Ahmed El-Geneidy
McGill University

Abstract

Traditionally, transit market research has categorized passengers into two distinct groups: captive riders and choice riders. Market analyses that depend on such broad categories are likely to overlook important details about the needs and desires of their customer base. This study attempts to better understand the complexities of the different groups who take transit by using information from five years of customer satisfaction questionnaires collected by two Canadian transit providers. Employing a series of clustering techniques, the analysis reveals that nine market segments are present across different modes in both transit agencies. Three different overarching groups of transit users are identified based on income and vehicle access: choice users (~69%), captive users (~18%), and captive-by-choice users (~13%). The groups are consistent across transit modes and in different geographical regions and are generalizable enough to be widely applicable as a conceptual framework for segmenting and understanding public transit users.

Keywords: Transit market, market segmentation, captive user, choice user, mode choice

Introduction

Although transportation agencies and public policymakers have brought attention to the importance of increasing transit mode share, transit usage still lags significantly behind that of the car. Thus, to increase ridership, transit agencies and governments first need to understand what motivates individuals to use environmentally- and socially-sustainable forms of transportation such as public transit. Although much research attempts to elucidate what motivates drivers to switch to taking transit (Abou-Zeid et al. 2012; Curtis and Headicar 1997), fewer studies attempt to understand how to maintain and increase ridership among existing transit users. It is important for transit agencies to focus on retaining existing users, as it is known that individuals stop using transit for many reasons, including changes in income, family size, the availability of another mode,
A New Market Segmentation Approach: Evidence from Two Canadian Cities

as well as reasons related to the quality of service (Evans 2004; Grimsrud and El-Geneidy 2014; Perk et al. 2008).

One way to motivate existing users to remain loyal to the transit system is through increasing their satisfaction by taking into account their needs, perceptions, and desires with respect to transit. It is important to understand how to motivate loyalty in transit as it “involves a commitment on the part of the customer to make a sustained investment in an ongoing relationship with transit service” (Transportation Research Board 1999, 18). However, before developing strategies that attempt to increase satisfaction and loyalty among current transit users, it is beneficial to segment the market. Traditionally, transit market research has categorized riders into two distinct groups: captive riders and choice riders. Captive transit riders are commonly defined as individuals who do not have an alternative transportation choice; choice riders are those who choose to use transit even though another mode, such as a car, is available to them (Beimborn et al. 2003; Jacques et al. 2013; Krizek and El-Geneidy 2007; Wilson et al. 1984). Although it is important for transit agencies to acknowledge the presence of these two groups, analyses that depend on these broad categories are likely to overlook details about the needs and desires of their customer base. Therefore, rather than taking an approach to market segmentation that relies only on an analysis of whether or not transit users have access to alternative modes, the present study attempts to better understand the complexities of different groups who use transit. This is executed by using information about transit user socioeconomic status, personal preferences, and perceptions of satisfaction with transit services.

Nearly a decade ago, Krizek and El-Geneidy (2007) identified the habits and preferences of captive and choice transit users. Since then, transit markets have changed and new groups have emerged; Figure 1 demonstrates their conceptual framework. This study uses their transit market segmentation as a base on which to expand knowledge about transit user markets. The purpose of this study is to expand the left side of Krizek and El-Geneidy’s (2007) framework by assessing the different types of current transit users present in the two geographically-distinct Canadian cities of Montreal and Vancouver and update their transit market segmentation model.

This paper begins with a review of the relevant literature related to market segmentation. Next, based on an analysis of customer surveys collected by transit agencies in both cities over a five-year period, statistical clustering techniques are used to uncover market segments that are consistent in both geographic contexts. This is followed by a discussion of policy recommendations aimed at increasing ridership.
among the different clusters. In doing so, this paper illustrates how already-existing data can be used productively to inform public transit research, policy, and managerial practice.

**Literature Review**

**Market Segmentation**

Transit agencies are showing growing interest in understanding consumer behavior and have recognized that market-orientated research in public transit is likely to result in increases in user satisfaction (Molander et al. 2012; Transportation Research Board 1998a, 1998b). A first step toward identifying ways to increase customer satisfaction is to develop a market segmentation strategy to understand the needs and desires of the different groups using transit. Whereas market segmentation analysis can be a difficult task for practitioners (Palmer and Millier 2004), it can serve as a research base on which other marketing strategies can be built (Weinstein 2004).

Within the field of transportation planning, there have been a limited number of studies assessing transit market segments. One of the earliest examples of grouping types of transit users is the Transportation Research Board’s report on customer satisfaction (1999), which made suggestions for developing analyses that group current transit users as “secure,” “favorable,” “vulnerable,” and “at risk” to accordingly develop appropriate marketing strategies.

Several empirical studies have attempted to segment the transit markets in various regions (Anable 2005; Beirão and Cabral 2008; Jensen 1999). For example, Beirão and Cabral (2008) determined six unique traveler segments with different attitudes, demographic profiles, and intentions for using public transit in Porto, Portugal. Furthermore, Wilson et al. (1984) developed four market segments to account for variation in choice and captive riders, and McLaughlin and Boyle (1997) identified transit-dependent populations in Los Angeles County by segmenting based on car availability and income. Beimborn and Greenwald (2003) segmented the transportation market in Portland, Oregon, into what they call choice and captive riders based on mode preference and mode options. These authors recommended that transit agencies use these categories to improve forecasting and service design. Based on this study, Krizek and El-Geneidy (2007) evaluated the habit and preferences of users and non-users of transit to segment the market in the Minneapolis–St. Paul, Minnesota, metropolitan area. They found eight different segments of transit users and non-users including captive and choice users and recommended that policies should be based on an understanding of commuter attitudes and preferences, emphasizing that the retention of current riders is as important as the attraction of new ones. Jacques et al. (2013) took the concept of choice vs. captive riders further and found four segments that they claim are more representative of the market: “convenience,” which describes choice riders; “true captivity,” which describes captive riders; and “utilitarian” and “dedication,” which are neither clearly captive or choice riders. These authors suggested
that segments should not be viewed as static groups, but that individuals can move between categories.

Most of the abovementioned studies were derived from a sample of transit users or non-users residing within one region and were based on convenience samples. The present study segments the transit market to avoid analyzing heterogeneous groups within a transit market. It adds to the literature by using a segmentation technique that identifies context-specific clusters, and then groups the identified clusters based on income and car access. Therefore, this study provides a nuanced approach to understanding current transit users that is generalizable enough to be widely applicable as a conceptual framework for segmenting and understanding public transit users. The findings can provide transit agencies with information necessary to better understand the needs and desires of different groups within a transit market (Demby 1994; Peter and Olson 1999; Weinstein 2004).

Data

The data used for this study were obtained from two large public transit agencies in Canada: Montreal’s Société de transport de Montréal (STM) and Vancouver’s TransLink under a data sharing agreement to be used in academic research. In 2011, the population of the Montreal census metropolitan area (CMA) was 3.8 million with a transit mode share of 22.2% for work trips. In Vancouver, the CMA population was 2.3 million with 19.7% using transit for work trips (Statistics Canada 2014). The transit agencies in both cities provided the results of five years of customer satisfaction questionnaires that were conducted three or four times per year using telephone interviews. Telephone numbers were selected randomly, and respondents were filtered based on whether or not they use public transit. Only public transit users were interviewed and included in the sample. (Because participation was voluntary, non-response bias may be present.) In both Montreal and Vancouver, these routine questionnaires are intended to evaluate the quality of the transit service provided by the transit agencies and are used by the transit agencies to better understand perception of service quality and also as insight into where changes and/or improvements to service attributes could be accomplished to increase customer satisfaction and, accordingly, increase overall ridership.

To assess customer satisfaction with the transit service, the STM asks survey participants to report their experience with transit in general over the last 30 days. TransLink, however, takes a different approach by asking participants to specifically report their experience on their last and second-to-last trip. Although both strategies are appropriate for collecting information concerning customer satisfaction, the STM’s approach to asking about individual experiences in general may lack detail, whereas TransLink’s method of asking about the previous trip could result in capturing irregular travel, but it is likely negligible compared to those reporting regular travel behavior. In addition, both agencies ask questions regarding travel frequency, making it possible to distinguish frequent vs. infrequent users. Both agencies also assess transit user access to a car. Furthermore, because the questionnaires asked similar as well as several identical
questions, the differences in the method of the data collection were not problematic for this study; only data that were consistent between the two cities were included.

The STM provided information for a total 18,595 interviews, and TransLink for 42,061 interviews from 2009 to 2013. Not all questions were asked every year, and, therefore, inconsistent survey questions were removed from the database and not included in the analysis. The data were not weighted, as it would require having auxiliary information for all transit users in the regions, and also because the sample did not contain geographic information such as origin and destination points. However, the data are collected by the STM and TransLink in an attempt to collect representative random samples by ensuring that every transit user in each region with phone access has the same chance of being selected to be part of the survey following the basic rules of obtaining a representative random sample (Dunlop and Tamhane 2000).

Additional data cleaning was required to remove entries that were missing relevant information as well as apparent mistakes in the data. The surveys asked information including, but not limited to, transit user socioeconomic status, personal preferences, perception of satisfaction, and travel habits. Information about household structure and the presence of children was not included.

Satisfaction questions were asked using a 10-point Likert scale, and categorical data were converted to a series of dummy variables before being included in the analysis. Tables 1 and 2 list the questions that were used from the surveys from each transit agency. Data were then separated into three modal categories: bus, metro/SkyTrain, and the modes in combination. To clarify, bus users were individuals who reported using only the bus, metro/SkyTrain users were those who traveled only by rail, and individuals who used both modes represent those who reported using both modes in the same trip. The analysis was conducted for every distinct modal category to account for the differences in mode-specific service attributes. After data preparations were completed, a total of 14,842 observations were found suitable for the STM analysis and 29,224 for TransLink. This sample size at the 95% confidence level represents a confidence interval of 1.8% for transit users in Montreal and 1.3% for users in Vancouver. For the STM, the analysis yielded 7,190 bus users, 3,778 metro users, and 3,874 individuals who used both modes in combination. For Translink, the sample included 9,850 bus users, 6,604 SkyTrain users, and 12,770 who used both modes.
### TABLE 1. Factor Loadings: STM, Montreal

<table>
<thead>
<tr>
<th>Survey Questions</th>
<th>Bus</th>
<th>Metro</th>
<th>Both</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Car Access</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I use public transit because I don’t have a car.</td>
<td>-.904</td>
<td>.882</td>
<td>-.904</td>
</tr>
<tr>
<td>I currently have car access.</td>
<td>.531</td>
<td>-.650</td>
<td>.547</td>
</tr>
<tr>
<td>I use public transit because I don’t like driving/traffic.</td>
<td>.551</td>
<td></td>
<td>.540</td>
</tr>
<tr>
<td><strong>Financial Situation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>My income is greater than $80,000.</td>
<td></td>
<td>.664</td>
<td>.648</td>
</tr>
<tr>
<td>Status = work (compared to student, other)</td>
<td></td>
<td>.747</td>
<td>.774</td>
</tr>
<tr>
<td><strong>Life Phase</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What is your age?</td>
<td></td>
<td>-.854</td>
<td>-.810</td>
</tr>
<tr>
<td>Status = student (compared to work, other)</td>
<td></td>
<td>.882</td>
<td>.866</td>
</tr>
<tr>
<td><strong>Travel Day</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>When during the week do you take the bus most often? (mainly on the weekend)</td>
<td></td>
<td>-.766</td>
<td>-.807</td>
</tr>
<tr>
<td>When during the week do you take the bus most often? (mainly during the week)</td>
<td></td>
<td>.800</td>
<td>.790</td>
</tr>
<tr>
<td><strong>Loyalty</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I have been using STM public transit for at least one year as frequently as I do now.</td>
<td></td>
<td>.697</td>
<td></td>
</tr>
<tr>
<td>I plan to keep using the STM public transit network for a few or many more years.</td>
<td></td>
<td>.810</td>
<td>.741</td>
</tr>
<tr>
<td>Getting a new job, moving, or having a child would make me use public transit less in the next year.</td>
<td></td>
<td>-.709</td>
<td>-.732</td>
</tr>
<tr>
<td><strong>Frequency (Regularity)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am using STM public transit less than I used to.</td>
<td></td>
<td>-.594</td>
<td>-.692</td>
</tr>
<tr>
<td>In the last 30 days, what percentage of your trips would you say you made using public transit?</td>
<td></td>
<td>.734</td>
<td>.763</td>
</tr>
<tr>
<td>How many times did you take transit in the last 30 days?</td>
<td></td>
<td>.734</td>
<td>.736</td>
</tr>
<tr>
<td><strong>Convenience</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I use public transit because it is punctual/efficient.</td>
<td></td>
<td>.899</td>
<td>.851</td>
</tr>
<tr>
<td>I use public transit because I don’t like driving/traffic.</td>
<td></td>
<td></td>
<td>-.822</td>
</tr>
<tr>
<td><strong>Importance Of Low Costs</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I use public transit because of the low costs.</td>
<td></td>
<td>.964</td>
<td>.965</td>
</tr>
</tbody>
</table>
### Survey Questions

<table>
<thead>
<tr>
<th>Survey Questions</th>
<th>Bus</th>
<th>Metro</th>
<th>Both</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is your level of satisfaction with the cleanliness inside the bus/ metro cars?</td>
<td>.518</td>
<td>.831</td>
<td></td>
</tr>
<tr>
<td>What is your level of satisfaction with the cleanliness inside the metro stations?</td>
<td></td>
<td>.838</td>
<td></td>
</tr>
<tr>
<td>What is your level of agreement with the statement: “In the last month, the metro service on the lines that I used was reliable.”</td>
<td>.518</td>
<td>.539</td>
<td>.512</td>
</tr>
<tr>
<td>Last month, what was your level of security at any time you were on the bus or in metro installations?</td>
<td>.759</td>
<td>.541</td>
<td>.748</td>
</tr>
<tr>
<td>What is your level of satisfaction, out of 10, with the way in which drivers start, drive, and stop their buses on the STM bus routes that you use?</td>
<td>.795</td>
<td>.830</td>
<td></td>
</tr>
<tr>
<td>What is your agreement with the statement: “I feel that the driver drives carefully while respecting traffic regulations.”</td>
<td>.822</td>
<td>.842</td>
<td></td>
</tr>
</tbody>
</table>

### Satisfaction with Services

- What is your level of satisfaction with the cleanliness inside the bus?
  - .592
- What is your level of satisfaction with the cleanliness inside the metro stations?
  - .865
- What is your level of satisfaction with the cleanliness inside the metro cars?
  - .881

### Total variance (%)

- 65%
- 67%
- 68%

*Blanks show that the question had a factor loading of <0.5 or that it did not factor with the question group.*
### TABLE 2. Factor Loadings: TransLink, Vancouver

<table>
<thead>
<tr>
<th>Survey Questions</th>
<th>Bus</th>
<th>SkyTrain</th>
<th>Both</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Car Access</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I use public transit because I do not have a car (I have no choice).</td>
<td>-.715</td>
<td>-.772</td>
<td>-.748</td>
</tr>
<tr>
<td>Which of the following best describes your total household income before taxes?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Under $35,000)</td>
<td>-.513</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I use public transit because parking costs too much.</td>
<td>.666</td>
<td>.531</td>
<td>.713</td>
</tr>
<tr>
<td>Do you have access to a car, van or truck as a driver or passenger for the trips</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>you make using public transit? Yes</td>
<td>.726</td>
<td>.715</td>
<td>.718</td>
</tr>
<tr>
<td><strong>Financial Situation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Which of the following best describes your total household income before taxes?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(More than $75,000)</td>
<td>-.559</td>
<td>-.781</td>
<td>.677</td>
</tr>
<tr>
<td>Which of the following best describes your total household income before taxes?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Between $35,000–$75,000)</td>
<td>.920</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Which of the following best describes your total household income before taxes?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Under $35,000)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Life Phase</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What is your age?</td>
<td>-.821</td>
<td>.793</td>
<td>.800</td>
</tr>
<tr>
<td>What is the highest level of education you have completed? Some high school or less</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.614</td>
<td>-.510</td>
<td></td>
</tr>
<tr>
<td>What is your present employment status? “Student”</td>
<td>.806</td>
<td>-.807</td>
<td>-.820</td>
</tr>
<tr>
<td><strong>Travel Day</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did you make your last one way trip on Monday–Friday between 5–9:30am or</td>
<td>-.802</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monday–Friday between 3–6:30pm?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did you make your last one way trip on Saturday, Sunday or holiday?</td>
<td>.784</td>
<td></td>
<td>.809</td>
</tr>
<tr>
<td>Did you make your last one way trip on Monday–Friday between 5–9:30am or</td>
<td>-.829</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monday–Friday between 3–6:30pm?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did you make your last one way trip on Saturday, Sunday or holiday?</td>
<td>.835</td>
<td></td>
<td>.814</td>
</tr>
<tr>
<td><strong>Loyalty</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Compared to six months ago, would you say you are now riding transit more</td>
<td>-.805</td>
<td>-.803</td>
<td>-.789</td>
</tr>
<tr>
<td>regularly, less regularly, or about the same? (Less regularly than 6 months ago)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How likely are you to continue to take transit as often as you do now in the</td>
<td>.697</td>
<td>.705</td>
<td>.695</td>
</tr>
<tr>
<td>foreseeable future? (Probably or definitely continue as often as I do now)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Frequency (Regularity)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approximately how long have you been riding transit on a regular basis? (Number</td>
<td>.723</td>
<td>.854</td>
<td>.743</td>
</tr>
<tr>
<td>of years and months)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regular user (yes/no)</td>
<td>.817</td>
<td>.817</td>
<td>.800</td>
</tr>
<tr>
<td><strong>Convenience</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I use public transportation because it is reliable and because it has a good</td>
<td>.674</td>
<td>.883</td>
<td>.512</td>
</tr>
<tr>
<td>schedule.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I use public transit because of the convenience of the stops and stations.</td>
<td>.730</td>
<td></td>
<td>.761</td>
</tr>
</tbody>
</table>
### TABLE 2. (cont’d.) Factor Loadings: TransLink, Vancouver

<table>
<thead>
<tr>
<th>Survey Questions</th>
<th>Bus</th>
<th>SkyTrain</th>
<th>Both</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Low Costs</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I use public transit because it is cheaper.</td>
<td>.837</td>
<td>.715</td>
<td>.853</td>
</tr>
<tr>
<td>I use public transit because of the convenience of the stops and stations.</td>
<td></td>
<td></td>
<td>.539</td>
</tr>
<tr>
<td><strong>Satisfaction with Services 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How would you rate the bus for having a direct route?</td>
<td></td>
<td></td>
<td>.676</td>
</tr>
<tr>
<td>Trip duration from the time you boarded to the time you got off the bus?</td>
<td></td>
<td></td>
<td>.720</td>
</tr>
<tr>
<td>How would you rate it in terms of providing on time reliable service?</td>
<td></td>
<td></td>
<td>.744</td>
</tr>
<tr>
<td>How would you rate it in terms of frequency of service?</td>
<td></td>
<td></td>
<td>.797</td>
</tr>
<tr>
<td>Feeling safe from crime onboard the bus?</td>
<td></td>
<td></td>
<td>.556</td>
</tr>
<tr>
<td>How would you rate it for feeling safe from crime at the bus stop or transit exchange where you boarded?</td>
<td></td>
<td></td>
<td>.599</td>
</tr>
<tr>
<td>How would you rate it in terms of being clean and graffiti free?</td>
<td></td>
<td></td>
<td>.684</td>
</tr>
<tr>
<td>How would you rate that station in terms of safety?</td>
<td></td>
<td></td>
<td>.776</td>
</tr>
<tr>
<td>How would you rate your trip in terms of feeling safe from crime onboard SkyTrain?</td>
<td></td>
<td></td>
<td>.795</td>
</tr>
<tr>
<td><strong>Satisfaction with Services 2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Having a courteous bus operator?</td>
<td>.561</td>
<td>.608</td>
<td></td>
</tr>
<tr>
<td>How would you rate it in terms on being clean and graffiti free?</td>
<td></td>
<td></td>
<td>.617</td>
</tr>
<tr>
<td>How would you rate it for feeling safe from crime at the bus stop or transit exchange where you boarded?</td>
<td></td>
<td></td>
<td>.586</td>
</tr>
<tr>
<td>Feeling safe from crime onboard the bus?</td>
<td>.785</td>
<td></td>
<td></td>
</tr>
<tr>
<td>How would you rate the bus for having a direct route?</td>
<td>.830</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trip duration from the time you boarded to the time you got off the bus?</td>
<td></td>
<td></td>
<td>.682</td>
</tr>
<tr>
<td>How would you rate it in terms of frequency of service?</td>
<td></td>
<td></td>
<td>.752</td>
</tr>
<tr>
<td>How would you rate it in terms of providing on time reliable service?</td>
<td></td>
<td></td>
<td>.767</td>
</tr>
<tr>
<td><strong>Satisfaction (SkyTrain Only)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How would you rate it in terms of frequency of service?</td>
<td></td>
<td></td>
<td>.727</td>
</tr>
<tr>
<td>How would you rate it in terms of being clean and graffiti free?</td>
<td></td>
<td></td>
<td>.728</td>
</tr>
<tr>
<td>How would you rate it in terms of providing on time reliable service?</td>
<td></td>
<td></td>
<td>.766</td>
</tr>
<tr>
<td>How would you rate that station?</td>
<td></td>
<td></td>
<td>.786</td>
</tr>
<tr>
<td>How would you rate your trip in terms of feeling safe from crime onboard SkyTrain?</td>
<td></td>
<td></td>
<td>.807</td>
</tr>
<tr>
<td><strong>Total variance</strong></td>
<td>64%</td>
<td>65%</td>
<td>61%</td>
</tr>
</tbody>
</table>

*Blanks show that the question had a factor loading of <0.5 or that it did not factor with the question group*
Analysis

Principal Component Factor Analysis

Using SPSS 17, principal component analysis (factor analysis) was employed for each modal category to understand how survey questions related to each other. This statistical method considers the complete set of questions from the survey as well as their responses and creates a certain number of groupings (factors) that capture the variability in the data and therefore aids in reducing the number of variables analyzed (Doloreuxa and Shearmur 2013; Krizek and El-Geneidy 2007; Song and Knaap 2007). Using varimax rotation to maximize the variance of the squared loadings and Eigen values greater than one, this type of factor analysis was employed for each modal category within each agency: bus, metro/SkyTrain, and users who combined modes. Tables 1 and 2 demonstrate the results of the principal component analysis for the STM and TransLink and provide the factor loadings for the specific analysis of each modal category. These tables present the variables and corresponding survey questions used to build the components needed for the next phase of analysis. The numbers in the tables indicate the weight of each of the respective components; these factor loadings were grouped together when they were greater than 0.5 or less than -0.5.

Tables 1 and 2 show that the categories for each of the grouped questions were given titles that could be applied to both the STM and TransLink data, where possible. However, variation in the wording of specific questions was observed even though the questionnaires from both transit agencies assess individual socioeconomic profiles, travel behavior, opinions about transit, and perceived satisfaction of transit. Furthermore, questions that could not be grouped due to statistically insignificant factor loadings were removed from the analysis. The next phase of the analysis used the groups of questions, or factors, to define the market segments present in each transit agency.

K-means Cluster Analysis

Based on the results of the principal component analyses for each agency, k-means cluster analyses were performed using SPSS 17 with the factors developed for each modal category in both cities. This type of analysis is common in the literature and has proven to be a good method for segmentation (Damant-Sirois et al. 2014; Doloreuxa and Shearmur 2013; Jain 2010; Krizek and El-Geneidy 2007; Song and Knaap 2007). The factor scores that were generated for each variable included in Tables 1 and 2 were grouped together to identify segments of transit users for each modal category in both cities. In other words, the goal of the cluster analysis was to identify different groups of transit users within the existing customer base of the STM and TransLink by grouping riders with similar socioeconomic profiles, personal values, levels of satisfaction, and travel habits. The analysis maximized the differences between groups while minimizing the differences within groups. As the method used is an exploratory form of cluster
A New Market Segmentation Approach: Evidence from Two Canadian Cities

analysis, it was important to set criteria to determine how many clusters to retain. Although there are many approaches to judging the quality of segments (Dibb and Simkin 2010), because this analysis aims to update Krizek and El-Geneidy’s (2007) Transit Segmentation Model, we used the transit-specific criteria set by these authors to guide our decision:

- statistical output (cluster characteristics)
- relevance and transferability to transport policy
- previous studies
- common sense and intuition

Clustering was tried with three to eight groups, as suggested by Damant-Sirois et al. (2014), and final clusters of six and seven groups were found to provide the best qualitative descriptions for the groups using different modes in each city (Figures 2 and 3). These clusters are not specific to individual modes and named based on the prevalence of different factors. The sample size of each cluster is included below the name, and the bars represent each of the factors presented in Tables 1 and 2. Positive bar values represent that this factor was positively associated with the cluster, and vice versa. For example, “economizing riders” are labeled as such because they tend to use transit due to the associated cost savings. Although the figures demonstrate that most categories were consistent across modes, some differences exist. For example, Figure 2 shows that for every cluster of bus and bus and metro users, the first bar in every group is colored in light pink and represents access to a car. However, this bar is not included for the metro users; instead, metro user car access is determined by a white-colored factor, representing that a user does not have access to a car. The reason for the difference between “car access” and “no car access” is due to the results of the factor analysis represented in Table 1.
FIGURE 2. K-means cluster analysis for STM
A New Market Segmentation Approach: Evidence from Two Canadian Cities

FIGURE 3. K-means cluster analysis for TransLink
Similar to the results of Krizek and El-Geneidy’s (2007) segmentation analysis, Figures 2 and 3 demonstrate whether a cluster is categorized as a choice or captive users based on their income and access to a car:

- Choice users: Car access
- Captive users: No car access, low income

However, the results of the present study revealed that the data described more than choice and captive users, identifying a group of transit users present in the two cities that, to our knowledge, has not been previously identified in the literature. This new group was named “captive-by-choice” to reflect that they are captive to transit because they do not have access to a car but likely have chosen this situation, as they appear not to have as much of an income barrier compared to other clusters:

- Captive-by-choice users: No car access, do not have low income

Figures 2 and 3 use the terms “captive,” “choice,” and “captive-by-choice” to describe the clusters present among all modes. Finally, a description of the results of the cluster analysis is provided in Table 3.

### Table 3. STM and Translink Clusters

<table>
<thead>
<tr>
<th>Rider Type</th>
<th>Bus Users</th>
<th>Metro/SkyTrain Users</th>
<th>Bus and Metro/SkyTrain Users</th>
</tr>
</thead>
<tbody>
<tr>
<td>Service-driven riders</td>
<td>Have access to a car, do not have low incomes, are loyal, and travel during the week. Are not influenced by cost or convenience, satisfied with services. [S,T]</td>
<td>Have access to a car, do not have low incomes, and tend to be loyal. Are older, use the system occasionally, and are not influenced by cost or convenience, satisfied with services. [T]</td>
<td>Have access to a car, tend to be high income and loyal. Are older users who travel during the week, are not motivated by cost savings, and are satisfied with services. [S,T]</td>
</tr>
<tr>
<td>Economizing riders</td>
<td>Have access to a car, do not have a low income, and regularly commute during the week. Are largely motivated by cost savings. [S,T]</td>
<td>Have access to a car and regularly travel during the week. Tend to be loyal and are strongly motivated by cost. [S,T]</td>
<td>Have access to a car, and are regular loyal users who are motivated by cost savings. [S,T]</td>
</tr>
<tr>
<td>Convenience riders</td>
<td>Tend to be older, do not have high incomes, and travel during the week. Are loyal and very motivated by convenience. [S,T]</td>
<td>Are older, loyal, satisfied with services, and very motivated by convenience. Do not have access to a car. [S,T]</td>
<td>Tend to be older, loyal, satisfied with services, and motivated by convenience. Have high incomes and do not have access to a car. [T]</td>
</tr>
<tr>
<td>Weekend riders</td>
<td>Occasional users who primarily take transit on the weekend, have access to cars, and tend to be loyal. Are generally satisfied with services. [S,T]</td>
<td>Occasional users who primarily take transit on the weekend, have access to cars, and tend to be loyal. Tend to be older and high income and are generally satisfied with services. [S]</td>
<td>Occasional users who primarily take transit on the weekend. Are older and satisfied with services, but are not loyal or motivated by cost savings or convenience. [S,T]</td>
</tr>
<tr>
<td>Occasional weekday rider</td>
<td>Occasionally use transit during the week. Have car access, high incomes, tend to be older, and are motivated by convenience, but not by cost savings. Are satisfied with the services. [S]</td>
<td>Occasional use transit during the week. Have car access, high incomes, tend to be older, and are motivated by convenience, but not by cost savings. Are satisfied with the services. [S]</td>
<td>Occasional use transit during the week. Have car access, high incomes, tend to be older, and are satisfied with the services. [S,T]</td>
</tr>
<tr>
<td>Frustrated riders</td>
<td>Are unsatisfied with transit services, do not have access to a car, and are not medium income. Tend to be young and regular users who are loyal to the system and are not motivated by cost savings or convenience. [T]</td>
<td>Are unsatisfied with transit services and do not use them due to associated cost savings or convenience. Are older, regular users who have car access. [S,T]</td>
<td>Are unsatisfied with services and not motivated by cost savings. Are low-income, older, regular users who are loyal to the system. [S,T]</td>
</tr>
</tbody>
</table>
TABLE 3. (cont’d.) STM and Translink Clusters

<table>
<thead>
<tr>
<th>Rider Type</th>
<th>Bus Users</th>
<th>Metro/SkyTrain Users</th>
<th>Bus and Metro/SkyTrain Users</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disloyal riders</td>
<td>Are not loyal to the system, even though they do not have access to a car. Tend to be younger, do not have low incomes, and are not motivated by cost savings. Are somewhat satisfied with services. [S,T]</td>
<td>Are not loyal to the system and do not have access to a car. Use transit during the week, are not motivated by cost savings, but are slightly motivated by the convenience of transit. [S]</td>
<td>Are not loyal to the system and do not have access to a car. Are not motivated by cost savings or convenience, tend to be older, have higher incomes. [S,T]</td>
</tr>
<tr>
<td>Young riders</td>
<td>Tend to be younger and have lower incomes. Are loyal, use transit regularly, and are not motivated by cost savings. Are somewhat satisfied with services. [S,T]</td>
<td>Tend to be younger and have lower incomes. Are loyal, use transit regularly, and are not motivated by cost savings. Do not have access to a car and are somewhat satisfied with services. [S,T]</td>
<td>Tend to be younger and have lower incomes. Are loyal, use transit regularly, and are not motivated by cost savings. Do not have access to a car and are somewhat satisfied with services. [S,T]</td>
</tr>
<tr>
<td>Carless riders</td>
<td>Do not have access to a car, do not have high incomes, and tend to be loyal to transit. Are older, travel during the week, and are somewhat satisfied with services. [S,T]</td>
<td>Do not have access to a car, do not have high incomes, and are regular users who travel during the week. Are not motivated by cost or convenience, only somewhat satisfied with services. [T]</td>
<td>Do not have access to a car, do not have high incomes, and tend to be loyal to transit. Are regular users who are not motivated by cost savings or convenience. [S]</td>
</tr>
</tbody>
</table>

S = STM, T = TransLink

Discussion

Based on the findings from the cluster analyses presented in Figures 1 and 2, we were able to update Krizek and El-Geneidy’s (2007) transit market segmentation model to account for the different types of transit users that have been identified in the present study. Figure 4, accordingly, demonstrates that choice and captive users are not always separate entities, but instead overlap, showing that some individuals, in fact, chose to be captive.

The group that is captive-by-choice may have the financial ability to access another mode, but might have chosen to give up their cars because they prefer the experience of taking transit over that of driving. However, it is important to recognize that because information on household structure is not available to include in the analysis, individuals living in larger households have a higher chance of being financially-constrained compared to those in the same income bracket living with fewer family members. Therefore, some captive-by-choice users who have many members in their household may be more financially-constrained compared to captive-by-choice users.
who are financially-responsible for fewer household members. Similarly, not all choice riders will have the same transportation options available to them, and some, regardless of choice, may be more restricted to using public transit than others. Nevertheless, Figure 4 demonstrates that given these findings, the conceptual model makes clear that different groups of people can be accounted for within the broader categories of captive, choice, and captive-by-choice.

**A New Conceptual Transit Segmentation Model**

The new model presented in Figure 5 could serve as a tool for transit agencies wishing to develop marketing strategies to increase satisfaction and loyalty among many users. More specifically, this broader segmentation strategy can be used as a framework to better understand the urgency of developing policy interventions geared at the different groups using transit. Figure 5 adds to the new transit market segmentation model by taking it one step further to demonstrate the predictability of future usage of the different groups:

![Predictability of transit usage by group](image)

Figure 5 demonstrates that whereas choice users are likely to continue using transit in the long term, they may not choose to use it for all trips in the short term, as they have alternative modes available to them. Captive users, however, do not have access to alternative modes and, therefore, in the short term are predicted to use transit, but in the long term might gain access to a car or increase their income and, consequently, become captive-by-choice or choice users. Therefore, while at any given point it is likely that there will always be captive users, choice users, and captive-by-choice users, individuals will likely move between categories throughout the course of their lives. The goal of transit agencies should be to maximize the number of choice riders in a city while also working to better serve captive and captive-by-choice riders who have fewer modal options and, therefore, may also have more limited access to opportunities. The following paragraphs provide specific policy interventions aimed at inspiring users in different categories to continue using transit as they go through different life phases.
**Choice Users (~69%)**

In Vancouver and Montreal, choice users make up the majority of the transit market, and, therefore, it is important to motivate these users to continue using transit in the future. Service-driven riders represent the largest group, and, therefore, their needs should be prioritized. Economizing riders, however, represent another large group of transit users, and policies should be carefully developed to encourage this group to continue using transit. However, the needs and desires of weekend riders and occasional riders should not be overlooked, as service improvements geared specifically at this group may result in increased usage.

**Service-driven riders** often use transit because they are satisfied with the services and with the characteristics associated with their trips. To motivate these users to continue using transit, agencies should focus on maintaining the cleanliness and the safety of services (de Oña et al. 2013; Weinstein 2000), develop service improvements such as real-time travel information, and communicate transit investment and plans for service improvements (dell’Olio et al. 2011; Tyrinopoulos and Antoniou 2008).

**Economizing riders** often use transit because they benefit from the associated cost savings. Providing a low-cost transit service is associated with ridership (D’Alessandro and Des 2008; Hodge et al. 1994), and to positively impact individual perception of service and ultimately motivate their loyalty, transit agencies should communicate the cost saving benefits associated with using transit compared to other modes (Lai and Chen 2011). Agencies would also benefit from developing policies that encourage ridership through financial motivation (such as reduced fares). Increases in fares will likely have a negative influence on this group’s transit ridership and, therefore, must be carefully planned. Finally, cities can help motivate this group to continue using transit by developing policies that increase the price of driving and parking cars.

**Weekend riders and occasional riders** are grouped together, as they represent irregular users. Transit agencies should ensure that these users develop a positive perception of the system with regard to efficiency, travel time, and reliability (Carreira et al. 2014; Chou et al. 2014; de Oña et al. 2013). In the long term, transit agencies should focus on improving the common negative cultural image that is often attributed to transit (Schweitzer 2014). Transit’s cultural stigma can be changed by the implementation of policies that promote the service as being more comfortable and more efficient than using a private motorized vehicle (Chou and Kim 2009; Chou et al. 2014; Lee et al. 2009). Individual attitudes and preconceived ideas about public transit can be improved through policies that promote the aspects of transit that are unique to the service such as the ability for commuters to save time by working, reading, using the internet, or relaxing while they travel (Cain et al. 2009).

Although not all **frustrated riders** are choice riders, the majority fit into this overarching category. They are regular users who are not motivated by cost savings and are not satisfied with the services provided by public transit agencies. To satisfy these users, transit agencies should ensure that the system is clean, safe, and reliable (Burkhardt 2003; de Oña et al. 2013; Susilo and Cats 2014; Weinstein 2000). Additionally, these riders
would benefit from the implementation of an advanced information system to better communicate waiting times (Politis et al. 2010), route information, and connections to alternative modes such as bicycle share to increase the ease of usage of the entire transit system. Finally, it is important to note that although these riders have been categorized as choice riders, they may not have as many options as other choice riders, and therefore, although not low-income and having access to a car, could be restricted to using public transit. This is an area of research that should be explored in the future.

**Captive Users (~18%)**

Captive users are often carless riders and young riders, and transit agencies should take special care to cater to the needs of these groups to increase rider satisfaction in the present and not lose them in the future. Life-cycle changes (e.g., student to employment, renting to home-ownership, changes in family size and structure, etc.) often result in travel behavior changes (Evans 2004; Grimsrud and El-Geneidy 2014; Perk et al. 2008). Therefore, if captive users are not satisfied with the services provided by the transit agency, they may consider switching their mode when they increase their income due to a change in employment.

*Carless riders* use transit because they do not have access to a vehicle and do not have high incomes. Transit agencies must assess the needs and desires of this group and engage in equitable planning that recognizes that this group is strongly reliant on public transit (Stanley and Lucas 2008). In addition, transit agencies should provide the safest services possible for this group, as they do not have alternative options; depending on the context, safety provisions may include the installation of platform screen doors, additional lighting or surveillance cameras, and even security guards.

Although not all *young riders* are captive, this group tends to take transit because of their low incomes. Transit agencies should aim to improve how young transit users experience transit by developing technologically-current online customer feedback tools such as social media, web-based forums, and customer information mobile applications that can provide useful information for riders (Ferris et al. 2010). Furthermore, in the long term, agencies should be prepared to accommodate these uses as they go through lifestyle changes. This may include increasing convenience by increasing spatial and temporal coverage density.

**Captive-by-Choice Users (~13%)**

The identification of the captive-by-choice segment provides an important conceptual step from the car-as-norm paradigm that is often dominating transport research and policy. This newly-identified group appears to view public transit neither as a last resort when no options are available nor a mere complement to other transport modes. Alternatively, the existence of this group suggests that these users view transit as a viable transportation alternative on its own; in Vancouver and Montreal, it includes convenience riders and disloyal riders. Because these groups are likely to have the financial accessibility to switch modes, it is in the best interest of transit agencies
to develop a transit system that takes into account the needs and desires of these users. For example, for captive-by-choice users, public transit is likely to be in direct competition with car-share programs such as Car2Go and ride-share services and Transportation Network Companies (TNCs) such as Uber and the lower-cost UberX (Car2Go 2015; Rayle et al. 2014; Uber 2015).

Convenience riders generally take transit because they benefit from the convenience of this mode compared to other modes. Well-integrated services provided at and around transit stations are likely to attract these users. For example, in many regions, free wi-fi is now offered on trains and buses as well as stations to provide an additional service that appeals to younger generations. Such improvements are likely to increase overall levels of satisfaction for all users and attract irregular commuters to begin enjoying commuting by transit regularly. Furthermore, transit users tend to have a biased, distorted perception regarding transit travel time and waiting time, and they often report travel and waiting times that are longer than reality (Diab and El-Geneidy 2014). Correcting this distortion by using polices that improve the awareness of transit service qualities, as well as by implementing technologies such as next-arrival services, may help in increasing transit use (Garvill et al. 2003; Kenyon and Lyons 2003; Mishalani et al. 2006; Rose and Ampt 2001).

To increase loyalty among disloyal riders, transit agencies should communicate the benefits of using transit to these groups and focus on maintaining a safe, clean, and convenient system (Figler et al. 2011; Lai and Chen 2011; Minser and Webb 2010). However, transit agencies should also invest in better understanding the specific needs and desires of this group, as it is not clearly understood why these users are strongly disloyal.

**Conclusion**

This cluster analysis of two Canadian transit agencies links customer points of view to transit performance to bridge an existing gap in public transit segmentation research. The analysis has made clear that although different segments exist within each modal category, the overarching categories of captive, choice, and captive-by-choice are helpful to develop policy recommendations that reach further than policies directed at a single cluster. Because the findings are consistent in both the geographically-distinct settings of Montreal and Vancouver, this research is expected to be replicable and applicable in other cities. However, future research would benefit from applying and testing a similar segmentation analysis in other cities, especially in the US, where transit mode shares tend to be lower and the percentage of captive riders tends to be higher. Furthermore, although choice, captive, and captive-by-choice users are expected to be present in all transit markets, the percentage of users per group is expected to vary depending on the context. System improvements that are targeted at a specific segment could improve the experience of other groups as well, thereby motivate ridership among different users.
In addition to the findings of the analysis, this paper has also demonstrated how existing data from transit agencies can be used productively to inform public transit research, policy, and managerial practice. In the future, to further help in the development of policies that aim to retain and/or increase transit ridership, research should include in-depth analyses focused on understanding the needs and desires of the different market segments and set out to better understand how to motivate non-users to use public transit.

**Acknowledgment**

The authors thank Brigitte Vallée and Isabelle Ouellet from the STM and Victor Gaspar from TransLink for providing access to their customer service questionnaire databases. We also thank the three reviewers for their feedback on the earlier version of the manuscript. This research is funded by the Natural Science and Engineering Research Council (NSERC) and the Social Sciences and Humanities Research Council (SSHRC) of Canada. The ideas and findings presented in this paper represent the authors' views in an academic exercise.

**References**


A New Market Segmentation Approach: Evidence from Two Canadian Cities


About the Authors

**Dea van Lierop** ([dea.vanlierop@mcgill.ca](mailto:dea.vanlierop@mcgill.ca)) is a Ph.D. Candidate in the School of Urban Planning at McGill University in Montreal. Her research interests include public transit satisfaction and loyalty, land use and transportation planning, pedestrian and cyclist safety, and innovations in transit-oriented development.

**Ahmed El-Geneidy** ([ahmed.elgeneidy@mcgill.ca](mailto:ahmed.elgeneidy@mcgill.ca)) is an Associate Professor at McGill University in Quebec, Canada. His research interests include land use and transportation planning, transit operations and planning, travel behavior analysis including both motorized (auto and transit) and non-motorized (bicycle and pedestrian) modes of transportation, travel behavior of disadvantaged populations (seniors and people with disabilities), and measurements of accessibility and mobility in urban contexts.
A Framework for Measuring the Spatial Equity in the Distribution of Public Transportation Benefits

Seyed Amir Hosein Mortazavi and Meisam Akbarzadeh
Isfahan University of Technology, Iran

Abstract

This paper proposes that an equitable transit system requires that the geographical distribution of transit service benefits conform to the geographical distribution of the citizens with the greatest need for public transportation. This is the essence of vertical equity. This study calculated “connectivity power,” which reflects public transit service quality in each traffic analysis zone (TAZ) in a city to indicate the amount of benefit that TAZ is receiving from the transit system. The number of carless citizens in each TAZ was also calculated as an index of need to the public transit services in that area. Conformity of need and supply was analyzed using Spearman’s rank correlation and the Gini index. This framework indicates that in Isfahan, Iran, adding three bus rapid transit (BRT) lines and completing the existing metro line would not improve the Gini index but would improve the Spearman’s rank correlation to an acceptable level.

Keywords: Equity, traffic analysis zone, Gini index, Spearman’s Rank Correlation.

Introduction

Public transportation has certain merits, including high productivity of inputs (e.g., fuel and urban infrastructure capacity) and compliance with sustainable development. Therefore, promoting public transportation is one of the priorities of modern urban management. To evaluate transportation equity, it is important to measure the distribution of public transit service benefits and costs. This requires a rigorous theoretical and computational framework so that urban management is able to evaluate the current status as well as the consequences of future plans. This paper proposes a framework for evaluating the equity in the distribution of the public transportation benefit and applies it to the public transportation system of Isfahan, Iran. Various studies throughout the world have been carried out and reported in the literature, including Ricciardi et al. (2015), Griffin et al. (2015), and Iseki (2016).
From the point of view of passengers, the benefits of public transportation include mobility (operational speed) and accessibility (access to stations, access the intended destination, waiting and transfer convenience). Various public transportation benefits as well as a variety of equity definitions and the difficulty of matching executive plans to definitions have led to a kind of obscurity in both theoretical and practical aspects of public transportation equity (El-Geneidy et al. 2013). To avoid this obscurity, a clear definition of equity and an exact realm of factors of interest for evaluation are discussed in this paper.

In terms of the availability of travel alternatives, citizens can be divided into two groups—“choice” and “captive” transit users. Choice users have at least two motorized alternative modes to accomplish their urban trips. Usually, this implies that choice users have a car at their disposal and also can select a transit or paratransit mode. On the other hand, captive users have only one motorized mode for their intra-urban trips, which can be due to the their lack of ownership or access to car and/or financial, health, or age situation or travel conditions.

Most captive users are lower-income (Welch 2013; Manaugh and El-Geneidy 2012). Consequently, a key function of a public transportation system is filling the utilization gap among different income classes. This function is carried out by providing access to destinations (especially to job opportunities) for captive users. The importance of the public transportation system is to the extent that, in some research, more suitable accessibility is considered a reason for gathering low-income households in neighborhoods near downtown (Kahn et al. 2008).

Spatial mismatch theory perceives the geographical distance between citizen residence and their available job opportunities as an urban issue. Transportation mismatch theory studies the mismatch among transportation infrastructure (including highways and public transportation facilities) with citizen job opportunity access demand (Ong and Miller 2005). Accordingly, covering urban expansion caused physical distance between origin and destination using public transportation is a progressive challenge of low-income citizens.

Various approaches for quantification of the equity concept are proposed in the literature. Litman (2002) gives a thorough review of these definitions. Horizontal and vertical equities are two main approaches in this regard. Equal distribution of benefits among all social classes is known as horizontal equity. Vertical equity in public transportation requires distribution of benefits according to the need of each social class to those services (Murray and Davis 2001). In this definition, cost of service and the purchasing power of citizens are not taken into account.

Martins et al. (2012) considered the social meaning of a good as its equitable distribution criterion. For essential goods (e.g., bread, water), equity would be served by a uniform distribution. However, for luxury goods, due to their social meaning and usage by specific social groups, supply zone distribution should be determined by demand market norms. Social meaning of public transportation service is not identical for all citizens. Public transportation is considered an essential good for captive users.
but only an option for choice users. Therefore, to consider the citizen dependence factor, it is better to distribute public transportation services based on vertical equity.

Social equity in public transportation has been the focus of several research papers. As an example of recent endeavors, Currie (2010) studied the gap in transit supply based on social needs. He developed indexes for supply and demand and evaluated the conformity of the supply and need throughout various districts of Melbourne, Australia. The supply index included the number of stations within a zone, the frequency of service in each station, and the ratio of covered area of a zone to its total area. The demand index was a linear function of various characteristics of a zone, including the number of adults without a car, persons over age 60, students, etc. The supply and need were plotted against each other, and regions with low conformity of supply and need were detected.

Based on supply-demand analysis using the Gini coefficient, Bertolaccini and Lownes (2013) investigated the effects of scale and boundary selection in assessing the equity of transit supply distribution. They calculated Gini coefficients for six urban transit systems within the United States within two boundaries (Metropolitan Statistical Area and Transit Service Area) at two scales (Census Tract and Block Group) and adopted two different demand measures (population and population plus employment). The results suggest that calculations with Gini coefficients on the basis of various boundary definitions can lead to significantly different comparative results, but the different scales and demand measures had insignificant impact on interregional comparisons.

Lemans (2016) adopted the Gini index, Theil index, and descriptive statistics to measure the equity of accessibility in an area of Utrecht, Netherlands. The descriptive statistics included the mean, variance, range, and coefficient of variation of the value of accessibility throughout the study area. Two situations (current as of 2015 and future as of 2020) were considered in the evaluation. A noticeable conclusion of the research was that the Gini index outperforms the Theil index and the descriptive statistics. The reason is that the Theil index used the average accessibility per resident in a zone, whereas the accessibility of each zone was determined by summing the contribution of stops located in it. The disadvantage of the descriptive statistics was the high sensitivity of the obtained scores to outliers in the study area.

**Methodology**

The methodology adopted in this paper is illustrated in Figure 1. In the first step, a clear definition of equity must be selected. As mentioned in the previous section, vertical equity was selected as the definition fitting the topic of public transportation benefits distribution. The definition of vertical equity has three keywords—benefit, need, and distribution. In the second step, measures were developed to pinpoint these keywords and quantify their values.
As the trip and socio-economic data are usually available for traffic analysis zones (TAZs), TAZs were selected as the geographical units for evaluating the distribution of benefit and need.

Benefit was formulated as “connectivity power,” which is a function of accessibility and mobility similar to (but not the same as) the formulation of Welch et al. (2013) and Kaplan et al. (2014). In public transportation systems, stations bridge demand with supply. Connectivity power of a station (transit service quality index) shows how well it is connecting the demand to the whole network and the urban area. The connectivity power of station \( n \) brought about by a line \( (P_{l,n}) \) is the product of vehicle capacity \( (C_l) \), frequency \( (f_l) \), service hours \( (H_l) \), speed \( (V_l) \), and the number of stations \( (D_l) \) of that line. The relation is presented in equation (1).

\[
P_{l,n} = C_l \cdot f_l \cdot H_l \cdot D_l \cdot V_l
\]

(1)
Total power of a station \( P_n \) equals the summation of power of the lines passing through it.

\[
P_n = \sum_{l,n \in l} P_{l,n} \quad (2)
\]

To calculate the amount of benefit a public transportation system is providing to a TAZ, one only needs to add up the power of all the stations within that TAZ.

Need in each TAZ was estimated by the population of captives, i.e., citizens without car. In each TAZ \( i \), number of transit-dependent residents or simply the index for need \( N_i \) is equal the product of the population \( P_i \) and the car ownership \( C_i \). Therefore:

\[
N_i = P_i \times (1 - C_i) \quad (3)
\]

Car ownership rate in each TAZ was obtained from the comprehensive survey recently conducted in Isfahan.

Once the need for public transportation service and provided benefits in each TAZ are determined, instruments are needed to evaluate the distribution of need and benefit throughout the urban area. The first step taken in this regard was to assess the conformity among the rank of TAZs according to need and benefit. This framework assumes that the distribution of transit services reflects local needs, so areas that have greater needs receive greater services. To assess conformity among the rank of TAZs in need and benefit lists, the list of TAZs was sorted according to need, and the rank of each TAZ was determined. Then, the list of TAZs was sorted according to benefit, and the rank of each TAZ was determined. Then, Spearman’s correlation coefficient was calculated, as shown in equation (4).

\[
r_s = 1 - \frac{6 \times \sum_{i=1}^{m} d_i^2}{m(m^3 - m)} \quad (4)
\]

Where \( d_i \) indicates the difference of ranking of a zone in the two directories and \( m \) indicates the number of zones. \( r_s \) follows Student’s \( t \) distribution with degrees of freedom equal to \( m - 2 \) and, therefore, its statistical significance can be determined (Siegel 1956).

In addition to assessing the conformity of orders of need and benefit, overall distribution of benefits among zones was analyzed by calculating the Gini index. From an economic perspective, complete equality means a situation in which all process gains are equally distributed among community members (Welch 2013). A profile of this status is the first quadrant coordinate system.

The Gini coefficient indicates the distance of the current situation from the ideal situation, which means complete equality. In fact, the Gini index illustrates the surface between two diagrams of complete equality and the current situation. The mathematical definition of Gini index is as follows:

\[
G_a = 1 - \sum_{k=1}^{n} (X_k - X_{k-1})(Y_k + Y_{k-1}) \quad (5)
\]

In complete equity, the value of this index is zero and in the worst condition is 1.
Case Study

Isfahan has a population of 1.7 million and an area of approximately 500 square kilometers and is located in central Iran. Isfahan’s bus system consists of 92 lines and 1,753 active stations, including one active bus rapid transit (BRT) line with 32 stations and one subway line (the first phase of Line 1) with 10 active stations. It is expected that within two years, three additional BRT lines will be added to the public transportation system of the city. Line 1 of the Isfahan BRT was launched in 2012; up until then, there was only regular bus line throughout the city. Figure 2 depicts the evolution of the Isfahan transit network since 2012 to 2018. The dark blue line shows the part of metro line currently in use, the green line shows the current BRT line, the light blue lines show three future BRT lines, and the red line shows the future extension of the metro line of Isfahan.

Data on TAZ boundaries, their population, and car ownership came from a comprehensive transportation study conducted by the Isfahan municipality. Data on public transportation services were obtained from the Isfahan Metropolitan Transit Organization.
Based on the proposed framework, the condition in the base year (2012), the current condition (2016), and the future condition (2018) after completion of all BRT lines and the metro line were viewed and compared from the equity point of view. Data on supply side were collected from the Isfahan Metropolitan Bus Company.

Figures 3 through 5 illustrate the distribution of station power throughout the central part of Isfahan in three conditions.

**FIGURE 3.**
Distribution of station power in base condition
A Framework for Measuring the Spatial Equity in the Distribution of Public Transportation Benefits

FIGURE 4.
Distribution of station power in current state
To calculate the amount of need in each traffic zone, the number of residents without a private vehicle was taken into consideration. This group was expected to be totally dependent on public transportation service. The cause of information aggregation in the scale of traffic zones was the availability of data. This information was collected through surveys conducted in comprehensive studies of Isfahan public transportation. To predict the population in the future, the average national population growth rate of Iran was used.

Figure 6 depicts the distribution of need throughout the city. Zones are divided into 10 deciles; the first decile (darkest color) includes zones with the highest need and the tenth decile (lightest color) includes zones with the lowest need.

To calculate the amount of supply of public transportation service in each traffic zone, the power of stations within each traffic zone was summed. The results of the base year, current situation, and future year are depicted in Figures 7, 8, and 9, respectively. To facilitate the display, traffic zones are divided into 10 deciles based on power, where the first decile has the best and the tenth decile has the worst condition in terms of power.
FIGURE 6.
Distribution of need
FIGURE 7.
Distribution of power in base situation
FIGURE 8.
Distribution of power in current situation
After calculating the values in all TAZs under all three time steps, the values of “need” were pooled into one group and the values of “supply” were pooled separately. Then, each group was sorted and divided into 10 deciles. Therefore, the scale of all the maps pertaining to “need” is the same and the scale of all the maps pertaining to “supply” is the same.

Intensity of deficit in each zone \( (B_i) \) was calculated by dividing the normalized value of need \( (N) \) to the normalized value of power \( (P) \). Normalization was done by dividing the values by their maximum, i.e., \( \frac{P_{\text{max}}}{P_{\text{max}}} \) and \( \frac{N_{\text{max}}}{N_{\text{max}}} \).

\[
B_i = \left( \frac{N_i}{N_{\text{max}}} \right) / \left( \frac{P_i}{P_{\text{max}}} \right)
\]  

(8)
The values of $B_i$ for each traffic zone are separated per decile and depicted in Figures 10, 11, and 12. Zones with higher disparity among need and power are colored and the first decile zones (lightest) have the most suitable public transportation services compared to the need of residents.
FIGURE 11.
Distribution of need intensity in current situation
Social Equity Assessment

To assess equity, the Spearman’s rank correlation and the Gini index were calculated. The Spearman’s rank correlation value indicates the ranking imbalance between zone power and need. Also, the magnitude of the Gini index indicates the distance of the situation from full equity (perfect equity) condition. Table 1 shows the value of these coefficients in base, current, and future situations.

<table>
<thead>
<tr>
<th></th>
<th>Spearman’s Rank Correlation</th>
<th>Corresponding t Statistic</th>
<th>Gini Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base condition</td>
<td>0.18</td>
<td>2.4</td>
<td>0.31</td>
</tr>
<tr>
<td>Current condition</td>
<td>0.18</td>
<td>2.4</td>
<td>0.33</td>
</tr>
<tr>
<td>Future condition</td>
<td>0.04</td>
<td>0.5</td>
<td>0.33</td>
</tr>
</tbody>
</table>
In the base situation, the Spearman’s rank correlation for Isfahan is 0.18 and its corresponding t statistic is 2.4. The acceptable limit for this statistic with 184 degrees of freedom (number of zones minus two) is approximately 1.7. Therefore, the difference between Isfahan zone need rank and received power rank was statistically significant. Consequently, distribution was not proportional. The Gini index in the base situation for Isfahan was 0.31. Figure 13 illustrates the Lorenz curve and the bisector line (full equity) for this situation.

In the current situation, the Spearman’s rank correlation coefficient for Isfahan is 0.18 and its corresponding statistic is 2.4. The critical value for this statistic with 184 degrees of freedom is about 1.7. Therefore, Isfahan zone ranking in term of need is not statistically proportional with its ranking in term of received power. The Gini index pertaining to the current situation of Isfahan is 0.33, which does not show much improvement compared to the base year.

In the future situation, the Spearman’s rank correlation for Isfahan is 0.036 and its corresponding statistic is 0.49. The acceptable limit for this statistic with 184 degrees of freedom is about 1.7. Therefore, Isfahan zone ranking in term of need is statistically proportional with its ranking in terms of received power. Consequently, power and need distribution will be in adequate conformity. The Gini index pertaining to Isfahan will be 0.33, which does not show much difference compared to the base and current years.

**Conclusion**

The main idea of this paper was that in an equitable system, the geographic distribution of benefits throughout the city should conform to the geographical distribution of resident needs. This is the essence of vertical equity. This study calculated the connectivity power of Isfahan bus stations by incorporating the service hours, frequency, speed, vehicle capacity, and number of accessible stations. The total power presented in each TAZ was calculated as the amount of benefit that TAZ is receiving
from the transit system. On the other hand, the number of carless citizens in each TAZ was calculated as an index of need to the public transit services in the TAZ.

In the second stage, three conditions of Isfahan were investigated in terms of equity. Until 2012, Isfahan had a transit network consisting of ordinary bus routes running under right-of-way C. Since then, BRT was introduced to the system, and part of the metro system started working. It is planned to complete three more BRT lines and the north-south metro line by 2018. To assess the conformity of demand and supply (i.e., equity), the Spearman’s rank correlation and Gini index were calculated throughout the TAZs. First, TAZs were sorted in two separate lists based on their need and power. Then, the rank of each TAZ in two lists was analyzed to determine if the same order was preserved in the two lists. Second, the Gini index was calculated for the whole city under each condition. Results show that although new developments have not contributed to the Gini index (as it remains equal to 0.33 in all three conditions), conformity of demand and supply will reach an acceptable level by 2018 as a consequence of launching new BRT and metro lines.

The proposed framework could be applied to all urban contexts. It may also be used to prioritize tentative scenarios in public transportation of an urban area. As an example for the case of Isfahan, diagonal or radial lines connecting the northwestern part of Isfahan to its southeastern part may serve the zones with highest ratio of need to power and improve the equity status in the city. Including speed, the number of upstream and downstream stations made the calculations on the supply side more realistic.

References


About the Authors

SEYED AMIR HOSEIN MORTAZAVI (amirh_mortazavi@yahoo.com) is an MSc candidate in the Department of Transportation Engineering at Isfahan University of Technology.

MEISAM AKBARZADEH (makbarzadeh@cc.iut.ac.ir) is an Assistant Professor in the Department of Transportation Engineering at Isfahan University of Technology. His research areas of interest include public transportation and complex transportation networks.
Identifying Key Factors of Rail Transit Service Quality: An Empirical Analysis for Istanbul

Erkan Isikli
Istanbul Technical University

Nezir Aydin
Yildiz Technical University

Erkan Celik
Munzur University

Alev Taskin Gumus
Yildiz Technical University

Abstract

Providing a high quality of service in public transportation is essential to reduce dissatisfactions stemming from traffic congestion and noise. Public transport providers need to find ways to dilute the effects of immoderate use of private cars in big cities while maintaining a sufficient level of customer satisfaction. This study aimed to identify the key service quality (SQ) factors that drive passenger satisfaction in Istanbul’s rail transit (RT) system using data obtained from an extensive survey conducted by the Istanbul Public Transportation Co. A total of 11,116 passengers who used rail transport from May 15–June 3, 2012, and June 17–July 3, 2013, were interviewed in person. The relative importance of the SQ factors was assessed so that service provision could be prioritized and the enhancement of passenger satisfaction can be achieved employing several social choice techniques. The results indicate that, from an overall perspective, waiting time, crowdedness in cars, and fare are the SQ factors that best reflect the public good.

Keywords: Service quality; public transportation; rail transit systems; stated preferences; fallback voting; Istanbul
Introduction

Public transportation (PT) is a cost-effective solution for traffic congestion, especially in crowded areas, and its improvement is of critical importance to city governance and decisionmakers. As with many other PT services, rail transit (RT) systems should also hear the voice of the customer since decisionmakers need to create an efficient system to promote public transport use (Gronau and Kagermeier 2007; Le-Klähn et al. 2014).

With its 14.3 million inhabitants and a high level of socio-economic development, Istanbul is the heart of Turkey. The daytime population of this metropolis increases as many people commute from neighboring cities to Istanbul, which increases traffic congestion. The city’s population is expected to increase to nearly 15 million by 2019 and 16 million by 2023, according to the Turkish Statistical Institute (2015). Economic recovery and improvement in the standard of living makes passengers expect better PT services (bus rapid transit, rail transit, etc.). People prefer PT to avoid traffic congestion, noise, and long waiting times, especially during rush hours.

When all drawbacks of traveling by a private car are considered, RT has been one of the most appropriate modes of travel for public transport users in Istanbul. The city’s six RT lines (M1, M2, M4, T1, T4, F1), which are operated by Istanbul Public Transportation Co., total 145.5 kilometers in length and carry more than 1.3 million passengers daily (www.metro.istanbul/en). Figure 1 shows the network maps and characteristics of RT lines in Istanbul.
Identifying Key Factors of Rail Transit Service Quality: An Empirical Analysis for Istanbul

FIGURE 1. Istanbul rail transit line network map
To increase the number of people using the city's RT system, it is critical to gather information from the recent users of the system regarding how much their expectations are met so decisionmakers can make changes to meet the passenger needs better (Andaleeb 2007; Le-Klähn et al. 2014). Customer surveys are especially important nowadays considering the depth and amount of information they can provide so public transport providers can understand which service aspects play a more critical role in passenger satisfaction (Le-Klähn et al. 2014).

The Istanbul Public Transportation Co. conducts a customer satisfaction survey annually to determine the needs and problems of its RT passengers to improve the system based on their demand. In this study, we analyzed the results of the 2012 and 2013 surveys using several voting rules to evaluate the priority of each service quality (SQ) factor for RT passengers in Istanbul by line and year. A total of 11,116 passengers completed the surveys, which were distributed among 6 RT lines.

The remainder of this paper is structured as follows. Section 2 reviews the literature highlighting stated importance methods, Section 3 reports the details on survey data, and Section 4 provides a summary of the voting methods employed with exemplification. In Section 5, results from different categories of methods are clustered and interpreted. Finally, Section 6 concludes with an overall discussion emphasizing possible avenues for future research.

**Related Literature**

Proposing higher SQ levels in PT to reduce dissatisfaction (traffic congestion, noise, etc.) resulting from immoderate use of private cars in big cities is one of the most important issues for public transport providers. Thus, PT service planners pursue lessening the use of private cars by developing quality improvement plans that will initiate higher customer satisfaction. Increasing customer satisfaction or SQ levels results in a higher use of the service, involvement of new customers, and a better public image (de Oña et al. 2012; Çelik et al. 2013).

To reach an appropriate SQ level, service providers should consider several SQ factors associated with PT. Mouwen and Rietveld (2013) considered several factors to determine if competitive tendering increases SQ for PT in the Netherlands and determined that frequency of service, time accuracy, travel speed, and vehicle tidiness were the most effective. Waiting time, cleanliness, and comfort were observed to be the most valued PT factors in a study by dell'Olio et al. (2011). Redman et al. (2013) presented a comprehensive review on SQ factors in PT and determined that reliability, frequency, price, speed, access, comfort, and convenience were the factors that attract car users to use PT. Hassan et al. (2013) asserted that the most desirable SQ factors of PT services were reliability, frequency, capacity, price, cleanliness, comfort, security, staff, information, and ticketing system, with loading/ridership, travel time, travel distance, and service duration indicated as “efficiency” indicators.

Currently, in big and crowded cities, RT systems are preferred as one of the easiest ways of avoiding traffic congestion and noise. Therefore, analyzing service quality in
RT systems has gained importance. Gerçek et al. (2004) evaluated three alternative RT networks based on four main factors—financial, economic, system planning, and policy. Awasthi et al. (2011) integrated SERVQUAL and TOPSIS to evaluate the SQ of Montreal metro services. Nathanail (2008) evaluated the performance of Hellenic Railways based on 22 factors group into six major factors—itinerary accuracy, system safety, cleanliness, passenger comfort, servicing, and passenger information. The author concluded that the RT systems that paid attention to itinerary accuracy and system safety would perform best.

Brons et al. (2009) aimed to determine the significance level of the access-to-the-station effect on passenger overall satisfaction and the balance between the factors of the RT services. They concluded that, in several parts of the RT network, improving access services to the railway stations could substitute for improving the services provided on the rail network, which would attract passengers who used other transportation modes. Eboli and Mazzulla (2012) analyzed how RT passengers perceived different SQ factors, noting that promptness, consistency, frequency, and cleanliness had the highest positive influence for RT services. However, dell'Olio et al. (2010) noted that passenger perceptions on SQ might change depending on the type of passengers under consideration. Cascetta and Carteni (2014) provided a comparison between perceived and calculated SQ for a metro line servicing in the Campania region of Italy. In a study by de Oña et al. (2014b), passengers were clustered to determine the most important SQ factors, concluding that different factors may be determined as the most important for different groups of passengers. Punctuality was selected as the most important SQ factor for the first group (young female students who do not have a private car), and frequency was selected for the second group (women of medium age who frequently use public transport service for reaching jobs). From a general perspective, comfort, personnel, information, and service were determined as the most important factors (de Oña et al. 2014a).

As mentioned in Berry et al. (1990), since passengers are the only rulers of the systems in terms of SQ, their perception on SQ factors should be contemplated when evaluating the SQ level of a system. (Tyrinopoulos and Antoniou 2008; Filipović et al. 2009; Eboli and Mazzulla 2009, 2011). A study by de Oña et al. (2012) classified the methods proposed to evaluate the perceived importance of SQ factors into two main categories: "stated importance methods and derived importance methods. In the former, customers were asked to rate each factor on an importance scale, whereas in the latter, the importance of factors was determined by analyzing the relationship of each factor with the overall customer satisfaction via statistical testing.

In this study, the stated importance approach was adopted; however, as discussed in the related literature, it has several drawbacks (Eboli and Mazzulla 2008a, 2008b, 2010; Cirillo et al. 2011; Dell'Olio et al. 2011). First, stated importance methods may greatly suffer if passengers rate almost all of the criteria/items close to the top scale (e.g., 5 on a 5-point Likert scale). This results in an inadequate differentiation among mean importance ratings. In addition, such methods require that the survey cover a relatively longer period, which may reduce the overall response rate and the accuracy
of the survey. Some criteria found important may, in fact, have little effect on overall satisfaction (de Oña et al. 2012).

Despite the notable increase in the number of studies employing derived importance methods (Eboli and Mazzulla 2007; Dell’Olio et al. 2010; Jen et al. 2011) due to the issues summarized above, the stated importance approach has advantages over the derived importance approach. First, it is understood by decisionmakers and public policymakers more easily. It also requires fewer analytical skills and less expertise to employ (Van Ryzin and Immerwahr 2007). Nevertheless, interested readers are referred to Van Ryzin and Immerwahr (2007), Eboli and Mazzulla (2007), Dell’Olio et al. (2010), and Jen et al. (2011) to gain more insight on different applications of the derived importance methods.

Although many studies have focused on the evaluation of criterion-wise satisfaction levels or overall satisfaction level, few have paid attention to the relative importance of service quality. When determining the key SQ factors, research to date has neglected to consider customer preference rankings that are information-rich and can be easily processed and interpreted. If customers rate their satisfaction with only a specific SQ factor, the path followed by research done so far is inevitable; however, when customers order SQ factors based on their preferences, they provide more information regarding on what decisionmakers should focus. Therefore, this study contributes to the existing literature by providing a different aspect to analyze passenger satisfaction using a considerably large sample and comparing results between years and RT lines with the help of different voting procedures that are easy to implement. Highly-prioritized SQ factors were determined using a representative sample consisting of 11,116 individuals. The joint investigation of traditional and non-traditional voting methods for ranking the most important SQ factors also added value. In addition, determining high-priority SQ factors for each line separately provides more insight on potential differentiation between the lines considered. Finally, the procedures provide valuable information regarding SQ factors that should be primarily focused on to provide a better service in RT lines for future investments.

**Survey Data**

The survey was composed of four parts: Station and Ticketing, Rail Transit Usage, Overall and Criterion-Based Satisfaction, and Demographics. The survey questions measured each SQ factor on a 6-point Likert scales with “extremely satisfied” reflecting the highest favorable response and “extremely dissatisfied” indicating the least favorable response to each statement.

To determine the importance of SQ factors for RT lines in Istanbul, we analyzed passenger satisfaction surveys that were conducted among 4,966 passengers in 2012 (from May 15 to June 3), and 6,150 passengers in 2013 (from June 17 to July 3). The distribution of the 11,116 survey participants across years and lines are shown in Table 1. Since the M4 line was not open during the time the survey was conducted in 2012, there were no data available regarding that year.
All passengers were interviewed in person. A multistage stratified sampling procedure was employed in which tiers were formed considering the differences at rush hours and off-peak traffic hours among the transit lines. The quotas for the tiers were set according to the following criteria:

1. Day of use: weekdays, Saturday, Sunday
2. Time slot: morning rush, morning, noon, evening rush, evening
3. Station-wise crowdedness
4. Ticket type: token, full fare, discount fare, free

The participants in the survey were selected as follows to achieve randomness: A pollster waiting at an exit asked the 6th (5th in 2013) passenger who passed the turnstiles to participate in the survey; if that passenger was not willing, then the next passenger was asked to participate, and so on. At that point, the sequence of passengers was no longer important. Note that the pollster was not asked to follow a systematic sampling procedure in the first place, as it would be impractical to select every nth passenger for the survey, especially during rush hours.

Table 2 shows details on the survey data regarding demographics and travel characteristics. Note that median monthly household income of the survey participants was 1,782 Turkish lira (TL) (approx. $970 based on the Central Bank of Turkey’s exchange rate in May 28, 2012) in 2012, and the full transit fare was 1.65 TL during that time. This increased to 2,431 TL (approx. $1,250 based on the Central Bank of Turkey’s exchange rate in June 24, 2013) in the next survey year, and the full transit fare increased to 1.95 TL. Some notable differences between survey years regarding demographics appear in education level and household income level. The percentage of participants who had a primary school degree significantly decreased in contrast to the percentage of participants with an undergraduate degree. The frequency distribution of household income also changed; it was right-skewed in 2012, but was fairly symmetric in 2013 (with a higher median value compared to the previous survey year). This might be attributed to the introduction of a new line (M4) into the RT system by the time survey was conducted in 2013 since this line provides service in the Anatolian part of Istanbul, unlike the other five.
### TABLE 2. Respondent Profile

<table>
<thead>
<tr>
<th>Demographic Characteristics</th>
<th>2012</th>
<th>2013</th>
<th>Travel Profile Characteristics</th>
<th>2012</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td>Car Ownership</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>77.4%</td>
<td>74.6%</td>
<td>Yes</td>
<td>40.2%</td>
<td>43.0%</td>
</tr>
<tr>
<td>Female</td>
<td>22.6%</td>
<td>25.4%</td>
<td>No</td>
<td>59.8%</td>
<td>57.0%</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td>Time to Station</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15-25</td>
<td>46.7%</td>
<td>48.4%</td>
<td>Less than 5 min</td>
<td>31.3%</td>
<td>26.7%</td>
</tr>
<tr>
<td>26-35</td>
<td>26.4%</td>
<td>28.6%</td>
<td>6-10 min</td>
<td>25.8%</td>
<td>25.2%</td>
</tr>
<tr>
<td>36-45</td>
<td>13.6%</td>
<td>13.9%</td>
<td>11-15 min</td>
<td>13.6%</td>
<td>13.1%</td>
</tr>
<tr>
<td>46-55</td>
<td>8.0%</td>
<td>5.9%</td>
<td>16-20 min</td>
<td>9.1%</td>
<td>12.5%</td>
</tr>
<tr>
<td>Older than 55</td>
<td>5.3%</td>
<td>3.3%</td>
<td>21-25 min</td>
<td>2.8%</td>
<td>2.8%</td>
</tr>
<tr>
<td><strong>Educational Level</strong></td>
<td></td>
<td></td>
<td>Total Time of Travel</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary school not completed</td>
<td>0.6%</td>
<td>0.5%</td>
<td>Less than 10 min</td>
<td>8.0%</td>
<td>14.6%</td>
</tr>
<tr>
<td>Primary school degree</td>
<td>11.4%</td>
<td>8.1%</td>
<td>11-20 min</td>
<td>26.6%</td>
<td>37.9%</td>
</tr>
<tr>
<td>Secondary school degree</td>
<td>9.9%</td>
<td>9.8%</td>
<td>21-30 min</td>
<td>23.1%</td>
<td>18.7%</td>
</tr>
<tr>
<td>High school student</td>
<td>12.4%</td>
<td>9.2%</td>
<td>31-40 min</td>
<td>13.3%</td>
<td>10.6%</td>
</tr>
<tr>
<td>High school degree</td>
<td>22.1%</td>
<td>23.5%</td>
<td>41-50 min</td>
<td>10.2%</td>
<td>7.3%</td>
</tr>
<tr>
<td>Vocational school student/ degree</td>
<td>2.5%</td>
<td>3.0%</td>
<td>51-60 min</td>
<td>7.1%</td>
<td>4.6%</td>
</tr>
<tr>
<td>Undergraduate student</td>
<td>19.1%</td>
<td>17.3%</td>
<td>61-70 min</td>
<td>2.8%</td>
<td>1.7%</td>
</tr>
<tr>
<td>Undergraduate degree</td>
<td>18.7%</td>
<td>23.6%</td>
<td>71-80 min</td>
<td>2.0%</td>
<td>1.9%</td>
</tr>
<tr>
<td>Post graduate student/degree</td>
<td>3.3%</td>
<td>4.8%</td>
<td>81-90 min</td>
<td>2.6%</td>
<td>0.9%</td>
</tr>
<tr>
<td><strong>Monthly Household Income</strong></td>
<td></td>
<td></td>
<td>Frequency Of Use</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than 500 TL</td>
<td>1.2%</td>
<td>0.6%</td>
<td>At least once a week</td>
<td>25.7%</td>
<td>24.3%</td>
</tr>
<tr>
<td>501-1000 TL</td>
<td>13.9%</td>
<td>5.3%</td>
<td>Once a day</td>
<td>16.1%</td>
<td>9.7%</td>
</tr>
<tr>
<td>1001-1500 TL</td>
<td>22.0%</td>
<td>11.0%</td>
<td>Twice a day</td>
<td>48.0%</td>
<td>53.8%</td>
</tr>
<tr>
<td>1501-2000 TL</td>
<td>20.7%</td>
<td>15.2%</td>
<td>Three times a day</td>
<td>2.9%</td>
<td>4.5%</td>
</tr>
<tr>
<td>2001-2500 TL</td>
<td>12.5%</td>
<td>12.9%</td>
<td>More than three times a day</td>
<td>7.2%</td>
<td>7.7%</td>
</tr>
<tr>
<td>2501-3000 TL</td>
<td>10.2%</td>
<td>12.1%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3001-3500 TL</td>
<td>4.0%</td>
<td>7.2%</td>
<td>Ticket Type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3501-4000 TL</td>
<td>3.3%</td>
<td>5.5%</td>
<td>Full</td>
<td>52.5%</td>
<td>58.3%</td>
</tr>
<tr>
<td>4001-4500 TL</td>
<td>1.9%</td>
<td>3.4%</td>
<td>Discount (student, teacher, or social)</td>
<td>34.6%</td>
<td>34.0%</td>
</tr>
<tr>
<td>4501-5000 TL</td>
<td>3.1%</td>
<td>3.6%</td>
<td>Token or Free</td>
<td>13.0%</td>
<td>7.7%</td>
</tr>
<tr>
<td>More than 5001 TL</td>
<td>4.6%</td>
<td>9.5%</td>
<td>Full</td>
<td>52.5%</td>
<td>58.3%</td>
</tr>
<tr>
<td>Median</td>
<td>1782 TL</td>
<td>2431 TL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Marital Status</strong></td>
<td></td>
<td></td>
<td>Main Purpose of Travel</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>62.5%</td>
<td>66.6%</td>
<td>Commute</td>
<td>44.8%</td>
<td>53.1%</td>
</tr>
<tr>
<td>Married</td>
<td>37.5%</td>
<td>33.4%</td>
<td>Go to or return from school</td>
<td>20.9%</td>
<td>12.2%</td>
</tr>
<tr>
<td><strong>Employment Status</strong></td>
<td></td>
<td></td>
<td>Work-related activities</td>
<td>13.3%</td>
<td>8.1%</td>
</tr>
<tr>
<td>Unemployed or student</td>
<td>35.7%</td>
<td>30.2%</td>
<td>Entertainment or social activities</td>
<td>12.1%</td>
<td>20.2%</td>
</tr>
<tr>
<td>Employed</td>
<td>64.3%</td>
<td>69.8%</td>
<td>Other</td>
<td>8.9%</td>
<td>6.4%</td>
</tr>
</tbody>
</table>
Methodology

As mentioned earlier, we employed various voting procedures from the social choice literature to find the criterion that had the highest importance for Istanbul Public Transportation Co. passengers by RT line and year. One question in the surveys asks passengers to rank, in order of importance, five of the SQ factors given. These SQ factors were treated as candidates aiming to be the winner of an election and the passengers as voters. Given a set of SQ factors, each passenger prioritized the five based on their preferences.

The most appropriate method in the social choice literature to analyze such data is Fallback Voting, mainly because customers are almost always allowed to provide partial preference; however, one can argue that Approval Voting, in which a voter may cast one vote for as many candidates as desired without ranking them, or even Condorcet-consistent methods, which work mostly with complete preferences, also could work.

In this section, we briefly cover the relevant voting procedures without going into too much detail. Interested readers should refer to Tideman (1987) and Plassmann and Tideman (2014) for more information on the fundamentals of voting systems and social choice functions.

Assume that individual j is endowed with a preference relation $\succsim_j$ that is defined over $C$, a finite set of candidates competing in the election. A voting system is a function that takes an election as input and produces a set of winners, a subset of $C$. The preference relation $\succsim_j$ is desired to have some characteristics such as completeness, reflexivity, and transitivity. Completeness requires that given two different candidates, say A and B, from $C$, either $A \succsim_j B$ (A beats B according to j’s preference or they are tied) or $B \succsim_j A$. In other words, a voter’s ranking contains all available candidates. Incompleteness (partial preferences) corresponds to the case where the voters rank only a subset of candidates contained in $C$. Reflexivity states that any candidate $A \in C$ is as preferable as itself; that is, $A \succsim_j A$. Finally, given three candidates $A, B,$ and $C \in C$ such that $A \succsim_j B$ and $B \succsim_j C$, transitivity implies that $A \succsim_j C$. Note that the preference relation $\succsim_j$ is called weak since it allows for ties (indifference). A strict preference relation, which can be denoted by $\succ_j$, is irreflexive and individual j is assumed to rank one of the two arbitrary candidates in $C$ over the other; that is, individual j is never indifferent between any two candidates.

Bulk literature exists on consumer choice modeling that is based on the utility concept and is directly related to these preference relations with the aforementioned properties. They primarily assume that an individual gains an economic utility when he/she selects an alternative. However, these models are mostly probabilistic and make certain assumptions regarding individual characteristics and/or candidate characteristics. This stream of research is not covered here; however, interested readers should refer to Anderson et al. (1992) for a comprehensive review and detailed discussion on the origin and the evolution of statistics-oriented choice models and utility maximizing voters.

The preference relation explained above should carry extra properties to have a fair voting system; however, note that there is no ideal scheme to decide a winner in an election, as Arrow’s Impossibility Theorem proves (Kelly 1978).
1. Pareto optimality (unanimity): For all pairs of candidates A and B, if A is preferred to B by all the voters, then B should not be declared as the winner.

2. Monotonicity: Increasing (decreasing) the number of votes for a winning (losing) alternative cannot make it a loser (winner).

3. Anonymity: Voters are treated the same.

4. Independence of Irrelevant Alternatives (IIA): Suppose that a group of individuals decided that A should be ranked before B. If a new candidate, say N, which is outside C, was introduced, then the group decision would not change, provided that the relative ordering of A and B was preserved.

5. Non-dictatorship: "No voter should decide the outcome of an election" (Menton 2013).

Traditional voting rules can be grouped mainly into two categories according to their starting point: Condorcet-consistent methods and Condorcet-inconsistent methods. In the former, the main objective is to find a Condorcet winner if one exists; in the latter, the winner may be determined by “the points allocated to candidates according to their ranking on individual voters’ ballots” (Cox 1989). Such methods are called scoring-based methods. The rest of this section provides an overview on traditional methods along with recently-proposed voting methods, followed by main assumptions.

Condorcet (1789) asserted that the candidate that is preferred pairwise to every other candidate by a majority of voters wins the election. Such a candidate is called a Condorcet winner. If no such winner exists, all candidates tie for the win (Mattei 2012).

As an illustrative example, adapted from Schulze (2003), suppose that there are four cities (A, B, C, D) vying to host a special event and the 30 members of the international organizing committee are asked to rank each of these cities from the most favorable to the least favorable in terms of suitability to stage the event. The aggregated ranked ballots are as follows:

| A>D>C>D>B | 3   | B>C>D>A | 5   | C>D>A>B | 5   | D>A>B>C | 2   |
| A>D>B>C | 5   | B>A>C>D | 4   | C>A>D>B | 2   | D>B>A>C | 4   |

Most of the methods discussed here use a pairwise preference matrix that shows how many times candidates were preferred over one another. The original matrix is shown below (Table 3), and the first line reads: City A was preferred to cities B, C, and D in 17, 18, and 14 instances, respectively.

<table>
<thead>
<tr>
<th>FOR</th>
<th>AGAINST</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td>17</td>
<td>18</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>13</td>
<td>20</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>12</td>
<td>10</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>16</td>
<td>21</td>
<td>11</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
City A wins over cities B and C, but loses to city D. City D wins over cities A and B, but loses to city C. Thus, there is no Condorcet winner since none of the candidates won every comparison with all the other candidates. Note that if two of the voters who preferred the ranking D>B>A>C changed to A>C>D>B, city A would be the Condorcet winner since it would win all of its pairwise comparisons.

On the other hand, de Borda (1781) argues that a majority winner may not always exist and proposes a paired-comparisons procedure that assigns points to candidates in head-to-head elections. Each candidate's total score is calculated based on points associated with each rank. The winner that has the highest score is then declared the winner.

For the illustrative example above, city A has a Borda score of 49 ( = 8x3 + 8x2 + 9x1), city B has a Borda score of 42 ( = 9x3 + 4x2 + 7x1), city C has a Borda score of 41 ( = 7x3 + 8x2 + 4x1), and city D has a Borda score of 48 ( = 6x3 + 10x2 + 10x1). Thus, city A wins the election.

The Borda method may elect a candidate that was not ranked first by any of the voters (Mattei 2012) and it does not satisfy the IIA property. Eğecioğlu and Giritgil (2011) addressed the difficulty encountered when one aims to implement this method in case of partial preferences.

Condorcet-Consistent Methods

The rules summarized below assume complete linear orderings and select the Condorcet winner if one exists.

1. Baldwin’s iterative procedure employs the Borda count and eliminates the candidate(s) with the lowest Borda score(s) at each step and recalculates Borda scores for the remaining candidates; the procedure proceeds until a group of candidates with the same Borda score can be formed (Hwang and Lin 1987).

For the illustrative example mentioned earlier, at the first step, city C is eliminated since it has the lowest Borda score. Following the elimination, the Borda scores of cities A, B, and D become 31, 22, and 37, respectively. Thus, city B is eliminated. City D wins eventually as its reduced Borda score (16) is greater than that of city A’s (14).

2. Black (1958) elects a Condorcet winner if one exists; otherwise, the Borda count winner is elected.

3. Copeland’s rule (Copeland 1951) works with pairwise comparisons; it counts the number of wins and losses for each candidate competing in the election. For each win (loss), a candidate gains (loses) one point. The candidate with the highest total score wins the election. It allows for ties (no points assigned to the candidates that are tied), but it may be indecisive.

In the pairwise comparison matrix for the illustrative example above, we see that cities B and C are eliminated immediately since the former wins only over city C, whereas the latter wins only over city D. Cities A and D are tied since they
both have a Copeland score of 2. A tie-breaking rule is necessary at this point to
determine the winner.

4. Dodgson’s method determines the fewest number of pairwise interchanges
needed to make a candidate the Condorcet winner. The candidate with the fewest
interchanges is declared the winner (Black 1958). Determining a Dodgson winner
is a rather complicated procedure that is NP-hard and in which the time required
to determine the winner is polynomial (Caragiannis et al. 2012). Like Bucklin’s, this
method also works with full preference information.

City A needs only 2 swaps to win the election, whereas cities B and C both need
16 swaps, and city D needs 8. City A is the Dodgson winner as it needs the least
number of swaps to win the election.

5. Similar to Baldwin’s, Nanson’s rule eliminates at each step the candidates with a
Borda score smaller than the average Borda score (calculated considering all the
candidates at a step). The Borda scores are then revised, taking only the remaining
candidates into account. The procedure repeats until a Borda winner can be
determined (Nanson 1883).

**Condorcet-Inconsistent Methods**

The rules summarized below are not guaranteed to select the Condorcet winner if one
exists.

1. Bucklin’s method elects the candidate that was ranked first by the majority of
voters as the winner. If there exists no such candidate, the candidate that was
ranked either first or second by the majority of voters is declared the winner. The
procedure continues, expanding the number of levels to consider every time a
majority winner cannot be determined, until one of the candidates has more than
half the number of votes. Bucklin requires complete preference information as
well (Hoag and Hallett 1926).

Consider the illustrative example above. According to Bucklin’s rule, one of the
cities would need to be ranked first by at least 16 of the committee members to
win the election. However, the number of times cities A, B, C, and D preferred as
the organizer is 8, 9, 7, and 6, respectively. Thus, at the second stage, we count the
total number of times a city was ranked either first or second. In the end, cities A
and D are tied as they take the first or second places 16 times, whereas city B (city
C) appeared in the top two only 13 (15) times.

2. Coombs (1964) proposed a recursive elimination method that discards at each
step the candidate who was ranked last the most number of times. This rank
scoring procedure repeats until someone can be declared winner.

Approval Voting is also a rank scoring rule that allows individuals to vote for a
predetermined number of candidates available. For instance, under \( k \)-Approval Voting,
each ballot contains at most \( k \) candidates, but the voter is not asked to rank them. The
candidate that appears the most in the ballots wins the election.
Note that Felsenthal and Tideman (2014) report that Nanson, Dodgson, and Coombs are all vulnerable to monotonicity failure.

**Non-traditional Methods**

1. As an attempt to avoid cloning in elections, Schulze (2011) introduced a new Condorcet-consistent method based on a weighted majority graph using a best-path finding algorithm, which is solved in polynomial time. On a majority graph, vertices represent the candidates and edges correspond to the relative performance of pairs of candidates (Menton 2013). The method describes “the strong paths from each winning candidate to every other candidate” (Menton 2013). One needs to first determine the number of voters who strictly prefer one candidate over another. Then, all possible paths to reach from one candidate to another must be identified. The weakest link in a path determines the strength of that path. If there are more than one possible path from one candidate to another, then the path with the largest strength is chosen and it is called the strongest path. If the path from one candidate (X) to another (Y) is stronger than (no ties) the path backwards (Y to X), then X disqualifies Y. If X disqualifies every other candidate, then X wins the election outright (Schulze 2003).

Consider the example introduced earlier. The directed graph given in Figure 2 is constructed using the pairwise comparison matrix. There are two possible ways to reach B in this case: a direct path from A to B (with a strength of 17) and an indirect path from A to C to D to B (with a strength of 18). The strength of the latter is determined by the weakest link, which is A to C. The strongest path is the one with the largest relative pairwise performance; that is, A~C~D~B. Here, “~” denotes a direct link from one candidate to another. There is only one path from B to A with a strength of 16; B to C to D to A. Since the strength of the path from A to B is larger than that of B to A, A disqualifies B.

![Weighted majority graph for illustrated example](image)

Table 4 compares the strongest beatpaths. City A wins the election since it loses none of the beatpath comparisons.
2. Tideman’s ranked pairs method is very similar to that of Schulze’s and tells one “what edges are considered in what order, and whether and how the edges are set in the election graph.” (Menton 2013). The method “requires the collective ranking of the candidates to be consistent with the paired comparisons decided by the largest and second largest margins, and then, if possible, with the paired comparison decided by the third largest margin, and so on.” (Tideman 1987). The candidates are first ordered from top to bottom based on margin of victory in head-to-head elections. The ranking with the largest margin is determined and locked. Then, all rankings that contradict it is eliminated. The procedure continues with the next largest margin of victory until one ranking remains (Levin and Nalebuff 1995). For the example above, we start with B and D since they have the largest margin of victory (21 – 9 = 12). The ranking DB is locked. The second largest margin is between B and C (20 – 10=10) which lets us lock BC. Since DB and BC, the ranking DC is also locked. Finally, we lock AC, AB, and DA. Therefore, D wins the election based on the final ranking: DABC.

3. Fallback Voting (FV) is an extension of Bucklin’s procedure that does not need complete orderings, yet it does not allow for ties. FV combines Bucklin’s method with approval voting (Erdélyi et al. 2015), and, as Brams and Sanver (2009) summarized, it proceeds as follows.

First, voters rank a set of candidates they approve in order of preference. The set of approved candidates is allowed to be empty or to consist of all the candidates competing in the election. If a candidate was ranked first by a majority of voters, this candidate is called a level 1 FV winner. If no candidate can be declared a level 1 winner, the candidate that is ranked either first or second by a majority of voters is considered, and this candidate is declared the winner. If there are more than one such candidates, then the candidate with the largest majority is called a level 2 FV winner. If there is no level 2 winner, the voters descend—one level at a time—

### TABLE 4.
Strong Paths between Each Candidate

<table>
<thead>
<tr>
<th>From</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>A~B: 17</td>
<td>A~C: 18</td>
<td>A<del>C</del>D: 18 = min(18,19)</td>
<td>A<del>C</del>D~B: 18 = min(18,19,21)</td>
</tr>
<tr>
<td>B</td>
<td>B<del>C</del>D~A: 16 = min(20,19,16)</td>
<td>B~C: 20</td>
<td>B<del>C</del>D: 19 = min(20,19)</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>C<del>D</del>A: 16 = min(19,16)</td>
<td>C<del>D</del>B: 19 = min(19,21)</td>
<td>C~D: 19</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>D~A: 16</td>
<td>D~B: 21 = min(16,17)</td>
<td>D<del>A</del>B: 16 = min(16,17)</td>
<td>D<del>A</del>B~C: 16 = min(16,17,20)</td>
</tr>
</tbody>
</table>

#### Tideman’s ranked pairs method

Tideman’s ranked pairs method is very similar to that of Schulze’s and tells one “what edges are considered in what order, and whether and how the edges are set in the election graph.” (Menton 2013). The method “requires the collective ranking of the candidates to be consistent with the paired comparisons decided by the largest and second largest margins, and then, if possible, with the paired comparison decided by the third largest margin, and so on.” (Tideman 1987). The candidates are first ordered from top to bottom based on margin of victory in head-to-head elections. The ranking with the largest margin is determined and locked. Then, all rankings that contradict it is eliminated. The procedure continues with the next largest margin of victory until one ranking remains (Levin and Nalebuff 1995). For the example above, we start with B and D since they have the largest margin of victory (21 – 9 = 12). The ranking DB is locked. The second largest margin is between B and C (20 – 10=10) which lets us lock BC. Since DB and BC, the ranking DC is also locked. Finally, we lock AC, AB, and DA. Therefore, D wins the election based on the final ranking: DABC.

#### Fallback Voting (FV)

Fallback Voting (FV) is an extension of Bucklin’s procedure that does not need complete orderings, yet it does not allow for ties. FV combines Bucklin’s method with approval voting (Erdélyi et al. 2015), and, as Brams and Sanver (2009) summarized, it proceeds as follows.

First, voters rank a set of candidates they approve in order of preference. The set of approved candidates is allowed to be empty or to consist of all the candidates competing in the election. If a candidate was ranked first by a majority of voters, this candidate is called a level 1 FV winner. If no candidate can be declared a level 1 winner, the candidate that is ranked either first or second by a majority of voters is considered, and this candidate is declared the winner. If there are more than one such candidates, then the candidate with the largest majority is called a level 2 FV winner. If there is no level 2 winner, the voters descend—one level at a time—to
lower and lower ranks of approved candidates, stopping when, for the first time, one or more candidates are approved by a majority of voters, or no more candidates are ranked. If exactly one candidate receives majority approval, this candidate is the FV winner. If more than one candidate receives majority approval, then the candidate with the largest majority is the FV winner. If the descent reaches the lowest rank of all voters and no candidate is approved by a majority of voters, the candidate with the most approval is the FV winner.

In the illustrative example, since none of the candidates take the majority of the votes, a level 1 FV winner is not found. Next, we add the second-rank counts and see that cities A and D are tied. Descending to the third level leads city D to victory since it appears in the top three 26 times, once more than city A. Hence, city D is a level 3 FV winner.

Black, Copeland, Dodgson, Schulze, Tideman, Nanson, and Baldwin choose the Condorcet winner if one exists. One disadvantage of Condorcet arises when the group decision is not transitive, even though the individual preferences are (Mattei 2012). Most of the traditional methods enjoy completeness; however, it is often highly impractical to ask individuals compare alternatives in pairwise fashion (a preference or a tie). Thus, in customer satisfaction surveys, where there are too many alternatives, respondents are usually asked to rank a subset of them. This avoids cognitive complexity and waste of time, yet results in incomplete preferential votes. FV is designed to work with incomplete information and asks voters to select a set of candidates they approve and then rank them (Brams and Sanver 2009). The social choice literature on voting rules is expanding continually and alternative methods are being introduced. Recently, Camps et al. (2013) provided a continuous rating method for the social acceptance of different alternatives in case the individuals do not express a comparison between every pair of alternatives available or they provide an ordered list restricted to a subset of the most preferred options.

Finally, we list below our main assumptions that will provide us flexibility when interpreting the results in Section 5:

1. The respondents did not choose strategically; he/she is not be interested in what other respondents think or how they decide. In short, the voters are assumed to be sincere.

2. When employing the traditional methods, we assumed that the ballots are completely filled.

3. Since multiple winners would not be an issue, we did not work through a tiebreaking procedure.

Results and Discussion

The survey question we considered asked passengers to rank, based on their preferences, the five most important SQ factors listed in Table 5. They were allowed to report incomplete rankings; however, fewer than 2% of participants provided a ranking
with less than five SQ factors. In total, 26 SQ factors were assessed to determine those of higher priority—based on individual rankings—for each line and year.

### TABLE 5.
**SQ Factors**

<table>
<thead>
<tr>
<th>SQ Factors</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Waiting time</td>
<td>Lighting</td>
</tr>
<tr>
<td>Access to stations</td>
<td>Fares</td>
</tr>
<tr>
<td>Security at stations</td>
<td>Ventilation systems in cars</td>
</tr>
<tr>
<td>Security in cars</td>
<td>Moving stairways/walkways and escalators</td>
</tr>
<tr>
<td>Attitude of security personnel</td>
<td>Token machines</td>
</tr>
<tr>
<td>Travel (in-vehicle) time</td>
<td>Working conditions of turnstiles</td>
</tr>
<tr>
<td>Crowdedness in cars</td>
<td>Comfort level at stations</td>
</tr>
<tr>
<td>Cleanliness of stations</td>
<td>Transport information systems</td>
</tr>
<tr>
<td>Timeliness of cars</td>
<td>Notifications at stations in case of delay</td>
</tr>
<tr>
<td>Information systems at stations</td>
<td>Notifications in cars in case of delay</td>
</tr>
<tr>
<td>Token sale</td>
<td>Transfer fares</td>
</tr>
<tr>
<td>Cleanliness of cars</td>
<td>Notifications on website</td>
</tr>
<tr>
<td>Noise and vibration in cars</td>
<td>Notifications from phone line</td>
</tr>
</tbody>
</table>

We used several traditional voting procedures and three recently-proposed voting procedures (Schulze, Tideman, and Fallback Voting) to determine the highly-prioritized criteria by line and year. We reported the first, second, and third priorities identified using the methods explained in Section 4. As mentioned earlier, traditional procedures can be mainly grouped into two categories with respect to Condorcet-consistency. It would not be surprising to see that two traditional methods from different categories chose different candidates as winners. Yet, they agreed with each other at almost every instance, as seen in Tables 6 and 7, when only the primary (first, second, or third) priorities were considered. In contrast with other traditional methods, Dodgson and Simpson chose security at stations as the third priority for the M2 line in 2012. Hence, we combined the results of the traditional methods other than Borda in one table.

Note that the ballots in our study are truncated. The passengers ranked, at most, five of the SQ factors available in order of importance. When employing the traditional methods, we assumed that the voters strictly ranked the first five candidates and they were indifferent with the rest, which let us work with completely-filled ballots in return. Investigating the second and third priorities reveals that the Borda method is significantly affected by this assumption; there are nine such instances on which Borda and the other traditional methods do not agree. Regarding the first priorities, the Borda method and the traditional methods disagree only for M1 in 2013. The former favors fares, whereas the latter favors waiting time in that case. That is, when the first priorities are considered, these two clusters of methods differ from each other less significantly compared to the case when the second-ranked or third-ranked priorities are taken into account.

Waiting time appears to be a consistent problem for M2 line. In both years, this criterion is observed as a first priority for M2 passengers. Another interesting finding belongs to F1 and T4 lines. The priorities of F1 and T4 passengers changed through survey years.
Waiting time and crowdedness in cars appear to be the most important SQ factors on which decisionmakers should focus given the survey results of 2013. As for the second and third priorities, Borda and the other traditional methods point out different criteria at almost every instance for T1, M4, and F1 lines.

**TABLE 6.**
Passenger Priorities by Line and Year – Borda

<table>
<thead>
<tr>
<th>Priority</th>
<th>Year</th>
<th>Rail Transit Line</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>T1</td>
</tr>
<tr>
<td>First</td>
<td>2013</td>
<td>Crowdedness in cars</td>
</tr>
<tr>
<td></td>
<td>2012</td>
<td>Waiting time</td>
</tr>
<tr>
<td>Second</td>
<td>2013</td>
<td>Fares</td>
</tr>
<tr>
<td></td>
<td>2012</td>
<td>Security at stations</td>
</tr>
<tr>
<td>Third</td>
<td>2013</td>
<td>Waiting time</td>
</tr>
<tr>
<td></td>
<td>2012</td>
<td>Fares</td>
</tr>
</tbody>
</table>

**TABLE 7.**
Passenger Priorities by Line and Year – Other Traditional Methods

<table>
<thead>
<tr>
<th>Priority</th>
<th>Year</th>
<th>Rail Transit Line</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>T1</td>
</tr>
<tr>
<td>First</td>
<td>2013</td>
<td>Crowdedness in cars</td>
</tr>
<tr>
<td></td>
<td>2012</td>
<td>Waiting time</td>
</tr>
<tr>
<td>Second</td>
<td>2013</td>
<td>Fares</td>
</tr>
<tr>
<td></td>
<td>2012</td>
<td>Security at stations</td>
</tr>
<tr>
<td>Third</td>
<td>2013</td>
<td>Waiting time</td>
</tr>
<tr>
<td></td>
<td>2012</td>
<td>Fares</td>
</tr>
</tbody>
</table>

* Excluding Dodgson and Simpson procedures.
As seen in Tables 8 and 9, in 13 instances, Tideman, Schulze, and FV disagree with each other when determining a high-priority SQ factor. In 2013, in contrast with FV, Tideman and Schulze favored waiting time and crowdedness in cars over fares as the first priority, for the M1 and F1 lines, respectively. On the other hand, FV and the Borda procedure addressed different SQ factor as a priority at a total of five instances (T1, M2, and F1 lines); however, only the orderings differ for M2 and F1 lines. Thus, these two methods refer to the same set of SQ factors for both of these lines in a given year when ordering is overlooked. This is not the case when the results of FV are compared with those from the Tideman and Schulze methods. Distinctions occur mostly for the second and third priorities in 2013 (M1, M4, and F1 lines). The major difference in the set of highly-prioritized criteria in 2012 is observed for M2 line; FV elects cleanliness of cars (travel time) as the second (third) priority, whereas both Tideman and Schulze elect travel time (security at stations) as the second (third) priority.

In 2013, the methods are quite consistent regarding first priorities. For M1 line, FV and Borda elect fares, whereas Tideman and Schulze favor waiting time along with the other traditional methods. Similarly, for F1 line, all the methods excluding FV elect crowdedness in cars. On the other hand, the methods lead exactly to the same set of primary priorities with subtle differences in the ordering for F1 in 2012.

As mentioned in Section 4, it is not always possible to find an FV winner. For example, distinctions were detected between the traditional methods and the recent methods in 2012 for M1 and F1 lines. The passengers of these lines prioritized fares above others. However, we were not able to determine an FV winner for M1 line even after the first five ranks were considered in 2012. Thus, fares is the SQ factor that has the most approvals among the others in that case and is also the 5-approval winner.

<table>
<thead>
<tr>
<th>Priority</th>
<th>Year</th>
<th>T1</th>
<th>T4</th>
<th>M1</th>
<th>M2</th>
<th>M4</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td>2013</td>
<td>Crowdedness in cars</td>
<td>Crowdedness in cars</td>
<td>Fares</td>
<td>Waiting time</td>
<td>Waiting time</td>
<td>Fares</td>
</tr>
<tr>
<td></td>
<td>2012</td>
<td>Crowdedness in cars</td>
<td>Waiting time</td>
<td>Fares</td>
<td>Waiting time</td>
<td>Fares</td>
<td></td>
</tr>
<tr>
<td>Second</td>
<td>2013</td>
<td>Fares</td>
<td>Fares</td>
<td>Waiting time</td>
<td>Crowdedness in cars</td>
<td>Security at stations</td>
<td>Crowdedness in cars</td>
</tr>
<tr>
<td></td>
<td>2012</td>
<td>Fares</td>
<td>Security at stations</td>
<td>Crowdedness in cars</td>
<td>Cleanliness of cars</td>
<td>Crowdedness in cars</td>
<td></td>
</tr>
<tr>
<td>Third</td>
<td>2013</td>
<td>Waiting time</td>
<td>Waiting time</td>
<td>Travel time</td>
<td>Fares</td>
<td>Crowdedness in cars</td>
<td>Access to stations</td>
</tr>
<tr>
<td></td>
<td>2012</td>
<td>Waiting time</td>
<td>Crowdedness in cars</td>
<td>Waiting time</td>
<td>Travel time</td>
<td>Waiting time</td>
<td></td>
</tr>
</tbody>
</table>
The average level of satisfaction for each SQ factor can be seen in Table 10. The number in parenthesis shows the rank of an SQ factor among all the others based on its satisfaction level for a given line and year. Crowdedness in cars, with which customers from all lines are dissatisfied, was found to be a high-priority SQ factor in all lines with two exceptions: M1 in 2013 and M2 in 2012. Thus, service providers should focus resources on improving this SQ factor to significantly increase ridership. On the other hand, even though the satisfaction level for waiting time was relatively high for F1 line in 2012, it was addressed as a third priority for this line that year. This SQ factor has a lower satisfaction level whenever it is highly prioritized. Thus, spending time and/or money on its improvement can also enhance the overall passenger satisfaction. Even though waiting time was perceived as a highly important SQ factor by passengers, travel time (in-vehicle time) was prioritized in 2012 and 2013 by only M2 and M1 passengers, respectively. Yet, the passengers in these cases seem to be satisfied with this SQ factor as seen in Table 10.

The frequency distributions for overall customer satisfaction levels, which should definitely be taken into account to make a better conclusion, are given in Figure 3. The percentage of extremely satisfied passengers shows a significant increase in both M2 and T4 lines from 2012 to 2013. Therefore, one should compare not only the rankings, but also the average satisfaction levels of an SQ factor from different years since the change in the overall satisfaction may be attributed to sample-based differences in these years rather than a significant increase in customer satisfaction. For instance, waiting time should be improved in M2 line as it is a high-priority SQ factor with an increasing satisfaction level (from 4.82 to 5.04) but a decreasing relative satisfaction (from 7th to 14th). As for T4 line, crowdedness in cars definitely needs attention since it appeared as a high-priority SQ factor with a decreasing satisfaction level despite the increase in its overall satisfaction level.
TABLE 10. Satisfaction Levels for Highly-Prioritized SQ Factors

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Waiting time</td>
<td>4.31 (19)</td>
<td>4.41 (14)</td>
<td>4.58 (19)</td>
<td>4.9 (18)</td>
<td>4.63 (19)</td>
<td>4.67 (16)</td>
<td>4.82 (7)</td>
<td>5.04 (14)</td>
<td>N/A</td>
<td>5.53 (8)</td>
<td>5.12 (5)</td>
<td>4.92 (1)</td>
</tr>
<tr>
<td>Access to stations</td>
<td>4.61 (8)</td>
<td>4.53 (9)</td>
<td>4.81 (7)</td>
<td>5.13 (11)</td>
<td>4.82 (12)</td>
<td>4.75 (9)</td>
<td>4.79 (13)</td>
<td>5.06 (11)</td>
<td>N/A</td>
<td>5.34 (20)</td>
<td>5.12 (6)</td>
<td>4.73 (9)</td>
</tr>
<tr>
<td>Security at stations</td>
<td>4.58 (10)</td>
<td>4.33 (16)</td>
<td>4.74 (14)</td>
<td>5.16 (8)</td>
<td>4.75 (16)</td>
<td>4.63 (17)</td>
<td>4.69 (19)</td>
<td>5.01 (16)</td>
<td>N/A</td>
<td>5.49 (14)</td>
<td>4.91 (14)</td>
<td>4.71 (12)</td>
</tr>
<tr>
<td>Travel (in-vehicle) time</td>
<td>4.37 (17)</td>
<td>4.26 (18)</td>
<td>4.67 (18)</td>
<td>4.82 (19)</td>
<td>4.95 (6)</td>
<td>4.87 (2)</td>
<td>4.96 (1)</td>
<td>5.06 (10)</td>
<td>N/A</td>
<td>5.52 (10)</td>
<td>5.14 (3)</td>
<td>4.57 (18)</td>
</tr>
<tr>
<td>Cleanliness of cars</td>
<td>2.40 (24)</td>
<td>2.94 (22)</td>
<td>4.00 (25)</td>
<td>3.69 (26)</td>
<td>3.47 (26)</td>
<td>3.92 (23)</td>
<td>3.93 (25)</td>
<td>4.09 (25)</td>
<td>N/A</td>
<td>5.11 (24)</td>
<td>3.58 (25)</td>
<td>3.79 (25)</td>
</tr>
<tr>
<td>Ventilation sys. in cars</td>
<td>4.57 (11)</td>
<td>4.58 (3)</td>
<td>4.79 (11)</td>
<td>5.10 (13)</td>
<td>4.76 (15)</td>
<td>4.73 (11)</td>
<td>4.81 (8)</td>
<td>5.00 (18)</td>
<td>N/A</td>
<td>5.48 (15)</td>
<td>4.92 (13)</td>
<td>4.62 (17)</td>
</tr>
<tr>
<td>Fares</td>
<td>3.54 (23)</td>
<td>3.65 (21)</td>
<td>4.19 (24)</td>
<td>3.97 (25)</td>
<td>3.68 (25)</td>
<td>3.83 (24)</td>
<td>4.11 (23)</td>
<td>4.16 (24)</td>
<td>N/A</td>
<td>5.38 (18)</td>
<td>3.64 (24)</td>
<td>3.76 (26)</td>
</tr>
<tr>
<td>Ventilation sys. in cars</td>
<td>3.71 (22)</td>
<td>4.21 (20)</td>
<td>4.51 (20)</td>
<td>4.73 (20)</td>
<td>4.37 (21)</td>
<td>4.49 (20)</td>
<td>4.45 (21)</td>
<td>4.68 (21)</td>
<td>N/A</td>
<td>5.25 (22)</td>
<td>4.42 (22)</td>
<td>4.51 (20)</td>
</tr>
</tbody>
</table>
FIGURE 3. Distribution of overall satisfaction levels by year and line
As stated previously, the purpose of this study was to provide a relatively simple way to decide which SQ factors perceived more important by RT passengers considering preference rankings. In line with dell'Olio et al. (2011) and Celik et al. (2014), we found that waiting time is a highly-prioritized SQ factor on which the service provider should focus resources. Similarly, crowdedness in cars was also addressed as a highly important SQ factor in de Oña et al. (2014) and Aydin et al. (2015). Givoni and Rietveld (2007) and Brons et al. (2009) emphasized the importance of access to stations, whereas Bhat and Sardesai (2006) highlighted travel (in-vehicle) time as an important SQ factor. However, in our case, these two factors were favored by the respondents at considerably few instances. However, the service provider should never overlook their importance.

Assuming that the respondents had no knowledge of the ranking patterns, there was no concern about the effects of manipulation or bribery on the overall conclusion (Mattei 2012). On the other hand, even though the results of the traditional methods coincide with those obtained from the FV, the reader should note that the complete preferential votes could change the outcome.

Incorporating preference rankings into the analysis avoids missing valuable information. Such information should not easily be disposed, as this may distort the overall conclusion. Employing k-approval voting could also have been considered; however, the relative ranking of the five most favored SQ factors would have been ignored in that case. For instance, one could employ 5-approval voting in this case, which would ignore the ordering of SQ factors unlike FV and different SQ factors might be declared winners.

Note also that we do not report the priorities by year using the overall data, as some of the voting rules employed might suffer from the multiple districts paradox, which describes the case in which a candidate that won an election in distinct electoral districts is not declared the winner when the districts are joined together (Young 1974; Plassmann and Tideman 2014).

Conclusions

Offering high-quality service in PT allows passengers to avoid traffic congestion and noise, especially in big and crowded cities such as Istanbul. Hence, determining the key SQ factors that passengers value most is an essential task for PT service providers and policymakers. This paper reports on the results obtained by analyzing data from a passenger satisfaction survey conducted annually by Istanbul Public Transportation Co. Several voting rules available in the literature were employed and compared to decide which SQ factors would be perceived as more important by RT passengers.

The findings indicate that improving waiting time, crowdedness in cars, and fares can increase passenger satisfaction with RT services. Since a considerably high percentage of passengers prefer RT to commute or for work-related activities, they ranked waiting time and crowdedness in cars higher than most of the other SQ factors; they would rather get to work on time, comfortably. Hence, policymakers should focus on improving the comfort in cars and increasing the frequency of cars to decrease waiting time.
In addition, customers reported that they were dissatisfied with fares and they favored it as one of the most important SQ factors. However, the median household income level appears to increase through survey years. Even though this does not clearly reflect purchasing power, the service provider should either find a way to decrease fares or increase the level of service customers are provided to a level such that price paid is not seen as a problem. This would be a rational decision considering more than half of the respondents pay full fare.

Arrow’s Impossibility Theorem states that there is no ideal scheme to decide a winner in an election (Kelly 1978); however, following the footsteps of Camps et al. (2013) and re-analyzing the data in hand may lead decisionmakers to more reliable results even though the overall conclusion presented here repeats. This is highly recommended, especially when the decisionmaker would like to learn about how much social acceptance a certain criterion is provided rather than whether it is a primary priority for RT passengers. Note that passengers are one of the stakeholders in transportation planning. Other stakeholders such as motorists and transit agencies should also be brought into the discussion to make better decisions.

A possible avenue for future research is to investigating the vulnerability of the methods employed to sampling procedures. We assumed that the samples in each year are the best representations of the population. However, one might be interested in checking this, especially whether monotonicity property is violated, since Coombs, Nanson, and Dodgson are all vulnerable to monotonicity failure (Felsenthal and Tideman 2014). Another possibility for future research lies within a machine-learning setting that finds rank orderings, as mentioned in Dobrska et al. (2011). Investigating the effect of demographics on the priorities did not provide an enhancement of the results mentioned in Section 5, mainly due to the similitude of RT lines from this perspective. Further investigation using multivariate techniques such as multiple discriminant analysis might be considered to assess the importance of demographics.

Acknowledgments

The authors express their utmost gratitude to Istanbul Public Transportation Co. (İstanbul Ulaşım A.Ş.) for their understanding, support, and the data provided.

References


**About the Authors**

**ERKAN ISIKLI** (isiklie@itu.edu.tr) is currently an instructor in the Department of Industrial Engineering at the Istanbul Technical University in Turkey. He received BSc. and MSc. degrees in Mathematical Engineering and Business and Administration from Istanbul Technical University in 2005 and 2007, respectively, and a Ph.D. degree in Industrial and Systems Engineering from Wayne State University in Detroit. He taught several graduate and undergraduate level courses during his Ph.D. studies and currently...
is conducting projects related to decisionmaking, statistics, data mining, mathematics, and public transportation.

**Nezir Aydin** (nzraydin@yildiz.edu.tr) is currently an Assistant Professor in the Department of Industrial Engineering at the Yildiz Technical University in Istanbul, which he joined in 2014. He received BSc. and MSc. Degrees in Industrial Engineering from Yildiz Technical University in 2005 and 2007, respectively, a Ph.D. degree in Industrial and Systems Engineering from Wayne State University. He worked as a researcher and taught several graduate and undergraduate level courses during his Ph.D. studies and currently is conducting projects related to optimization, mathematical modeling, humanitarian logistics, decisionmaking, supply chain management, and public transportation.

**Erkan Celik** (erkancelik@munzur.edu.tr) is an Assistant Professor in the Department of Industrial Engineering at Munzur University in Turkey. He received BSc. and MSc. degrees in Industrial Engineering from Selcuk University in 2008 and 2011, respectively, and a Ph.D. degree from the Department of Industrial Engineering at Yildiz Technical University in 2015. His research interests are in decision analysis, public transportation, humanitarian logistic, and fuzzy sets.

**Alev Taskin Gumus** (ataskin@yildiz.edu.tr) is an Associate Professor in the Department of Industrial Engineering at Yildiz Technical University. She received BSc. and MSc. degrees in Industrial Engineering from Yildiz Technical University, an MBA from Istanbul Technical University, and a Ph.D. in Industrial Engineering from Yildiz Technical University. She completed postdoctoral research at Zaragoza Logistics Center in Spain, and her research interests are in supply chain management, public transportation, production and inventory systems, decision analysis, artificial intelligence, and fuzzy logic applications in industrial engineering and management sciences.
The Causal Effect of Bus Rapid Transit on Changes in Transit Ridership

Orion T. Stewart and Anne Vernez Moudon
University of Washington

Brian E. Saelens
University of Washington and Seattle Children’s Research Institute

Abstract

Numerous studies have reported ridership increases along routes when Bus rapid transit (BRT) replaces conventional bus service, but these increases could be due simply to broader temporal trends in transit ridership. To address this limitation, we compared changes in ridership among routes where BRT was implemented to routes where BRT was planned or already existed in King County, Washington. Ridership was measured at 2010, 2013, and 2014. Ridership increased by 35% along routes where BRT was implemented from 2010 to 2013 compared to routes that maintained conventional bus service. Ridership increased by 29% along routes where BRT was implemented from 2013 to 2014 compared to consistent existing BRT service. These results provide stronger evidence for a causal relationship between BRT and increased transit ridership and a more accurate estimate of the independent effect of BRT on ridership.

Keywords: Longitudinal study, quasi-experimental, transportation system change, land use

Introduction

Metropolitan areas across the world are working to increase transit ridership to improve mobility and economic vitality. Bus rapid transit (BRT) is a particularly attractive method to add transit capacity and potentially increase ridership (Currie and Delbosc 2013). BRT promises the speed and reliability of rail while retaining the operating flexibility and lower cost of conventional bus service (Deng and Nelson 2011). This is achieved by running high-capacity buses with streamlined boarding systems along prioritized surface routes at frequent intervals. BRT was pioneered as a “surface metro” in Curitiba, Brazil, in the early 1970s and has since expanded to at least 204 cities worldwide (Across Latitudes and Cultures - Bus Rapid Transit 2016; Cervero 1998).
The exact mix of BRT components varies widely from system to system (Cervero 2013), yet studies consistently suggest that the increased service, reduced travel times, and improved facility identity that occur when BRT replaces conventional bus service result in increases in ridership (Kittelson & Associates Inc. et al. 2007; Levinson et al. 2003; Peak et al. 2005; US Government Accountability Office 2012). Increases in corridor-level ridership over one year can reach 80% (US Government Accountability Office 2012). Furthermore, transit surveys show that new BRT service attracts choice transit riders—those who previously made the trip by a non-transit mode—as well as new transit riders who previously did not make the trip at all (Peak et al. 2005).

Despite these positive findings, there is limited evidence for a causal relationship between BRT implementation and increases in transit ridership for three main reasons. First, most studies only evaluate ridership along routes where BRT was implemented and fail to account for potential increases in ridership among nearby non-BRT routes due to transfers to or from BRT or potential decreases in ridership to nearby routes due to shifts to the BRT route. Second, there is a degree of variability in transit ridership from stop to stop along a corridor, and few studies apply inferential statistics to determine if observed changes in ridership are beyond what may be due to chance by this stop-to-stop variation in ridership. Finally and most important, transit ridership along corridors where BRT was implemented could have increased to the same extent under continued conventional bus service. This counterfactual scenario is impossible to observe, but it can be approximated by comparing corridors where BRT was implemented to similar control corridors where no changes in transit service occurred over the same time period. This concept is illustrated in two studies of Adelaide, Australia, and Oakland, California, which respectively observed 76% and 66% increases in ridership along corridors where BRT was implemented during a time when the overall transit system experienced a decline in ridership (Kittelson & Associates Inc. et al. 2007; Peak et al. 2005). The entirety of a transit system, however, may not be a good basis for comparison. BRT may be implemented along certain corridors because these same corridors are experiencing increased demand for transit. Hence, projected increases in transit use may cause the BRT to be implemented rather than the BRT causing the increased transit use.

This study took advantage of an incremental roll-out of BRT in King County, Washington, to compare changes in ridership at stops along traditional bus corridors where BRT was implemented to corridors where BRT was either planned but not yet implemented or already existed. These comparison groups are appropriate because they consist of valid candidates for BRT intervention. We further added to the rigor of the assessment by measuring ridership at all transit stops serving a corridor where BRT was implemented. This helped account for increases in ridership at other routes due to transfers to or from BRT or decreases at other routes due to ridership shifting to the BRT route. Finally, we applied a longitudinal regression model to estimate differences in changes in ridership among corridors where BRT was implemented and corridors where no changes occurred. This model accounted for correlation among stops to provide a robust estimate of changes in ridership and to estimate if these changes are beyond the
realm of chance (Locascio and Atri 2011). This study was intended to strengthen the evidence for a causal association between BRT and changes in transit ridership.

**Methods**

This study used a quasi-experimental stepped wedge study design to assess changes in ridership, as King County Metro replaced conventional bus service with BRT along six transit corridors over a four-year period. Stepped wedge studies involve the sequential roll-out of an intervention to all participants over a number of time periods and often are used for ethical reasons when there is a good reason to believe that the intervention will do more good than harm and for practical reasons when it is impossible to deliver the intervention simultaneously to all participants (Handley et al. 2011). Analysis in stepped wedge studies involves comparing outcomes among those who received the intervention and those who did not at a given time (Brown and Lilford 2006). In this study, changes in ridership at transit stop locations that were upgraded to BRT service were compared to transit stop locations where no changes occurred during the same time period. The evaluation is considered quasi-experimental because the location of BRT service and timing of the roll-out of BRT to the six bus corridors was not chosen at random.

**Study Setting**

King County Metro implemented “RapidRide” branded BRT service in the Seattle metropolitan area starting in October 2010. RapidRide service replaced existing traditional bus service along six existing corridors:

- RapidRide A line replaced bus route 174 starting on October 2, 2010
- RapidRide B line replaced bus routes 230 and 253 starting on October 1, 2011
- RapidRide C line replaced bus routes 54 and 54 express starting on September 29, 2012
- RapidRide D line replaced bus routes 15 and 18 starting on September 29, 2012
- RapidRide E line replaced bus route 358 express starting on February 15, 2014
- RapidRide F line replaced bus routes 110 and 140 starting on June 7, 2014

RapidRide BRT implementation featured changes to vehicles, stops, routes, and service (King County Metro 2016). RapidRide buses were designed to minimize boarding time through three doors, interiors that enable riders to quickly move to seats, and wheelchair restraints that do not require assistance from the bus driver. RapidRide “stations,” which account for 48% of RapidRide stops, feature shelters, lighting, pre-pay kiosks, and real-time information systems indicating when the next bus will arrive. RapidRide routes use a combination of transit priority features, including high-occupancy vehicle (HOV) and business access and transit (BAT) lanes, bus bulbs, queue jumps, and signal prioritization. Service was changed from a fixed schedule for traditional buses to BRT 10-minute headways during peak periods and 15-minute...
The Causal Effect of Bus Rapid Transit on Changes in Transit Ridership

headways during off-peak periods. The RapidRide system features distinct branding from the conventional King County Metro bus system. Compared to other BRT systems, RapidRide qualifies as a BRT “lite” primarily because routes comprise varying levels of priority lanes rather than exclusive transit ways and stations are more similar to traditional bus stops as opposed to rail station platforms (Cervero 2013). A 2014 performance evaluation found that route-level travel time had generally decreased and ridership had generally increased along RapidRide corridors compared to times immediately prior to implementation (Parametrix 2014). This prior evaluation, however, did not assess changes to ridership at connecting or competing bus routes, compare changes along RapidRide routes to other similar routes where no service changes occurred, nor attempt to determine if observed changes were beyond the realm of chance.

Unit of Analysis: RapidRide Stop Places

This analysis used geocoded bus stop locations and corresponding stop-level total boarding and alightings (ridership) collected by King County Metro during three time periods to assess changes in ridership. Analyzing longitudinal changes in ridership at the bus-stop level can be problematic. Individual stops are sometimes closed and replaced by new stops with new IDs in similar locations, which makes it difficult to track changes in ridership across minor bus stop relocations or upgrades. Such changes often occurred as part of RapidRide implementation. Analyzing all bus ridership within a buffer of each stop is one solution to this problem. However, multiple stops often are very near one another; for example, stops across the street may serve different directions of the same route, which results in very similar measurements of ridership and violates the assumption of independent observations required for most regression models. Conceptually, individual bus stops (or buffers around them) also may not be the most appropriate unit of analysis. Because of transfers to nearby bus stops serving different routes and round trips with origins and destinations at the same place, broader “catchment areas” around groups of bus stops may more appropriately capture how riders interact with the transit system. Thus, for this analysis, the unit of analysis was the location of groups of nearby RapidRide stops, or “RapidRide stop places,” that were present in Fall 2014 after all RapidRide lines were in service. These RapidRide stop places were applied retrospectively to take measurements over the study period of 2010 to 2014.

To delineate RapidRide stop places, RapidRide stops within 500 Euclidean feet of one another were grouped together. This effectively combined RapidRide stops for the same route in the same service location, but serving different directions (e.g., northbound and southbound) and which may be used for the same round trip. The 500-foot threshold was chosen assuming that “paired” RapidRide stops would be no further than about a block apart. Five hundred feet is roughly the sum of a downtown Seattle city block (300 feet) plus two street widths (100 feet). A visual review of the data showed that this worked well in most locations (Figure 1, top right panel).
Stop Place Measures of Ridership and Residential Access

For each RapidRide stop place, weekday ridership was summed for King County Metro bus stops within 1/8 mile (660 feet) along the street network from any individual RapidRide stop that comprised the stop place (Figure 1, middle right panel). If a bus stop was within 1/8 mile of two or more RapidRide stop places, its ridership was assigned to the closest. Operationally, this was executed in ArcGIS 10.2 using the Network Analyst OD cost matrix function to measure the distance from each bus stop to all RapidRide stops within 1/8 mile, then joining the closest bus stops to each RapidRide stop and summing the ridership for all joined stops by RapidRide stop place. Ridership was measured as average weekday boardings and alightings during Spring 2010, Fall 2013, and Fall 2014. The 1/8-mile ridership catchment area was used to capture ridership at bus stops closed or relocated by RapidRide, as well as changes in ridership at bus stops serving nearby routes that may be due to transfers or displacement to RapidRide stops.

Counts of residential units within walking distance of RapidRide stop places were used to control for increased development that often corresponds with BRT implementation.
This variable was measured as the count of residential units within ¼ mile of the closest RapidRide stop place. Operationally, this was executed in ArcGIS 10.2 using hybrid Euclidean-Thiessen buffers to identify the area within ¼ mile of the closest RapidRide stop place (Figure 1, bottom right panel). Euclidean-Thiessen buffers were ¼-mile Euclidean buffers clipped by Thiessen Polygons, whose boundaries defined the area closest to each RapidRide stop relative to all other RapidRide stops. A ¼-mile residential catchment area was used because it is commonly used as a “rule of thumb” walking distance to bus transit (Kittelson & Associates Inc. et al. 2013). Euclidean distances rather than network distances were used because the formal street network may be an incomplete representation of the informal paths that exist for pedestrians to most directly access transit in suburban areas served by high-capacity transit (Moudon et al. 1998). Residential unit data were derived from the King County Assessor’s parcel data for the years 2010, 2013, and 2014. These data included counts of residential units for all residential land uses, including multi-family dwellings such as apartments, condominiums, and mixed-use buildings. Residential units were summed for all parcels that intersected each stop place residential catchment buffer. If a parcel partially intersected a buffer, the proportion of units equal to the proportion of area inside the buffer was counted.

Analysis

A total of 167 RapidRide stop places along the A, B, C, D, E, and F lines were identified. Stop places with no ridership data for any of the three time points were excluded (n=11) because they likely represented places newly served by RapidRide rather than areas where RapidRide replaced existing service. Stop places serving multiple RapidRide lines were also excluded (n=6), because they experienced RapidRide interventions at multiple time points, which would make analysis difficult. Also, however, they represented unique transit hubs (e.g., the downtown bus corridor and the Tukwila International Boulevard Link light rail station), where the effects of RapidRide service could be diluted by other changes to the transit system.

The remaining analytic sample of 150 RapidRide stop places was divided into two groups according to when RapidRide service began (Figure 1, left panel). The first group consisted of stop places serving the A, B, C, and D lines, which all opened between 2010 and 2013. The second group consisted of stop places serving the E and F lines, which opened between 2013 and 2014. Mean stop place ridership and residential units are presented for each group and for each RapidRide line by time period. Absolute and percent changes in mean ridership and residential units were calculated for each of the two time intervals, 2010 to 2013 and 2013 to 2014.

Statistically significant (p<0.05) differences in longitudinal changes in ridership between the two groups were assessed using a mixed effects negative binomial regression model. This model treats ridership at each time period as the dependent variable. The mixed effects component of the regression model accounts for correlation in observations among each stop place over the three time periods. The negative binomial link in the regression model accounts for overdispersion in the distribution of ridership count.
The Causal Effect of Bus Rapid Transit on Changes in Transit Ridership

Data (i.e., count data with many small values but also some very large values, which results in a standard deviation greater than the mean) and results in coefficients that, when exponentiated, take the form of incident rate ratios (IRRs). In this case, IRRs can be interpreted as ratios of ridership among groups that differ by one unit of the dependent variable. Dependent variables include a dummy variable representing group membership (ABCD group = 0, EF group = 1), a categorical time variable (values of 2010, 2013, and 2014), and a categorical interaction term of group by time. Thus, the group membership IRR represents the ratio of ridership among the EF group compared to the ABCD group at 2010; the time IRRs represent the ratio of ridership among ABCD groups at 2013 and 2014 compared to 2010; the group by 2013 interaction term IRR represents the ratio of the change in ridership from 2010 to 2013 among the EF group compared to the ABCD group; and the group by 2014 interaction term IRR represents the ratio of the change in ridership from 2010 to 2014 among the EF group compared to the ABCD group.

The interaction terms are used to test the hypothesis that changes in ridership were greater among stop places that experienced RapidRide intervention compared to stop places that had no change during the same time period. The group by 2013 interaction term directly tests whether the change in ridership from 2010 to 2013 was different among the EF group, which had traditional bus service during this time, compared to the ABCD group, which experienced RapidRide implementation. The linear combination of the group by 2014 interaction term minus the group by 2013 interaction term tests whether the change in ridership from 2013 to 2014 was different among the EF group, which experienced RapidRide implementation during this time period, compared to the ABCD group, which had existing RapidRide service. For interpretability, IRR are presented comparing the group that experienced RapidRide implementation compared to the group that experienced no change.

Models were repeated including residential units as a time-varying control variable to assess whether any changes in ridership were due to corresponding changes in the number of residential units served by each stop place.

Exploratory analyses were conducted to determine if the effect of RapidRide was different for lines serving downtown Seattle compared to lines serving outlying communities. Analyses were repeated separately for the CDE lines serving downtown Seattle and the ABF lines serving the outlying communities. All analyses were conducted using Stata 13.0.

Results

Mean stop place ridership increased along all RapidRide corridors from Spring 2010 to Fall 2013 and, with the exception of the B line, from Fall 2013 to Fall 2014 (Table 1). Both absolute and percent changes in mean ridership from 2010 to 2013 were greater among the ABCD group, during which time RapidRide was implemented, compared to the EF group, which had consistent conventional bus service during that time period. Similarly, both absolute and percent changes in mean ridership from 2013 to 2014 were greater.
among the EF group, during which time RapidRide was implemented, compared to the ABCD group, during which time had consistent existing RapidRide service.

Mean residential units within ¼ mile were slightly greater among the ABCD lines stop places than the EF lines (Table 2). However, changes in residential units were similar among both groups—about a 6% increase from 2010 to 2013 and a 1% increase from 2013 to 2014.

Results from the longitudinal regression model showed no differences in 2010 rates of ridership among the EF group compared to the ABCD group (IRR: 0.94; 95% CI: 0.73, 1.23; p=0.664) (Table 3). Rates of ridership among the ABCD group increased by 88% from 2010 to 2013 (IRR: 1.88; 95% CI: 1.73, 2.05; p<0.001). During the same time period, changes in ridership among the EF group were significantly lower, only 70% that of the change in the ABCD ridership (IRR: 0.70; 95% CI: 0.61, 0.81; p<0.001). From 2010 to 2014, rates of ridership among the ABCD group increased by 107% (IRR: 2.07; 95% CI: 1.90, 2.25; p<0.001), which were not significantly different from changes in ridership among the EF group from 2010 to 2014 (IRR: 0.92; 95% CI: 0.80, 1.06; p=0.232). This is explained by the 31% greater change in ridership from 2013 to 2014 among the EF group compared to the ABCD group (IRR: 1.31; 95% CI: 1.16, 1.49; p<0.001). Controlling for residential units only slightly attenuated the observed changes in ridership.
The Causal Effect of Bus Rapid Transit on Changes in Transit Ridership

TABLE 3.
Mixed Effects Negative Binomial Regression Model of Stop Place Ridership

<table>
<thead>
<tr>
<th></th>
<th>Crude</th>
<th>Adjusted*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IRR (95% CI)</td>
<td>p value</td>
</tr>
<tr>
<td>EF (reference = ABCD)</td>
<td>0.94 (0.73, 1.23)</td>
<td>0.664</td>
</tr>
<tr>
<td>Time: 2010</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td></td>
<td>2013</td>
<td>1.88 (1.73, 2.05)</td>
</tr>
<tr>
<td></td>
<td>2014</td>
<td>2.07 (1.90, 2.25)</td>
</tr>
<tr>
<td>Residential units (100)</td>
<td>1.08 (1.06, 1.10)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>EF X 2013</td>
<td>0.70 (0.61, 0.81)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>EF X 2014</td>
<td>0.92 (0.80, 1.06)</td>
<td>0.232</td>
</tr>
<tr>
<td>EF X 2014 - EF X 2013</td>
<td>1.31 (1.16, 1.49)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

* adjusted for residential units

Translating the model results to directly compare changes in ridership among stop place catchment areas where RapidRide was implemented to those where no change occurred resulted in an estimated 43% increase in ridership compared to consistent traditional bus service and a 31% increase in ridership compared to consistent RapidRide existing service (Table 4). Controlling for the effect of concurrent residential development only slightly reduced these estimates to 35% and 29% increases, respectively. When the sample was stratified by routes serving downtown Seattle and routes serving outlying communities, a stronger effect was observed among routes serving outlying communities.

TABLE 4.
Mixed Effects Negative Binomial Regression Model Results Modified to Compare RapidRide Intervention Group to No Change Group

<table>
<thead>
<tr>
<th>Location</th>
<th>Intervention</th>
<th>Comparison</th>
<th>Comparator</th>
<th>Crude</th>
<th>Adjusted*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IRR (95% CI)</td>
<td>p value</td>
</tr>
<tr>
<td>All</td>
<td>ABCD line BRT implementation</td>
<td>EF line bus service prior to BRT implementation</td>
<td>Change in ridership from 2010 to 2013</td>
<td>1.43 (1.24, 1.65)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>EF line BRT implementation</td>
<td>ABCD lines existing BRT service</td>
<td>Change in ridership from 2013 to 2014</td>
<td>1.31 (1.16, 1.49)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Inside Seattle</td>
<td>CD line BRT implementation</td>
<td>E line bus service prior to BRT implementation</td>
<td>Change in ridership from 2010 to 2013</td>
<td>1.16 (1.00, 1.35)</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td>E line BRT implementation</td>
<td>CD lines existing BRT service</td>
<td>Change in ridership from 2013 to 2014</td>
<td>1.14 (1.01, 1.29)</td>
<td>0.040</td>
</tr>
<tr>
<td>Outside Seattle</td>
<td>AB line BRT implementation</td>
<td>F line bus service prior to BRT implementation</td>
<td>Change in ridership from 2010 to 2013</td>
<td>1.73 (1.33, 2.26)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>F line BRT implementation</td>
<td>AB lines existing BRT service</td>
<td>Change in ridership from 2013 to 2014</td>
<td>1.55 (1.24, 1.94)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

* adjusted for residential units
Discussion

We estimated that implementation of BRT service leads to a 35% increase in transit ridership compared to continued conventional bus service. This estimate more accurately captures the causal effect of BRT on ridership than simple before/after comparisons of ridership along conventional bus routes where BRT is implemented, which appears to be the industry standard (Kittelson & Associates Inc. et al. 2007; Levinson et al. 2003; Parametrix 2014; Peak et al. 2005; US Government Accountability Office 2012). The estimate accounts for temporal trends in ridership, shifts in ridership due to BRT-related service changes, and nearby residential development that may accompany BRT service.

The 35% increase in transit ridership due to BRT implementation compared to continued conventional bus service from Spring 2010 to Fall 2013 was greater than the 29% increase observed when BRT implementation was compared to continued BRT service from Fall 2013 to Fall 2014. This could be due to the longer interval during which BRT implementation was compared to continued conventional bus service (3.5 vs. 1 year). It also could be due to continued gains in ridership during the 2013 to 2014 interval among the BRT lines that were implemented during the 2010 to 2013 interval. In either event, this suggests that major ridership gains from BRT implementation occur immediately, but also continue to accrue years after the service change. Residual longer-term gains in ridership associated with BRT may be due to residential or commercial development that occurs after BRT implementation (US Government Accountability Office 2012) and as people who wish to use transit move closer to the BRT corridor to take advantage of the service. Unfortunately, this analysis cannot pinpoint the precise temporal changes in ridership associated with BRT due to the limited number of time periods during which ridership was observed.

Unsurprisingly, the number of residential units within ¼ mile of stop places was positively associated with ridership. Controlling for change in residential units in the longitudinal analysis attenuated somewhat the effect of BRT implementation on increased ridership. This suggests that some of the increased ridership due to BRT was the result of increased residential density along BRT corridors. Transit planners who wish to get the most out of BRT implementation should work with land use planners to focus transit-oriented development (TOD) along the corridors (Cervero and Dai 2014), as it appears that the increased capacity of BRT is capable of handling the increased residential demand for transit service. The study was limited due to its inability to control for changes in employment density. Employment data at a spatial and temporal resolution suitable for this analysis were not available. It is possible that much of the effect of RapidRide on ridership could be due to employers choosing to locate along these BRT lines.

A stronger effect of BRT implementation was observed for the ABF lines outside of Seattle than for the CDE lines serving downtown Seattle. Ridership for routes outside Seattle were estimated to increase 61% with BRT implementation compared to conventional bus service, whereas ridership for routes serving downtown Seattle were estimated to increase 17%. It may be that BRT is more effective in attracting riders in
The Causal Effect of Bus Rapid Transit on Changes in Transit Ridership

places where transit use is less common or in areas where the initial improvement in service frequency and span was more substantial.

The stepped wedge design employed in this study is a robust alternative to randomized controlled trials—the gold standard study design for estimating a causal effect—when the timing of the intervention is assigned randomly (Bonell et al. 2011). BRT implementation in this quasi-experimental study was not assigned randomly and, therefore, the timing of BRT implementation across corridors could have biased the estimate if BRT was rolled out to correspond with increases in ridership due to exogenous events. This is unlikely, as there were no major commercial developments or infrastructure projects completed in the vicinity of the BRT corridors during this time, and the analysis controlled for residential development.

This analysis used data from King County, Washington, and evaluated RapidRide BRT implementation that rolled out between 2010 and 2014. It may be of limited generalizability to other metropolitan area, BRT systems, or time periods. King County is a major metropolitan area that is largely reliant on bus service for transit. The RapidRide BRT service does not compete with rail transit for riders; in fact, all but one of the RapidRide corridors provide transfer service to the single light rail corridor in the region. Similar increases in ridership may not be realized in major metro areas where BRT must compete with existing, extensive rail transit systems or in smaller cities where transit is less competitive with driving. The RapidRide service includes many of the features commonly found in BRT systems worldwide, such as frequent service and a streamlined entry system, yet it qualifies as BRT lite only due to the lack of dedicated travel lanes and subway-like transit platforms (Cervero 2013). More or less extensive BRT systems may result in greater or lesser changes in ridership. Finally, during the study period King County’s population increased by an estimated 86,000 from 1.93 million to 2.02 million (Office of Financial Management 2016), and median housing prices increased by 16%, from $349,000 to $406,000 (Zillow 2016). BRT that is implemented during periods of slower growth may see smaller changes in ridership.

This study also was limited to the use of average weekday ridership as its single evaluation metric. RapidRide service changes were most dramatic during weekend service periods, and any resulting changes in weekend ridership were not captured in this study. We also did not capture changes in service quality. The increases in ridership associated with RapidRide BRT implementation we observed during weekdays likely were due to a combination of more spacious buses, shorter headways, extended service hours, and more welcoming stop infrastructure—all for the same fare price as traditional bus service. These enhancements would conceivably result in a quicker and more comfortable trip, even for an individual who would have ridden the bus anyway.

Finally, during the study period, King County Metro changed automatic passenger count systems. The older system under-counted by about 3% and the newer system over-counted by about 4%. This means that the changes in ridership over time presented in Table 1 are slightly inflated. However, the primary analysis compared the changes in ridership over time between routes with and without RapidRide
implementation, which would be subject to the same measurement errors over time and thus still result in a robust estimate.

Conclusion

This study used a quasi-experimental stepped wedge study design to assess the effect of incremental RapidRide BRT implementation in King County, Washington. The analysis was intended to add to the evidence for a causal association between BRT implementation and increased transit ridership by accounting for temporal changes in ridership, shifts in ridership to or from other bus routes, and residential development that may correspond with BRT implementation. Independent of these factors, BRT implementation was associated with a 35% increase in ridership compared to consistent conventional bus ridership and a 29% increase in ridership compared to consistent existing BRT service. These estimates should help transit planners develop more reliable estimates of ridership changes due to planned BRT systems and make a stronger argument for the ability of BRT to increase transit ridership and contribute to the mobility and vitality of the urban population they serve.

References


About the Authors

ORION T. STEWART (orions@uw.edu) is an epidemiology Ph.D. candidate at the University of Washington. His research focuses on quantifying the built environment and understanding its association with health-related behaviors.

ANNE VERNEZ MOUDON (moudon@uw.edu) is Professor Emerita of Architecture, Landscape Architecture, and Urban Design and Planning at the University of Washington, where she also directs the Urban Form Lab.

BRIAN E. SAELENS (brian.saelens@seattlechildrens.org) is Professor of Pediatrics and Psychiatry and Behavioral Sciences at the University of Washington and Seattle Children’s Research Institute. His research focuses on environmental and policy factors in obesity prevention and childhood obesity treatment.
Impact of Different Bus Stop Designs on Bus Operating Time Components

Xiaodong Liu, Yao Yang, Meng Meng, Andreas Rau
TUM CREATE Ltd., Singapore

Abstract

The design of bus stops significantly affects bus operation. The delay time caused by inappropriate bus stop design adversely influences the efficiency of the system. This paper aims to examine the influence of bus stops on bus operating time components through statistical analysis, using Singapore as a case study. Two common types of bus stops, bus bay and curb-side stop, were investigated during the field survey to obtain actual data of bus operation at stops. Sixteen stops were chosen in pairs to compare the differences in operating time at bus stops. Bus operating times, including acceleration time, dwell time, deceleration time, and delay time, were recorded, with five types of delay time categorized. A total of 2,653 valid data records were collected and processed. The results showed that buses have better operational performance at curb-side stops than at bus bays in terms of average passenger boarding and alighting time and acceleration time. These findings have operational and planning implications for transport authorities and operators with regard to evaluating the performance of bus operation and improving the design of bus stops.

Keywords: Bus efficiency, bus travel time, dwell time, bus bay, curb-side stop

Introduction

Bus operation efficiency is important to urban traffic systems. It involves macroscopic planning, but operational aspects at the microscopic level are crucial as well. The reduction of bus travel times will improve the quality of service for the passengers and also reduce the operating cost for operators. Travel times can be reduced by mitigating time losses at intersections, introducing dedicated bus lanes, etc. Many studies have been conducted to investigate the operating time for entire bus operation (Shrestha and Zolnik 2013) or specific time components including running time (Surprenant-Legault and El-Geneidy 2011), delay at intersection (Rakha and Zhang 2004) and arrival and dwelling at stops (Yu et al. 2011; Dueker et al. 2004; Tirachini 2013; Sun et al. 2014; Rajbhandari et al. 2003; Zhou et al. 2009).
Bus stop design has a strong influence on the operating time components of buses. To improve the quality of public transport operation, it is important to understand the differences in bus stop designs in terms of operating time components at stops; this forms the research question of this paper. The investigation has important implications for public transport stakeholders, including transport authorities and operators, on operation and planning.

Currently, there are three common types of bus stops: bus bay, curb-side stop, and bus bulb (Fitzpatrick et al. 1996). As shown in Figure 1, bus bays have a dedicated bus-stopping area away from the road lanes used for traveling. These are the prevailing bus infrastructure in many megacities in Asia where bus transit is a major transport mode, including Beijing, Singapore, Hong Kong, and Tokyo. It is generally recommended that bus bays be provided at locations along roads with high traffic volume and arterial roadways with higher cruising speeds (Koshy and Arasan 2008). Curb-side stops and bus bulbs use a marked or signed portion of the through-traffic lanes for the boarding and alighting of passengers. Curb-side bus stops are widely used in many cities and municipalities, such as Auckland, Washington, London, British Columbia, and Tauranga city. A bus bulb (nub) is an extension of the sidewalk from the curb of a parking lane into the edge of the through lane on a road. Thus, bus bulbs have similar performance as curb-side stops.

FIGURE 1.
Selected street-side bus stop design

The bus operating time at a bus stop consists of deceleration time, dwell time, and acceleration time. Deceleration time is the time a bus spends to decelerate from cruising speed to a full stop when approaching a bus stop. Dwell time is the time when a bus dwell at a stop and includes the time needed for doors to open and close and
the time for passengers to board and alight. Acceleration time is the time taken for a bus to leave a stop and merge back into the through lane. Buses may experience delays at all three stages, which leads to additional time needed for them to complete their journeys. Therefore, a fourth time component, delay time, is added in this study to fully understand bus operating time at bus stops. Five common types of delay time are considered in this study:

- **Re-entry delay** – This occurs when a bus leaves the stop but fails to immediately merge back into the through lane due to high traffic volume. It occurs only at bus bays because buses remain in the through lanes while dwelling at curb-side stops.

- **Queuing delay** – During peak hours, it is usual that multiple buses arrive at a stop at the same time. Due to the limited available berths, some buses may have to queue.

- **Boarding and alighting delay** – This type of delay is caused by passengers during the boarding and alighting process. For example, additional time is needed for passengers in wheelchairs to board or alight the bus.

- **Delay caused by stopped or parked vehicles** – When a bus leaves or approaches a stop, it may be obstructed by illegally-parked vehicles and may require additional maneuvers.

- **Delay caused by vehicles queueing in front of a traffic signal** – During peak hours, accumulated vehicle queues from downstream intersections may prevent a bus from leaving a stop, resulting in additional waiting time.

During operation at stops, buses may encounter more than one type of delay. Such situations are also considered in this study.

This paper investigates the differences in bus operating times at bus bays and curb-side stops via a survey in Singapore. The bus system in Singapore accounts for 49% of the 3.75 million passenger trips made by public transport per day (Land Transport Authority 2015). By 2014, there were about 4,700 bus stops island-wide (Land Transport Authority 2015), with two common types: bus bays and curb-side stops. Bus bulbs are not used in Singapore, as discovered during the field observations, due to limited on-street parking.

**Literature Review**

Several surveys in different cities have been conducted to identify the operational differences at bus bays and curb-side stops. All had different survey designs and purposes and, therefore, cannot be compared easily. However, the experiences of other cities could help to develop a better understanding of the operational impact of different bus stop designs. A study in Beijing (Xu et al. 2010) investigated the differences in acceleration, deceleration, and dwell times at bus bays and curb-side stops under optimal operation conditions (no queuing, no re-entry delays, etc.). The results showed an average deceleration time of 9.0s at bus bays and 8.5s at curb-side stops. The average acceleration time ranged from 10.7s at bus bays to 10.9s at curb-side stops. Boarding
and alighting times were not published. Another survey in Beijing in 2013 (Chen et al. 2013) analyzed the situation by including all kinds of possible delays. The average deceleration time increased to 11.1s at bus bays and 9.7s at curb-side bus stops. The study showed an average acceleration time of 11.1s (including re-entry delays) at bus bays and 10.2s at curb-side stops. The boarding and alighting times per passenger ranged from 2.3s to 2.4s at bus bays and from 1.8s to 2.5s at curb-side stops, both depending on the load factor of the bus.

A study in London (Transport for London 2006) analyzed the changes in bus operation by converting three bus bays into curb-side stops. As a result, the re-entry delays were reduced by 3% to 13%. The absolute values of the changes in acceleration, deceleration, and re-entry delays were not published. The original boarding time at the bus bays ranged from 2.6s to 3.8s per passenger and improved by 0.5s to 1s per passenger after the conversion. A survey in Ottawa (Genivar 2011) showed that the deceleration time at bus bays ranges from 1s to 2s and the acceleration plus re-entry delay is 4s to 5s longer than at curb-side stops. This study could not identify any differences in the dwell time per passenger between the two bus stop designs. Wang et al (2016) proposed a method to estimate bus dwell time at a bus stop, where the average boarding and alighting time was explicitly calibrated considering different channel doors. The average deceleration and acceleration time were not estimated.

Table 1 provides a summary of the results from the different surveys. The results show that bus bays have longer deceleration and acceleration delays than curb-side bus stops. All studies except the one in Ottawa indicate longer boarding and alighting times per passenger at bus bays.

<table>
<thead>
<tr>
<th>Study</th>
<th>Bus Stop Design</th>
<th>Average Deceleration Time</th>
<th>Average Acceleration Time</th>
<th>Boarding/Alighting Time per Passenger</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xu, Kwami, &amp; Yang, 2010</td>
<td>Bus bay</td>
<td>9.0s</td>
<td>10.7s</td>
<td>2.1s (single-channel door) 1.7s (double-channel doors)</td>
</tr>
<tr>
<td></td>
<td>Curb-side stop</td>
<td>8.5s</td>
<td>10.9s</td>
<td></td>
</tr>
<tr>
<td>Chen, Zhou, Zhou, &amp; Mao, 2013</td>
<td>Bus bay</td>
<td>11.11s</td>
<td>11.12s</td>
<td>2.22s (load factor &lt; 0.7) 2.37s (load factor &gt;=0.7)</td>
</tr>
<tr>
<td></td>
<td>Curb-side stop</td>
<td>9.74s</td>
<td>10.2s</td>
<td>1.82s (load factor &lt;0.55) 2.49s (load factor &gt;= 0.55)</td>
</tr>
<tr>
<td>Transport for London, 2006</td>
<td>Bus bay</td>
<td>-</td>
<td>3–13% reduction by curb-side stops, compared with bus bays</td>
<td>2.8–3.8s</td>
</tr>
<tr>
<td></td>
<td>Curb-side stop</td>
<td>-</td>
<td>-</td>
<td>0.5–1s faster than bus bay</td>
</tr>
<tr>
<td>Genivar, 2011</td>
<td>Bus bay</td>
<td>1–2s longer at bus bays compared to curb-side</td>
<td>4–5s longer at bus bays compared to curb-side</td>
<td>3.5s, with no difference between curb-side and bus bay</td>
</tr>
<tr>
<td></td>
<td>Curb-side stop</td>
<td>-</td>
<td>-</td>
<td>Boarding time: 2.5–4.0s (single-channel door); 0.6–2.5s (multiple-channel doors) Alighting time: 2.1–3.3 s (single-channel door); 0.5–3.3s (multiple-channel doors)</td>
</tr>
<tr>
<td>Wang et al. (2016)</td>
<td>Bus bay</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Curb-side stop</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>
Field Survey

To analyze the operational differences between bus bays and curb-side stops in Singapore, eight locations were chosen. The survey was conducted at each location by comparing the bus bay and curb-side stop in a “bus stop survey pair” (BSSP), which consisted of two successive stops: a bus bay and a curb-side stop (Figure 2).

To minimize the influence of factors other than the stop design (e.g., traffic volume, driving characteristics of drivers, bus loading factor, etc.) as much as possible, the BSSPs were selected according to the following criteria:

- A BSSP consists of two successive stops: a bus bay and a curb-side stop.
- Both stops have one berth.
- No traffic signals directly near the stops.
- No heavy congestion at both stops.
- Similar numbers of bus service lines at both stops.
- Similar passenger demand, but no crowding, at both stops.

The locations of the BSSPs were chosen to ensure that all criteria listed above were satisfied and that the locations were evenly distributed over Singapore. Table 2 shows a summary of the selected stops. Most of the selected stops are connected to residential areas and are a certain distance from both upstream and downstream intersections.

The survey was carried out during morning peak hours (7:30–9:30 AM) and evening peak hours (5:30–7:30 PM) on weekdays from July 16–31, 2014. Each location was surveyed for two days to gather sufficient data, including:

- Basic bus information, including service number, type of bus (single-decker, double-decker, or articulated bus); bus delay type: re-entry delay, queuing delay, boarding and alighting delay, delay caused by stopped or parked vehicles, and delay by vehicles queuing in front of traffic signal
- Position of bus in the service queue, if berth occupied by other buses
- Bus operating times at stops, including deceleration time, dwell time, acceleration time
- Passenger volume – number of passengers boarding and alighting from each door of bus
TABLE 2. Characteristics of Selected Bus Stop Survey Pairs

<table>
<thead>
<tr>
<th>BSSP</th>
<th>Stop Type</th>
<th>Number of Through Lanes per Direction</th>
<th>Upstream Signalized Intersection Distance*</th>
<th>Downstream Signalized Intersection Distance</th>
<th>Surrounding Areas</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bus bay</td>
<td>2</td>
<td>200m</td>
<td>-</td>
<td>Residential area</td>
</tr>
<tr>
<td></td>
<td>Curb-side stop</td>
<td>2</td>
<td>140m</td>
<td>110m</td>
<td>Open area with pedestrian walk</td>
</tr>
<tr>
<td>2</td>
<td>Bus bay</td>
<td>2</td>
<td>210m</td>
<td>160m</td>
<td>Residential area</td>
</tr>
<tr>
<td></td>
<td>Curb-side stop</td>
<td>2</td>
<td>100m</td>
<td>-</td>
<td>Residential area</td>
</tr>
<tr>
<td>3</td>
<td>Bus bay</td>
<td>2</td>
<td>-</td>
<td>110m</td>
<td>Residential area</td>
</tr>
<tr>
<td></td>
<td>Curb-side stop</td>
<td>2</td>
<td>260m</td>
<td>-</td>
<td>Residential area</td>
</tr>
<tr>
<td>4</td>
<td>Bus bay</td>
<td>2</td>
<td>-</td>
<td>240m</td>
<td>Residential area</td>
</tr>
<tr>
<td></td>
<td>Curb-side stop</td>
<td>2</td>
<td>90m</td>
<td>-</td>
<td>Residential area</td>
</tr>
<tr>
<td>5</td>
<td>Bus bay</td>
<td>2</td>
<td>230m</td>
<td>130m</td>
<td>Residential area</td>
</tr>
<tr>
<td></td>
<td>Curb-side stop</td>
<td>2</td>
<td>190m</td>
<td>-</td>
<td>Residential area</td>
</tr>
<tr>
<td>6</td>
<td>Bus bay</td>
<td>2</td>
<td>180m</td>
<td>180m</td>
<td>Park</td>
</tr>
<tr>
<td></td>
<td>Curb-side stop</td>
<td>2</td>
<td>380m</td>
<td>-</td>
<td>Open area</td>
</tr>
<tr>
<td>7</td>
<td>Bus bay</td>
<td>2</td>
<td>-</td>
<td>140m</td>
<td>Residential area</td>
</tr>
<tr>
<td></td>
<td>Curb-side stop</td>
<td>2</td>
<td>410m</td>
<td>-</td>
<td>Residential area</td>
</tr>
<tr>
<td>8</td>
<td>Bus bay</td>
<td>2</td>
<td>160m</td>
<td>240m</td>
<td>Residential area</td>
</tr>
<tr>
<td></td>
<td>Curb-side stop</td>
<td>2</td>
<td>140m</td>
<td>-</td>
<td>Residential area</td>
</tr>
</tbody>
</table>

*If there is another bus stop between a certain stop and its upstream/downstream intersection, distance marked as "-".

Result and Analysis

With incomplete data and skipped services filtered out, the data collection process led to 2,653 valid data records, comprising 1,256 valid data records at bus bays and 1,397 at curb-side stops. The analysis focuses on three parts: delay type, deceleration and acceleration times, and dwell time. The study on delay type includes basic statistics and their differences between delays at both bus bays and curb-side stops. With the assumption that the deceleration and acceleration times are different at bus bays and curb-side stops, hypothetical tests were applied for further data analysis. Finally, a regression model was applied to investigate the relationship between bus dwell time and the number of passengers boarding or alighting.

Delay-type Analysis

In total, 545 delayed records were collected at bus bays and 274 at curb-side bus stops. These delays were recorded according to the five categories mentioned above. The details of the distributions of the delay types are presented in Table 3.
Impact of Different Bus Stop Designs on Bus Operating Time Components

### TABLE 3.
Distributions of Different Delay Types at Bus Bays and Curb-side Stops

<table>
<thead>
<tr>
<th>Delay Type</th>
<th>Bus Bay</th>
<th>Curb-side Stop</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total valid data records</td>
<td>1,256</td>
<td>1,397</td>
</tr>
<tr>
<td>No Delay</td>
<td>703</td>
<td>1,110</td>
</tr>
<tr>
<td>Delayed</td>
<td>545</td>
<td>274</td>
</tr>
<tr>
<td>Re-entry delay</td>
<td>271</td>
<td>-</td>
</tr>
<tr>
<td>Queuing delay</td>
<td>118</td>
<td>106</td>
</tr>
<tr>
<td>Boarding and alighting delay</td>
<td>33</td>
<td>122</td>
</tr>
<tr>
<td>Delay by stopped or parked vehicle</td>
<td>9</td>
<td>17</td>
</tr>
<tr>
<td>Delay by vehicle queue in front of traffic signals</td>
<td>19</td>
<td>6</td>
</tr>
<tr>
<td>Delay by more than one type or other types</td>
<td>95</td>
<td>23</td>
</tr>
</tbody>
</table>

*Number of delayed buses / total valid data records.

In general, delays occurred more frequently at bus bays than at curb-side stops. The percentage of delayed buses at bus bays was more than two times that of the percentage at curb-side stops (43% vs. 20%). Re-entry delay was the major delay at bus bays. The proportion of bus queuing delays at bus bays was slightly higher than at curb-side stops (9% vs. 8%). The reason might be that re-entry delays at bus bays increase the amount of time a bus occupies the berth. A major delay at curb-side stops was due to passengers boarding and alighting; such delay occurred less frequently at bus bays (9% vs. 3%). Due to parking regulation that prohibits parking within 9 meters of a bus stop in Singapore, bus delays caused by stopped or parked vehicles at bus bays and curb-side stops are low (1% vs. 1%). According to the survey stop selection criteria, the percentage of delays caused by traffic signals is nearly zero. There is a large difference in terms of passenger boarding and alighting delay between bus bays and curb-side stops, as shown in Table 4.

### TABLE 4.
Boarding and Alighting Delay at Bus Bays and Curb-Side Stops

<table>
<thead>
<tr>
<th>BSSP</th>
<th>Average Number of Passengers</th>
<th>Average Number of Passengers (with Boarding and Alighting Delay)</th>
<th>Number of Boarding and Alighting Delays</th>
<th>Average Number of Passengers</th>
<th>Average Number of Passengers (with Boarding and Alighting Delay)</th>
<th>Number of Boarding and Alighting Delays</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pair 1</td>
<td>9.34</td>
<td>13.00</td>
<td>4</td>
<td>3.87</td>
<td>5.91</td>
<td>11</td>
</tr>
<tr>
<td>Pair 2</td>
<td>1.08</td>
<td>2.00</td>
<td>2</td>
<td>2.99</td>
<td>4.50</td>
<td>2</td>
</tr>
<tr>
<td>Pair 3</td>
<td>1.64</td>
<td>1.60</td>
<td>5</td>
<td>4.45</td>
<td>6.08</td>
<td>25</td>
</tr>
<tr>
<td>Pair 4</td>
<td>2.93</td>
<td>4.50</td>
<td>8</td>
<td>10.76</td>
<td>9.54</td>
<td>26</td>
</tr>
<tr>
<td>Pair 5</td>
<td>1.80</td>
<td>1.40</td>
<td>5</td>
<td>2.16</td>
<td>3.11</td>
<td>18</td>
</tr>
<tr>
<td>Pair 6</td>
<td>1.87</td>
<td>5.40</td>
<td>5</td>
<td>2.84</td>
<td>3.25</td>
<td>28</td>
</tr>
<tr>
<td>Pair 7</td>
<td>2.36</td>
<td>3.33</td>
<td>3</td>
<td>2.16</td>
<td>3.86</td>
<td>7</td>
</tr>
<tr>
<td>Pair 8</td>
<td>3.22</td>
<td>2.00</td>
<td>1</td>
<td>1.75</td>
<td>3.40</td>
<td>5</td>
</tr>
<tr>
<td>Overall</td>
<td>2.56</td>
<td>4.42</td>
<td>33</td>
<td>3.56</td>
<td>5.45</td>
<td>122</td>
</tr>
</tbody>
</table>
As shown in Table 4, the passenger demand at the bus bay and curb-side stops within one survey pair was not always similar, which could be one of the reasons for the difference in the boarding and alighting delays. Table 4 includes the overall average number of passengers boarding and alighting from each bus, the average number of passengers boarding and alighting from each bus with boarding and alighting delay, and the number of boarding and alighting delays encountered by each stop. It is possible that a high passenger volume at stops could increase the probability of encountering boarding and alighting delays—for example, at pairs 3, 4, 5, and 6. There are also some exceptions—for example, the bus bay in pair 1 has a larger passenger volume but a lower number of boarding and alighting delays. However, the survey data do not support any further investigation of this issue. According to daily observation, passengers waiting at curb-side stops were found to be better distributed in the waiting area or even outside. This could be another reason for the increased boarding and alighting delay at curb-side stops.

**Deceleration and Acceleration Times**

The data analysis of deceleration time was conducted for all recorded buses without any additional delay during the deceleration stage. The deceleration times of buses that could not enter the bus stop because it was occupied by another bus were not included. The results show an overall mean deceleration time of 8.84s at bus bays and 8.53s at curb-side stops. Table 5 and Figure 3 show the results of the mean deceleration times and their standard deviations for all BSSPs.

<table>
<thead>
<tr>
<th>BSSP</th>
<th>Bus Bay</th>
<th></th>
<th>Curb-side Stop</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean [s]</td>
<td>STD* [s]</td>
<td>Mean [s]</td>
<td>STD [s]</td>
<td></td>
</tr>
<tr>
<td>Pair 1</td>
<td>10.04</td>
<td>1.70</td>
<td>10.41</td>
<td>1.60</td>
</tr>
<tr>
<td>Pair 2</td>
<td>8.71</td>
<td>0.96</td>
<td>8.00</td>
<td>0.67</td>
</tr>
<tr>
<td>Pair 3</td>
<td>7.88</td>
<td>1.08</td>
<td>8.45</td>
<td>1.31</td>
</tr>
<tr>
<td>Pair 4</td>
<td>8.81</td>
<td>0.90</td>
<td>8.61</td>
<td>0.90</td>
</tr>
<tr>
<td>Pair 5</td>
<td>8.88</td>
<td>1.10</td>
<td>7.99</td>
<td>0.99</td>
</tr>
<tr>
<td>Pair 6</td>
<td>8.19</td>
<td>1.14</td>
<td>8.06</td>
<td>0.93</td>
</tr>
<tr>
<td>Pair 7</td>
<td>9.02</td>
<td>1.17</td>
<td>8.83</td>
<td>1.29</td>
</tr>
<tr>
<td>Pair 8</td>
<td>9.60</td>
<td>0.90</td>
<td>9.01</td>
<td>0.92</td>
</tr>
<tr>
<td>Overall</td>
<td>8.84</td>
<td>1.29</td>
<td>8.53</td>
<td>1.25</td>
</tr>
</tbody>
</table>

STD: Standard deviation
According to the results, the deceleration times vary among locations. In addition, within one survey pair, the average deceleration times and their standard deviations differ. For example, at pairs 1 and 3, a higher average deceleration time was observed at curb-side stops, whereas the other pairs show different results. A normality test was conducted, and the results show that a normal distribution is not plausible for the deceleration times. A two-tailed Wilcoxon signed rank test was applied to determine if there was any difference between bus bays and curb-side stops in terms of deceleration times and standard deviations (Rice 2007). The result \( T = 9 > 3 = T^* \) shows that there is no significant difference for both deceleration times and standard deviation at the 0.05 significance level.

As with the deceleration times, the acceleration times were analyzed for bus bays and curb-side stops. It is very difficult to differentiate between acceleration times and re-entry delays, since the acceleration will be slow if the bus cannot re-enter the road because of running traffic. Therefore, both time components were combined. Additionally, the acceleration times without any re-entry delay at bus bays were filtered and studied separately. The results are shown in Table 6 and Figure 4.

<table>
<thead>
<tr>
<th>BSSP</th>
<th>Bus Bay</th>
<th></th>
<th>Bus Bay, without Re-entry Delay</th>
<th></th>
<th>Curb-side Stop</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean [s]</td>
<td>STD* [s]</td>
<td>Mean [s]</td>
<td>STD [s]</td>
<td>Mean [s]</td>
<td>STD* [s]</td>
</tr>
<tr>
<td>Pair 1</td>
<td>11.99</td>
<td>3.05</td>
<td>10.61</td>
<td>1.21</td>
<td>11.98</td>
<td>1.68</td>
</tr>
<tr>
<td>Pair 2</td>
<td>12.94</td>
<td>4.12</td>
<td>10.50</td>
<td>1.14</td>
<td>10.58</td>
<td>1.26</td>
</tr>
<tr>
<td>Pair 3</td>
<td>9.66</td>
<td>2.08</td>
<td>9.26</td>
<td>1.57</td>
<td>9.79</td>
<td>2.17</td>
</tr>
<tr>
<td>Pair 4</td>
<td>10.49</td>
<td>2.26</td>
<td>9.74</td>
<td>1.13</td>
<td>9.77</td>
<td>1.65</td>
</tr>
<tr>
<td>Pair 5</td>
<td>11.76</td>
<td>5.23</td>
<td>9.25</td>
<td>1.36</td>
<td>8.75</td>
<td>1.60</td>
</tr>
<tr>
<td>Pair 6</td>
<td>11.07</td>
<td>2.33</td>
<td>10.15</td>
<td>1.20</td>
<td>9.45</td>
<td>1.46</td>
</tr>
<tr>
<td>Pair 7</td>
<td>13.35</td>
<td>5.12</td>
<td>10.01</td>
<td>1.73</td>
<td>10.16</td>
<td>2.07</td>
</tr>
<tr>
<td>Pair 8</td>
<td>10.28</td>
<td>3.05</td>
<td>8.97</td>
<td>1.35</td>
<td>9.02</td>
<td>1.75</td>
</tr>
<tr>
<td>Overall</td>
<td>11.35</td>
<td>4.32</td>
<td>8.84</td>
<td>1.29</td>
<td>9.73</td>
<td>1.91</td>
</tr>
</tbody>
</table>

STD: Standard deviation
Given that normality is not plausible for acceleration times, the one-tailed Wilcoxon signed rank test shows that the acceleration times at bus bays are longer than the ones at curb-side stops, with a statistical significance at the 0.05 significance level. This is caused by the large proportion of acceleration times at bus bays facing re-entry delays (17%, Table 1) during the acceleration stage. The greater mean acceleration times (11.35s vs. 9.73s at curb-side stops) and standard deviations at bus bays (4.32s vs. 1.91s at curb-side stops) decrease the operational efficiency and may delay the successive buses as well.

Using the two-tailed Wilcoxon signed rank test (0.05 significance level) to compare the acceleration times at bus bays without re-entry delay with the acceleration times at curb-side stops shows that there is no statistically-significant difference between both bus stop designs. This indicates that the re-entry delay during the acceleration is the critical time component differentiating the operational efficiency of bus bays and curb-side stops.

The study results show that there is no significant difference between bus bays and curb-side stops in terms of deceleration times. They fall into a stable range with an average value of 8.68s and a standard deviation of 1.28s among all stops regardless of the design type. On the other hand, there is a significant difference in acceleration times between bus bays and curb-side stops. The overall average acceleration time and the standard deviation at bus bays are 11.35s and 4.32s, respectively. These times are longer than at curb-side stops, with 9.73s and 1.91s, respectively. The difference is caused mainly by the frequently observed re-entry delays when buses are leaving the bus bays.

**Dwell Time**

The dwell time of each bus, together with the number of boarding and alighting passengers, was collected to analyze the differences in the average boarding/alighting time per passenger with respect to the different bus stop designs.
Due to the operation requirements for boarding and alighting, passengers can board a bus from the front door only, but can alight from any door. Assuming that the boarding and alighting times per passenger are similar, the following equations could be used to estimate the average boarding and alighting time per passenger:

\[ T = t \times P + C \]  
\[ P = \max \{ (B + A_1), A_2, (A_3) \} \]

where

- \( T \) is the total dwell time. It is dependent on the number of passengers boarding and alighting, average boarding or alighting time per passenger, and some constant time component which includes door opening and closing time.
- \( t \) is the boarding/alighting time per passenger.
- \( P \) is the maximal number of passengers boarding and alighting at one door.
- \( C \) is the unused dwell time, which quantifies the sum of the time gaps between bus arrival, passenger boarding and alighting, and bus’ departure.
- \( B \) is the number of passengers boarding from the front door.
- \( A_i \) is the number of passengers alighting from the \( i^{th} \) door of the bus; \( i = 1 \) for the front door, \( i = 2 \) for the rear door of single-deckers and double-deckers or the middle door of the articulated buses, and \( i = 3 \) for the rear door of the articulated buses.

Because the boarding/alighting time is related to the design of the buses, the study was conducted per different bus type. The data from all valid bus survey results without passenger boarding/alighting delays were used for the calculation of the average boarding and alighting time per passenger. The results of the linear regression model with outlying residuals filtered are shown in Table 7.

<table>
<thead>
<tr>
<th>Stop Design</th>
<th>Bus Type</th>
<th>Average Boarding/Alighting Time per Passenger [s]</th>
<th>Unused Dwell Time [s]</th>
<th>( R^2 )</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bus Bay</td>
<td>Single-decker</td>
<td>1.43</td>
<td>5.68</td>
<td>0.70</td>
<td>519</td>
</tr>
<tr>
<td></td>
<td>Double-decker</td>
<td>1.62</td>
<td>5.20</td>
<td>0.85</td>
<td>202</td>
</tr>
<tr>
<td></td>
<td>Articulated bus</td>
<td>1.48</td>
<td>6.49</td>
<td>0.76</td>
<td>142</td>
</tr>
<tr>
<td></td>
<td>Overall</td>
<td>1.52</td>
<td>5.60</td>
<td>0.77</td>
<td>863</td>
</tr>
<tr>
<td>Curb-side Stops</td>
<td>Single-decker</td>
<td>1.34</td>
<td>6.37</td>
<td>0.64</td>
<td>557</td>
</tr>
<tr>
<td></td>
<td>Double-decker</td>
<td>1.19</td>
<td>6.99</td>
<td>0.64</td>
<td>287</td>
</tr>
<tr>
<td></td>
<td>Articulated bus</td>
<td>1.30</td>
<td>7.64</td>
<td>0.95</td>
<td>121</td>
</tr>
<tr>
<td></td>
<td>Overall</td>
<td>1.33</td>
<td>6.52</td>
<td>0.86</td>
<td>963</td>
</tr>
</tbody>
</table>

As shown in Figure 5, the average boarding/alighting times per passenger at bus bays are always larger than at curb-side stops, by 14% overall (0.2s). The major reason is that the bus bays require drivers to make an additional maneuver to approach the curb. This results in a gap between the bus and the curb and requires the passenger to make additional steps onto the road before boarding and after alighting from the bus.
Bus bays usually have slightly shorter unused dwell times than curb-side stops—less than 1s (Figure 6). For all the data records used for this calculation, the average P value (Equations 1 and 2, the maximal number of passengers boarding and alighting at one door) is 3.4 at bus bays and 4.2 at curb-side stops. As the unused dwell time quantifies the sum of the time gaps between bus arrival, passenger boarding and alighting, and bus departure, the difference could have multiple reasons, including passenger volume, passenger distribution at stops, driver behaviour during arrival and departure, etc. The result is in line with the result of the share of different delay types in Table 3, which shows that curb-side stops have a higher percentage of boarding and alighting delays. This seems to be a specific result of this survey and should not be generalized. One possible explanation could be the larger passenger volume at the curb-side stops than at the bus bays in five of the BSSPs. More passenger boarding and alighting increases the possibility of longer unused dwell times. Further investigation into the reasons is beyond the scope of this study.
As a result of the delays, bus bays require longer boarding/alighting times per passenger for all bus types. This characteristic will increase the dwell time at bus bays compared with curb-side stops with similar passenger demand. Hence, the operational efficiency of bus bays in terms of dwell time is lower than that of curb-side stops.

Conclusion and Discussion

This paper aimed to compare the operational differences of bus bays and curb-side stops, using Singapore as a case study. A survey was conducted at eight different locations across Singapore to collect the bus operating time components, including encountered delay types, deceleration times, dwell times, passenger volumes, and acceleration times. The results show that bus bays are twice as likely to encounter delays than curb-side stops. This difference is caused mainly by re-entry delays during departure from the bus bays.

Compared to other surveys (TABLE 1), the survey conducted in Singapore provides a comprehensive data set on bus operation at stops. In terms of average acceleration time and deceleration time, international results show deceleration takes from 8.5s to 11.11s and acceleration takes from 10.2s to 11.12s. In Singapore, deceleration takes, on average, 8.84s at bus bays and 8.53s at curb-side stops, and average acceleration takes 11.35s at bus bays and 9.73s at curb-side stops. Both deceleration and acceleration times are slightly lower than the corresponding international survey results. The analysis shows that there is no statistically-significant difference between the deceleration times of the two designs. However, the survey also shows a great increment in acceleration times at bus bays; this result is in line with those of the other surveys. It can be concluded that commonly-observed re-entry delay is the major defect of bus bays.

As for the dwell time and average boarding and alighting time per passenger (shown in Table 7), the survey shows that curb-side stops require shorter boarding and alighting times (1.33s) than bus bays (1.52s). But the average boarding and alighting time per passenger in Singapore is shorter than those of all the other surveyed cities (ranging from 1.7s to 4.0s, Table 1). This is mainly because all buses are equipped with double-channel doors in Singapore. Additionally, this study reveals that in Singapore, boarding and alighting at bus bays is shorter than at curb-side stops. The numeric results, in comparison with international results, show that all operating time components of buses at stops vary from location to location. Thus, when planning public transport operation, it is recommended that agencies and operators conduct local surveys to carry out best-fitting local operational plans.

Curb-side stops, compared with bus bays, have similar deceleration times, shorter and more reliable acceleration times due to the absence of re-entry delays, and require less time for passengers to get on or off the bus. In terms of operational efficiency, curb-side stops, therefore, have better performance. However, they also have slightly longer unused dwell times than bus bays. The exact reasons for this will be studied via specific surveys in the future.
In practice, both types are commonly used. Curb-side stops have better efficiency but affect private traffic by blocking an entire lane. On the other hand, bus bays have less impact on private cars and guarantee better safety (Fitzpatrick et al. 1996); as trade-offs, they require longer operating times and reduce efficiency. As a result, choice of bus stop design should consider different aspects, including traffic volume, passenger demand, operational requirements (trunk service or feeder service), etc. For public transport prioritization, curb-side stops are suggested. Bus bays are suggested to be applied along major arterials with high-speed movements for safety reasons.

Acknowledgment

The content of the paper and any opinions expressed are the sole responsibility of the authors. This work was financially supported by the Singapore National Research Foundation under its Campus for Research Excellence and Technological Enterprise (CREATE) program. The data collection in this study was conducted by second author Yao Yang for her master’s thesis.

References


About the Authors

**Xiaodong Liu** (xiaodong.liu@tum-create.edu.sg) is with TUM CREATE Ltd. in Singapore.

**Yao Yang** (yao.yang@tum-create.edu.sg) is with TUM CREATE Ltd. in Singapore.

**Meng Meng** (meng.meng@tum-create.edu.sg) is with TUM CREATE Ltd. in Singapore.

**Andreas Rau** (andreas.rau@tum-create.edu.sg) is with TUM CREATE Ltd. in Singapore.
Measuring the Accuracy of Bus Rapid Transit Forecasts

John Perry

Abstract

The research of Dr. Bent Flyvbjerg in the 1990s and early 2000s showed that urban rail projects often cost more than estimated and carried fewer riders than projected, a troubling trend suggesting that the forecasts for urban rail projects were too optimistic in terms of cost and ridership. Inspired by that research, this analysis seeks to extend that framework to analyze Bus Rapid Transit (BRT). A study of forecast vs. actual costs and ridership was conducted for 19 BRT projects in the United States. From this, it was found that the cost projections for these projects tended to be quite accurate, but ridership projections tended to be quite inaccurate and showed a clear tendency towards an optimism bias. As BRT becomes a more common choice for rapid transit investment in the US, this analysis suggests that current ridership forecasting methods still leave much to be desired.

Keywords: Forecasting, inaccuracy, bus rapid transit, BRT

Introduction

When it was published, the research of Dr. Bent Flyvbjerg illustrated a consistent inaccuracy of urban rail transit project forecasts in regards to cost and ridership. Compared to road and highway projects, which showed a relatively even distribution in terms of inaccuracy between those projects that overestimated and underestimated the costs and usership of the finished project, urban rail projects were consistently projected under the actual cost and above the actual ridership. Based on these findings, Flyvbjerg called for greater accountability over forecasts of large-scale transportation projects (Flyvbjerg et al. 2005).

In the past several years, BRT has become an increasingly common choice for city and local governments when investing in rapid transit. At a time when many transit agencies are experiencing budget cuts and service reductions, the relatively low capital costs of BRT make it an attractive option for transit agencies looking to expand their rapid transit services. In 2014 alone, there were 24 BRT projects completed or under construction in the US (Freemark 2014).
Compared to the megaprojects and urban rail projects that Flyvbjerg studied, these BRT projects seem relatively inexpensive. But for many small and mid-size American cities, a BRT project may represent one of the most significant investments in public transportation in that community’s history. As such, it is just as important that planners strive for accuracy when projecting cost and ridership.

Much has been written about the characteristics of BRT and the current state of BRT development, and there is a substantial body of research regarding the economic and development impacts of a BRT system, but there is much less research regarding BRT ridership and cost forecasts. Also, although there is a large body of work on travel forecasts, most obviously Flyvbjerg’s research, most of it pertains specifically to freeway or urban rail megaprojects. Given the growing interest in BRT systems among local governments and transportation professionals, this is a gap in the current research.

**Literature Review**

There is a large and growing body of research on BRT, most of which generally fits into one of three categories: the basic characteristics of BRT, the current state of BRT development, and the economic and development impacts of BRT. In each of these categories, there is much research comparing the characteristics and impacts of BRT to those of other forms of rapid transit, particularly light rail.

A substantial amount of work has been written about the service characteristics of urban transportation modes, including BRT, and this work offers useful parameters for defining BRT systems and the opportunities for its development. This research attempted to define the level of service characteristics at which point a system can be considered “bus rapid transit” and found that BRT is growing in popularity due to its cost effectiveness and the fact that it can be adapted for use in conventional bus systems (Vuchic 1992; Jarzab et al. 2002; Levinson et al. 2002). Additionally, the Institute for Transportation Policy (ITDP), a non-profit organization that provides technical assistance on public transportation projects and advocates for BRT development, has developed a ranking system for comparing BRT systems and determining whether a system meets their standards for what can be considered “true” BRT (Weinstock et al. 2011). This research proved useful for understanding the characteristics of BRT and what could, even nominally, be considered as such.

Flyvbjerg’s work served as a model for this work. Building on a body of work from the 1990s and early 2000s, his research examined the accuracy of cost and ridership forecasts for large transportation projects, particularly urban rail and road projects, and found that urban rail projects frequently exhibited large cost overruns and typically presented very optimistic ridership forecasts compared with road projects, due, in part, to poor forecasting methods and to biases on the part of the forecasters to promote rail projects (Flyvbjerg et al. 2005); however, none of Flyvbjerg’s work dealt with bus transit projects. More recently, the work of Robert Bain has contributed significantly to this field of research, calling to attention widespread inaccuracy and optimism bias in traffic forecasts for toll road projects (Bain 2009).
Another resource dealing with cost and ridership projections was a series of before-and-after studies conducted by the Federal Transit Administration (FTA) regarding transit projects that have received New Starts funding (FTA 2006–2016). These studies also document a tendency to underestimate final construction costs, with the accuracy of ridership projections varying widely. However, these studies deal mostly with urban and commuter rail projects, with only a couple of bus transit projects included. This general lack of attention towards cost and ridership projections for BRT projects is a gap in the current literature.

Methodology

Funding and ridership information is collected by FTA, but most of this information deals with transit agencies as a whole rather than individual infrastructure projects. In cases in which cost and ridership estimates from the transit agency operating the BRT system were unavailable, the necessary information was obtained from media articles and government reports. BRT systems in operation in the US for which reliable and comparable data were not available are not included in this analysis. In total, 19 projects were included in the final analysis.

The methodologies Flyvbjerg employed in his research served as a guide for this research. Many projects go through multiple forecasts that change as a project moves forward through the design and construction phases. In his research, Flyvbjerg used the project forecasts from the time of the decision to build, arguing that this is the information available to decisionmakers when they agree to move forward on a project and, thus, are the most influential in determining the worthiness of a project. These figures were then compared to the actual figures from the completion of the project to determine their accuracy (Flyvbjerg et al. 2005). This is, in brief, what was attempted in this analysis, using figures as close to the time of the decision to build as were available.

Given the small number of what the ITDP would refer to as “true” BRT systems in the US—that is, systems that have all or nearly all of the features of BRT, such as dedicated lanes, pre-boarding payment, limited stops, and branded service—this analysis also includes some so-called “BRT-lite” systems (those with only a few of the features of BRT) and busway projects to produce a statistically-significant sample of BRT projects. Although not all of these systems fall into what the ITDP would deem “true” BRT, they do all represent significant investments in public transportation for each of the cities included here, so their value for an analysis of BRT cost and ridership projections should not be dismissed.

Individual Case Studies

Note that all dollar values are adjusted to values in the year of expenditure.
Cleveland, Ohio

The HealthLine is a 7.1-mile BRT line operated by the Greater Cleveland Regional Transit Authority (GCRTA). It opened in October 2008 and features exclusive bus lanes and median stations for 4.4 miles of the line, with the remaining 2.7 miles using mixed-traffic curb lanes and sidewalk stops. The HealthLine also uses distinctive station structures, off-board fare equipment, signal priority for vehicles at traffic intersections (FTA 2012). ITDP gave the HealthLine the highest ranking of any BRT system in the US, indicating that they consider it the most complete example of BRT in the US (Weinstock et al. 2011).

Projected capital costs for the HealthLine were $273.4 million according to a 1995 estimate; however, further revisions set estimates ranging from $248.2 million to $317.4 million. Due to cost-effectiveness requirements to receive federal funding, GCRTA trimmed costs for design elements and vehicle procurement and used management tools to monitor the project budget (GCRTA 2012). In the end, the actual capital costs were $197.2 million (GCRTA 2012).

Projected ridership for the HealthLine initially was 21,100 average weekday trips, although this later was revised to 13,500 (FTA, 2012). Actual ridership on the line was 14,300 average weekday trips as of 2012, well below the initial projection. At the time of opening, Cleveland was in the midst of a substantial contraction of the regional economy and a subsequent drop in system-wide transit ridership, which fell by 22% between 2007 and 2010 (FTA 2012). This may go some way towards explaining why the actual ridership was so far below early estimates.

El Paso, Texas

The Brio Mesa Corridor is an 8.6-mile BRT-lite line operated by Sun Metro. Completed in October 2014, the Mesa Corridor route was the first of four planned Brio routes scheduled to open within the coming years and is currently the only rapid transit service in El Paso (Sun Metro 2014). The line features branded and landscaped stations, pre-boarding fare payment, and traffic signal priority.

Projected capital costs for the line were $27.08 million (FTA 2010), with actual capital costs reported at $27.1 million (Sun Metro 2014). Projected ridership initially was 11,900 average weekday boardings within the opening year (FTA 2010), but this was revised to around 3,000 riders per day prior to Brio’s opening. Actual ridership was below even this lowered revision, with 52,000 average monthly boardings as of July 2016 (Wilcox 2016), an average of about 2,000 riders each operating day (Brio currently operates only six days per week).

Escondido, California

The Breeze Rapid is a 6-mile BRT-lite service operated by the North County Transit District (NCTD). Service began in June 2011 and features queue jump lanes at select intersections, traffic signal priority, bus station improvements, and branded service.
Projected capital costs for the system were $2.79 million, according to a NCTD 2006 concept study, with actual capital costs at $4.21 million. The Breeze Rapid came in over budget due, in large part, to revised plans for one intersection, at which a proposed queue jump lane was extended that required widening of the roadway for a full block to accommodate the new lane.¹

No formal ridership projection for the service was conducted; as the first instance of a rapid bus service in the region, the San Diego Association of Governments (SANDAG), the regional planning agency for San Diego County, treated it as a technology and concept demonstration that upgraded an existing local bus route rather than a project to significantly boost ridership.²

### Eugene, Oregon

The Emerald Express (EmX) is a BRT system operated by the Lane Transit District (LTD). An initial 4-mile segment opened in January 2007, with a 7.8-mile extension opening in January 2011. The EmX uses dedicated bus lanes for nearly 60% of its route, with a traffic signal priority system and branded stations with raised platforms, pre-boarding fare payment, and real-time bus arrival signs. The system has been noted as a BRT success story by the ITDP (Weinstock et al. 2011), constructed within budget and with ridership exceeding stated expectations.

Projected capital costs were $24.6 million for the initial segment and $43.1 million for the extension. These proved to be reasonably accurate, with actual capital costs being $24.6 million for the initial segment and $41.3 million for the extension.³

It was projected that the served corridor would see an increase of 40% over the first 20 years of service. In fact, there was a ridership increase of 63% in the corridor over the first year of service and 122% over the first four years. The 2011 extension was projected to increase ridership along the EmX line by 3,700 additional weekday boardings. After completion, the extension reached 80% of this projection within one year and exceeded it in the second year of service. Currently, weekday ridership averages 10,000+ during the school year and exceeds 11,000 in some months, 135% of the estimate predicted for the line after completion of the extension.⁴

### Fort Collins, Colorado

MAX is a 5-mile BRT line operated by Transfort. The line opened in May 2014 and features a dedicated transit-only busway for most of the route, branded service, pre-boarding fare payment, and platform-level boarding at all stations along the route.

¹ Information obtained through personal communication with D. Veeh, December 30, 2013.
² Ibid.
³ Information obtained through personal communication with A. Vobora, February 8, 2014.
⁴ Ibid.
Projected capital costs for the system were $81.98 million (FTA 2009), with actual capital costs coming in at $86.83 million (Duggan 2014). Projected ridership was 3,900 average weekday boardings (FTA, 2009). MAX managed to exceed this figure after its first year of service, with 4,680 average daily boardings in September 2015 (de la Rosa 2015).

**Grand Rapids, Michigan**

The Silver Line is a 9.8-mile BRT-lite service operated by the Interurban Transit Partnership. The line opened in August 2014 and features stations with a sidewalk snowmelt system, next bus signage, platform-level boarding, pre-boarding fare payment, and designated bus-only travel lanes along portions of the route during peak weekday travel periods.

Projected capital costs were $37 million (FTA 2010), with actual capital costs being $40 million (Krietz 2014). Projected ridership was 7,200 average weekday boardings in the first year (FTA 2010). Actual ridership fell well below this mark, with only 2,300 average weekday boardings as of March 2016 (Khut 2016).

**Kansas City, Missouri**

The Troost Avenue MAX is a 13-mile BRT-lite service operated by the Kansas City Area Transportation Authority (KCATA). Following the successful implementation of the MAX bus rapid transit line along Main Street in July 2005, KCATA began pursuing implementation of a second line along Troost Avenue, roughly one mile west of and parallel to the existing Main Street MAX line. As of 2007, the Main Street MAX line had resulted in a 20% growth in ridership along the Main Street corridor, and planners expected similar results from the new Troost Avenue line (FTA 2007). The Troost Avenue MAX opened in January 2011 and features dedicated bus lanes, traffic signal prioritization, branded buses and stations.

Projected capital costs for the line were $30.73 million (FTA 2007), with actual capital costs at $30.6 million (KCATA 2010). Projected ridership was 9,000 average weekday boardings after the first year of service (FTA 2007). This proved to be reasonably accurate, with actual ridership at 8,500 average weekday boardings following one year of service (KCATA 2012).

**Las Vegas, Nevada**

Since opening the MAX BRT line along North Las Vegas Boulevard in 2004, RTC Transit has developed an extensive express bus network that extends across much of the city, with multiple lines that incorporate varying features of BRT service. Notable additions to the system include the launch of the Strip & Downtown Express (SDX) service in 2010, the Boulder Highway Express service in 2011, and the Sahara Express service in 2012. Of these four lines, only the Sahara Express had reliable documentation of cost and ridership projections readily available for this analysis.
The Sahara Express is a 12-mile BRT line that opened in May 2012 and features sheltered stops with raised-level boarding, dedicated bus lanes along most of the route, double-decker buses, traffic signal priority, and landscaping improvements and widened sidewalks along the corridor. Projected capital costs were $43.56 million (RTCSNV 2009), with actual capital costs at $45.2 million. Projected ridership was 13,900 average daily boardings in 2013 (RTCSNV 2009). Actual ridership fell short of this projection, with only 10,000 average daily boardings in 2012 (Christensen 2012), with those figures remaining largely consistent through reported figures in April 2013 and April 2014 (RTCSNV 2014).

**Los Angeles, California**

The Orange Line is a BRT line operated by the Los Angeles County Metropolitan Transportation Authority, better known as Metro. An initial 14-mile segment opened in October 2005, with a 4-mile extension opening in June 2012. The Orange Line is regarded as one of the first “true” BRT systems in the US and uses a dedicated roadway along a former Southern Pacific Railroad branch line through the San Fernando Valley. The line features dedicated stations, pre-boarding fare payment, and a bikeway along the initial segment.

Projected capital costs were $340.4 million for the initial segment (Metro 2003) and $135 million for the extension (Guccione 2006; Callaghan and Vincent 2007). Actual capital costs were $323.6 million for the initial segment, with an extra $26 million for an additional station that opened in December 2006 (Callaghan and Vincent 2007) and $154 million for the extension (Bloomekatz 2012).

Projected ridership was 5,000 to 7,500 average weekday boardings for the first year of service and 22,000 average weekday boardings by 2020 on the initial segment (Callaghan and Vincent 2007). Actual ridership far outpaced these projections, with 21,828 average weekday boardings in May 2006 (Callaghan and Vincent 2007), a figure that grew to nearly 24,000 average weekday boardings in October 2010 and 26,614 average weekday boardings in October 2011 (Hymon 2012). Prior to completion of the extension, Metro projected that the entire line including the extension would carry 45,000 daily riders by 2030 (Anderson 2012); whether this goal will be met remains to be seen, but average weekday Orange Line ridership rose from 26,670 in May 2012, one month prior to completion of the extension, to 31,780 in October 2013.\(^5\)

**Minneapolis, Minnesota**

The METRO Red Line is an 11-mile BRT-lite service operated by Metro Transit. It opened in June 2013 and uses bus-only shoulder lanes between the Twin City suburbs of Bloomington and Apple Valley, with plans for a further extension south to the community of Lakeville.

The project was faced with construction setbacks, and the phasing of the project changed over time. The original plan was to implement the full Bloomington-Lakeville

---

\(^5\) Information obtained through personal communication with D. Mieger, December 18, 2013.
project in four phases with various elements of capital and operating investment in each phase. However, the most recent plan changed the investment strategy to three phases and altered the timing of the elements included in each stage. Operational costs for the system were cut back during construction, causing the initial roll-out of the system to be scaled back to a less-frequent service than originally planned. Projected capital costs under this revised plan were $118 million, with the actual capital cost at $112 million. Projected ridership was initially 2,250 average daily boardings, with a revised figure of 960 average daily boardings in the first year of service, following the reduction in service plans. Actual ridership was 975 average daily boardings in August 2014 (Van Berkel 2014), comparable to the revised figure but well below the initial projection.

**Pittsburgh, Pennsylvania**

The West Busway is a 5-mile dedicated busway used by the Port Authority of Allegheny County. Completed in 2000, the busway was originally planned to be 8.1 miles long and projected to cost $328.8 million to build (FTA 2003). However, the estimate rose to $515 million following issues with land acquisition from freight rail company CONRAIL and problems with the development of a proposed new HOV bridge over the Monongahela River into Downtown Pittsburgh. Ultimately, the CONRAIL land acquisition and proposed bridge elements were abandoned from the plan, and the project was scaled down from 8.1 miles to 5 miles, bringing the actual capital costs for the revised project down to $326.8 million (FTA 2003), technically within the projected cost for the project but only after these significant changes to the proposal were made. Projected ridership for the busway was 7,000 riders per day (FTA 2003), with actual weekday ridership being more than 8,700 riders in October 2002 (FTA 2003) and growing to a peak daily ridership of 10,000 in 2004 (Vincent 2004) before leveling off in later years.

Another Port Authority of Allegheny County busway project in Pittsburgh was the East Busway Swissvale extension, a 2.3-mile extension of the East Busway that was completed in 2003. Projected capital costs for the extension were $62.8 million (FTA 1998), with actual capital costs being $68.8 million (Grata 2003). Projected ridership for the project was an additional 3,800 daily riders on the East Busway by 2005 (FTA 1998). Instead, there were only an additional 2,000 daily riders as of 2004 (Vincent 2004), with ridership falling since, dropping from an average of 30,000 daily riders in 2004 (Vincent 2004) to 25,600 in 2011 (Weinstock et al. 2011). The fall in ridership along the East Busway occurred at a time when annual ridership had fallen overall for the Port Authority of Allegheny County, dropping from 66 million passenger trips in 2001 to 63.8 million in 2011 (NTD 2002 and 2012).

---

6 Information obtained through personal communication with C. Hiniker, December 17, 2013.
7 Ibid.


**Reno, Nevada**

The RTC RAPID is a 4.5-mile BRT-lite line operated by the Regional Transportation Commission of Washoe County (RTC). The service opened in Fall 2009, with additional phases completed in 2011 and 2013 that added more specialized BRT elements, and it currently features articulated buses, traffic signal priority, branded stations, floor-level boarding platforms, and off-board fare collection.

Projected capital costs for the system were $13.43 million, with actual capital costs at $15.35 million. Projected ridership for the line was 2,660,000 boardings along the served corridor in FY 2010 and 3,079,283 boardings in FY 2013. Actual ridership fell short of these projections, with only 1,665,702 boardings along the served corridor in FY 2010 and 1,822,018 in FY 2013. Following the Great Recession, there was a system-wide reduction in ridership and service cuts, resulting in falling ridership along the corridor served by RTC RAPID in its opening years. Ridership along the corridor failed to reach FY 2007 levels (prior to introduction of the RTC RAPID) until FY 2012. Since introduction of the RTC RAPID service, ridership has grown by almost 10% in the corridor, a modest increase but still falling below early projections.

**San Antonio, Texas**

VIA Primo is a 20-mile BRT-lite service operated by VIA Metropolitan Transit. Opened in December 2012, the service features branded stations and vehicles and a traffic signal priority system. Projected capital costs for the system were $40.1 million, with actual capital costs being $35 million. Projected ridership was 5,000 to 8,000 average daily riders following one year of service. Actual ridership fell within this range, with 5,800 average daily riders following one year of service.

**San Bernardino, California**

The sbX Green Line is a 15.7-mile BRT line operated by Omnitrans. Opened for service in April 2014, the line features dedicated bus lanes for portions of the route, sheltered stations with platform-level boarding and ticket vending machines, and branded service.

Projected capital costs for the line were $191.7 million (FTA 2010), which proved accurate judging from reports following completion (Starcic 2015). Projected ridership was 5,600 average daily boardings within opening year (FTA 2010). Actual ridership fell well short of this mark, with only 2,300 average daily boardings as of June 2015 (Wall 2015). Omnitrans officials have pointed to delays in the opening of a new transit center in Downtown San Bernardino, which originally was planned to open in tandem with the

---

8 Information obtained through personal communication E. Park, January 27, 2014.
9 Information obtained through personal communication R. Henson, February 26, 2014.
10 Ibid.
11 Information obtained through personal communication with J. Aguilera, February 7, 2014.
launch of the Green Line, as a factor in the line’s low first year ridership (Wall 2015). The transit center finally opened in September 2015.

**San Diego, California**

The Mid-City Rapid, also branded the Rapid 215, is a 10-mile BRT-lite line operated by the Metropolitan Transit System (MTS). Service began in October 2014 and features dedicated travel lanes along a small portion of the route, distinctive sheltered stations, a traffic signal priority system, and specially-branded articulated buses.

Projected capital costs for the service were $43.3 million (FTA 2008). Actual capital costs were $44 million (SANDAG 2014). Projected ridership was initially 15,000 average daily boardings upon opening (FTA 2008), although this figure was revised to 7,000 to 9,000 average daily boardings before the opening of the project (Keatts 2014). Actual ridership seems to have fallen short of this mark, with only 6,500 average daily boardings as of June 2015 (Schaver 2015). More recent ridership data have not been made publicly available as of this writing; without figures from after a full year of service, ridership data have been excluded from the final analysis.

**Snohomish County, Washington**

Swift is a 16.7-mile BRT-lite service operated by Community Transit. Opened in November 2009, the service features seven miles of transit-only lanes, traffic signal priority, articulated buses, and branded stations with pre-boarding fare collection.

Projected capital costs were $15–20 million (Community Transit 2005), with actual capital costs turning out to be $29 million, with 4 stations included in the original plan deferred until 2011 due to funding issues (Community Transit 2011). Projected ridership was 2,500 average weekday boardings after one year of service and 4,000 average weekday boardings after four years of service (Duke 2010). Actual ridership exceeded these projections, with 3,500 average weekday boardings after one year of service (Duke 2010), and 4,400 average weekday boardings after four years of service (Munguia 2013). Swift ridership grew despite a system-wide reduction in ridership (Munguia 2010) and service cuts in 2010 and 2012, which reduced service frequency and operating hours for the BRT service (Munguia 2012). Annual ridership for Community Transit dropped from 11.4 million in 2009 to 9.1 million in 2012 (NTD 2010, 2013).

The following table summarizes the above case studies, comparing the predicted and actual costs in constant US dollars per mile and the predicted and actual ridership.
### TABLE 1. Summary of Individual Case Studies

<table>
<thead>
<tr>
<th>City</th>
<th>Project name</th>
<th>System type</th>
<th>Year opened</th>
<th>Length (in miles)</th>
<th>Predicted cost per mile (in millions, adjusted to 2015 values)</th>
<th>Actual cost per mile (in millions, adjusted to 2015 values)</th>
<th>Actual total cost as % of Predicted total cost</th>
<th>Predicted ridership</th>
<th>Actual ridership</th>
<th>Actual ridership as % of Predicted ridership</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cleveland</td>
<td>HealthLine</td>
<td>BRT</td>
<td>2008</td>
<td>7.1</td>
<td>$42.20</td>
<td>$30.43</td>
<td>72%</td>
<td>21,100/day</td>
<td>14,300/day</td>
<td>69%</td>
</tr>
<tr>
<td>El Paso</td>
<td>BRT-lite</td>
<td>BRT-lite</td>
<td>2014</td>
<td>8.6</td>
<td>$3.17</td>
<td>$3.17</td>
<td>100%</td>
<td>11,900/day</td>
<td>2,000/day</td>
<td>17%</td>
</tr>
<tr>
<td>Escondido</td>
<td>Breeze Rapid</td>
<td>BRT-lite</td>
<td>2011</td>
<td>6</td>
<td>$0.49</td>
<td>$0.73</td>
<td>151%</td>
<td>Increase by 40% in 20 years</td>
<td>200+%</td>
<td>Increase by 222% in 4 years</td>
</tr>
<tr>
<td>Eugene</td>
<td>Emerald Express</td>
<td>BRT</td>
<td>2007</td>
<td>4</td>
<td>$7.00</td>
<td>$7.00</td>
<td>100%</td>
<td>20 years</td>
<td>3,700/day</td>
<td>5,000/day</td>
</tr>
<tr>
<td>Fort Collins</td>
<td>Emerald Express Gateway Ext.</td>
<td>BRT</td>
<td>2011</td>
<td>7.8</td>
<td>$5.80</td>
<td>$5.56</td>
<td>96%</td>
<td>3,900/day</td>
<td>3,680/day</td>
<td>95%</td>
</tr>
<tr>
<td>Grand Rapids</td>
<td>Silver Line</td>
<td>BRT-lite</td>
<td>2014</td>
<td>9.8</td>
<td>$3.80</td>
<td>$4.10</td>
<td>108%</td>
<td>7,200/day</td>
<td>2,300/day</td>
<td>32%</td>
</tr>
<tr>
<td>Kansas City</td>
<td>MAX TreoSt Line</td>
<td>BRT</td>
<td>2014</td>
<td>11</td>
<td>$2.47</td>
<td>$2.46</td>
<td>100%</td>
<td>9,000/day</td>
<td>8,500/day</td>
<td>95%</td>
</tr>
<tr>
<td>Las Vegas</td>
<td>Sahara Express</td>
<td>BRT</td>
<td>2014</td>
<td>12</td>
<td>$3.77</td>
<td>$3.90</td>
<td>103%</td>
<td>13,900/day</td>
<td>10,000/day</td>
<td>72%</td>
</tr>
<tr>
<td>Los Angeles</td>
<td>Orange Line</td>
<td>BRT</td>
<td>2005</td>
<td>14</td>
<td>$29.37</td>
<td>$27.92</td>
<td>95%</td>
<td>5,000-7,500/day</td>
<td>21,828/day</td>
<td>200+%</td>
</tr>
<tr>
<td>Los Angeles</td>
<td>Orange Line Canoga Ext.</td>
<td>BRT</td>
<td>2012</td>
<td>4</td>
<td>$34.69</td>
<td>$39.57</td>
<td>114%</td>
<td>2,250/day</td>
<td>975/day</td>
<td>43%</td>
</tr>
<tr>
<td>Minneapolis</td>
<td>METRO Red Line</td>
<td>BRT</td>
<td>2013</td>
<td>11</td>
<td>$10.87</td>
<td>$10.31</td>
<td>95%</td>
<td>7,000/day</td>
<td>7,000/day</td>
<td>124%</td>
</tr>
<tr>
<td>Pittsburgh</td>
<td>West Busway</td>
<td>Busway</td>
<td>2000</td>
<td>5</td>
<td>$55.62*</td>
<td>$89.57</td>
<td>109%</td>
<td>3,800/day</td>
<td>2,000/day</td>
<td>53%</td>
</tr>
<tr>
<td>Pittsburgh</td>
<td>East Busway Ext.</td>
<td>Busway</td>
<td>2003</td>
<td>2.3</td>
<td>$35.01</td>
<td>$38.19</td>
<td>109%</td>
<td>Additional</td>
<td>2,000/day</td>
<td>53%</td>
</tr>
<tr>
<td>Reno</td>
<td>RTC RAPID</td>
<td>BRT</td>
<td>2009</td>
<td>4.5</td>
<td>$3.28</td>
<td>$3.75</td>
<td>114%</td>
<td>3.1 million/yr</td>
<td>1.8 million/yr</td>
<td>59%</td>
</tr>
<tr>
<td>San Antonio</td>
<td>VIA Primo</td>
<td>BRT</td>
<td>2012</td>
<td>20</td>
<td>$2.07</td>
<td>$1.80</td>
<td>87%</td>
<td>5,000-8,000/day</td>
<td>5,000/day</td>
<td>100%</td>
</tr>
<tr>
<td>San Bernardino</td>
<td>sbX Green Line</td>
<td>BRT</td>
<td>2014</td>
<td>15.7</td>
<td>$12.27</td>
<td>$12.27</td>
<td>100%</td>
<td>5,600/day</td>
<td>2,300/day</td>
<td>41%</td>
</tr>
<tr>
<td>San Diego</td>
<td>Mid-City Rapid</td>
<td>BRT</td>
<td>2014</td>
<td>10</td>
<td>$4.35</td>
<td>$4.42</td>
<td>102%</td>
<td>15,000/day</td>
<td>4,400/day</td>
<td>110%</td>
</tr>
<tr>
<td>Snohomish County</td>
<td>Swift</td>
<td>BRT</td>
<td>2009</td>
<td>16.7</td>
<td>$1.32</td>
<td>$1.91</td>
<td>145%</td>
<td>4,000/day</td>
<td>4,400/day</td>
<td>110%</td>
</tr>
</tbody>
</table>

*Original West Busway proposal was an 8.1-mile project.*
Results

The BRT projects studied generally had more accurate cost estimates than the urban rail projects studied by Flyvbjerg. Of the 19 projects for which adequate cost information could be obtained, the actual capital costs of 10 came within 5% of their projected cost, and 16 came within 15% of their projected cost. However, although there was a high number of accurate or nearly accurate cost forecasts, the data also show a propensity towards projects coming in over their estimated budget, with 7 exceeding the projected cost by at least 5%, compared to only 2 that underestimated it by at least 5% (Figure 1). Of the projects described, the two that were far over budget (by at least 15%) were BRT-lite systems that had such low capital costs that even a difference of a few million dollars had a large proportional effect.

When it comes to ridership estimates, however, the overall picture is much different. Of the 16 projects for which adequate ridership data could be obtained, only 2 came within 10% of their projected ridership. The ridership projections also showed a propensity towards being over the actual ridership figures, with 8 projects seeing ridership at least 10% below what was projected versus only 6 seeing ridership at least 10% above the estimates (Figure 2).
Discussion

Whereas there was a propensity towards projects coming in over budget among the systems studied, the level of accuracy shown suggests that BRT projects in the US do not suffer from the same level of cost estimate inaccuracies as the urban rail projects studied by Flyvbjerg or FTA. In general, the accuracy of the estimated costs for these systems was very good.

However, although there was a high level of accuracy in the cost estimates, it is worth noting that in a few cases these BRT projects came within budget only due to a scaling back of the project from what was initially proposed. This was particularly evident in the case of the Western Busway in Pittsburgh, where the scope of the project was significantly reduced, and to a lesser extent with the HealthLine in Cleveland, when the expense of certain design elements was scaled back.

The ridership estimates, on the other hand, not only showed a high level of inaccuracy but also a clear propensity towards predicting ridership higher than the actual results. This shows that current ridership forecasting methods still leave something to be desired and suggests that many US public transit agencies may be too optimistic as to the ridership outcomes of their BRT projects. In some cases, the underperformance of these projects has been attributed to outside factors; in San Bernardino, for instance, delays in the opening of a new transit center were blamed for their BRT system’s low initial ridership.

A common theme among several of these projects was the effect of the Great Recession on ridership: in Cleveland, Minneapolis, Pittsburgh, and Reno, overall transit ridership
dropped as a result of the recession and subsequent cuts to transit service, which likely was a factor in causing new BRT projects in these cities to fail to live up to ridership expectations. However, the recession cannot account for all of the underperforming systems included here; some of the newest systems studied, such as those in El Paso, Grand Rapids, and San Bernardino, performed below expectations despite the fact that their ridership estimates were generated well after the start of the Great Recession. Additionally, there is little to suggest that ridership forecasts have gotten more accurate over time (Figure 3).

It is worth noting that ridership alone is not the only measure of success that can be applied to a public transportation infrastructure project, and, with few exceptions, the figures shown here reflect only the initial years of operating service. But transportation professionals should strive for a high degree of accuracy when discussing the potential benefits of a public transportation project, especially if they seek public confidence to expand public transportation infrastructure and service in the US in the years to come. It is not the intent of this research to question the worthiness of any of these BRT projects, but to note that a very common failing is occurring in the process of justifying these projects.

When applying for federal funding from FTA, three approaches can be taken to provide a ridership forecast: 1) using a region-wide travel model, 2) using incremental data-driven methods, which rely on existing ridership data and make projections by estimating the effects of proposed or expected changes, or 3) using FTA’s Simplified Trips-on-Project Software (STOPS) created by FTA, a simplified version of the standard four-step travel model (FTA 2016). When creating ridership projections for future years,
all three of these approaches rely on data input by local or regional agencies, including expected population and employment patterns. This brings us to a key weakness with standard travel forecasting methods: local governments and agencies in the US tend to be optimistic about future growth in population, employment, and transit ridership in their communities.

In his work, Flyvbjerg advocated for the adoption of “reference class forecasting,” in which an outside view of a proposed project would be enforced by comparing it to the outcomes of a reference group of similar projects (Flyvbjerg et al. 2005). Although there are inherent difficulties in determining which projects serve as adequate reference points and compensating for the unique characteristics of any single project, such an approach would avoid many of the drawbacks of current standard approaches, in particular the tendency towards optimistic future growth forecasts. With a growing number of BRT projects in the US, there is a growing reference class of projects to use.

Flyvbjerg also pointed out that current funding mechanisms, in which transportation projects across the country compete against each other for crucial federal funding, create an incentive for local planners to oversell the benefits of their projects (Flyvbjerg et al. 2005). This could be addressed either by adopting different funding mechanisms, in which transportation projects do not have to compete directly against each other at the federal level, or by applying more rigorous scrutiny to ridership forecasts, perhaps by comparing them to similar projects as suggested above.

Conclusions

In recent years, BRT has become an increasingly-common option for local and regional agencies when investing in public transit infrastructure. The results of this analysis show that BRT projects in the US do very well when holding to their cost projections, but fall short where it comes to ridership projections. Although they skewed slightly towards being completed over budget, the cost estimates of the BRT projects studied tended to be far more accurate than the urban rail projects Flyvbjerg studied. But the widespread inaccuracy of the ridership estimates among the projects studied demonstrates that a more critical eye should be directed towards ridership projections. Although there is much inherent difficulty in accurately predicting future transit ridership, the tendency towards overestimating ridership shown here suggests a bias similar to that demonstrated by Flyvbjerg’s research.

BRT offers an excellent opportunity for many communities to invest in high-quality public transportation. However, the results of this research suggest that there may be a tendency to oversell the benefits of these projects. With many new BRT projects under construction and opening in the years to come, it is important that transportation professionals apply more rigorous methodology to the ridership projections for these projects.
References


FTA. 2009. “Mason Corridor BRT, Fort Collins, Colorado.”


FTA. 2010. “E Street Corridor sBX BRT, San Bernardino, California.”

FTA. 2010. “Mesa Corridor BRT, El Paso, Texas.”


Measuring the Accuracy of Bus Rapid Transit Forecasts


About the Author

JOHN PERRY (abqstreetcar@gmail.com) received an MA in City & Regional Planning with a specialization in Transportation Planning from the University of North Carolina at Chapel Hill in 2014 and a BA in Environmental Planning and Design with an emphasis in Community Planning from the University of New Mexico in 2012. While
in Albuquerque, he discovered a passion for urban and transportation planning, with a particular interest in public transit and active transportation systems. He has worked as an advocate for the improvement of sustainable transportation in the US. He currently lives in Ames, Iowa.
Effect of Price Reduction and Increased Service Frequency on Public Transport Travel

Inge Brechan
Institute of Transport Economics, Oslo

Abstract

A random effects meta-analysis of the results from 15 projects involving price reduction and 9 projects involving increased service frequency showed that both price reduction and increased service frequency generated public transport travels. On average, the increased service frequency projects generated more travels by public transport than the price reduction projects. In the increased service frequency projects the proportion of travels generated by the increased frequency was strongly influenced by the size of the frequency increase. In the price reduction projects, we did not find a significant effect of the size of the price reduction on the proportion of travels generated by the price reduction. Finding that people’s use of public transport was related to the extent of the service offered suggests they have a need for transport that can be fulfilled with public transport. Although people appreciate lower fares in general, finding that use of public transport was not significantly related to the size of a price change suggests the effect of price change is uncertain.

Keywords: Price, service frequency, meta-analysis, travel mode choice, elasticity

Introduction

With the aim of generating more travel by public transport, several avenues may be considered. Two obvious options are to offer more service and to charge less for the service currently offered. Which of the two options brings the most success in terms of journeys generated? This research question is of interest to academia, policymakers, the public transport industry, environmental organizations, and general society. Research on longitudinal economic data (1987–1996) from France and England differ in their conclusions. Price elasticity was found to be greater than service (i.e., vehicle kilometers) elasticity in France, whereas the opposite was found in England (Bresson et al. 2003, 2004; Dargay and Hanly 2002). According to Preston (2014), research reviews
suggest that service elasticities in general are larger than price elasticities, but also that elasticities vary a lot. This paper investigated the research question by performing a meta-analysis of the results from customer survey data from the Norwegian trial scheme for public transport.

**Use of Public Transport Related to Price and Service Frequency**

At the most general level, use of public transport can be predicted by what customers give (i.e., cost) relative to what they get (i.e., service). Research on more specific factors associated with use of public transport are often related to fares (e.g., type of fare, price of petrol, income) and quality of service (e.g., intervals, reliability, interchanges) (see Balcombe et al. 2004). Although service quality may also include non-essential attributes, such as cleanliness, research has indicated that the problem-solving capability of the travel mode (i.e., taking people from where they are to where they want to go at the time they need to be there) is essential (Brechan 2006). Thus, routes and schedules are the primary service factors.

Results from meta-analyses on the effects of price and service frequency on use of public transport vary. Holmgren (2007) found average short-term service (vehicle kilometers) elasticity (1.05) to be more extreme than price elasticity (Europe -0.75 and America/ Australia -0.59). Holmgren (2007) found the service elasticity to vary more than the price elasticity. Hensher (2008) found average service (headway) elasticity (-0.29) to be less extreme than price elasticity (-0.40). Hensher (2008) found the price elasticity to vary more than the service elasticity. According to Paulley et al. (2006), the average short-term service elasticity (0.4) and price elasticity (-0.4) are similar in absolute strength. Paulley et al. (2006) concludes that there is a wide range of fare elasticities and suggests that the impact of prices is higher in the long run (vs. short run), in rural (vs. urban) areas, for leisure and shopping (vs. work and education) purposes and in off-peak (vs. traffic peak) hours. Similarly, they also suggest the impact of service (intervals) is higher in the long run (vs. short run), in rural (vs. urban) areas, and in off-peak (vs. traffic peak) hours.

**Norwegian Trial Scheme for Public Transport**

The Norwegian government’s Ministry of Transport established the Norwegian trial scheme for public transport in 1991, with the aim of developing public transport solutions that are more need-oriented, resource-efficient, and environmentally-friendly. Approximately 500 projects were awarded 461 million Norwegian kroner (NOK), the equivalent of approximately 60 million euro (EUR) or $70 million US, during the period 1991–1995. In total 24% of the projects were classified as route trials, including trials with increased service frequency, and 10% were classified as fare trials (i.e., trials with price reduction). Other projects in the trial scheme concerned bus terminals, information and marketing, ticketing, organization and administration, quality of vehicles, fuels, and telecommunication. Among the route and fare trials, 101 projects
Effect of Price Reduction and Increased Service Frequency on Public Transport Travel

were evaluated partly by means of an on-board customer survey. The results from these customer surveys constitute the data available for the current meta-analysis.

According to the database of Statistics Norway (www.ssb.no/statistikkbanken), the inland (excluding journeys starting or ending abroad) motorized passenger transport volume in Norway was 56,132 million passenger kilometres in 1995 (see Figure 1), which equals 35 km per person per day. This was divided among bus transport (7%), other road transport (81%—private car 78%, motorcycle 1%, taxi 1%, rental car 1%), rail (5%), air (6%), and sea (1%). Based on the Norwegian National Travel Study (www.toi.no/rv) from 1992 and 1998, we estimated that people in Norway on average walked 1 km and bicycled 0.5 km per day in 1995. Thus, the non-motorized passenger travel volume was 1.5 km per person per day, compared to the motorized passenger travel volume of 35 km per person per day. There are large regional and seasonal differences. Norway is a sparsely-populated (14.29 pop. per sq. km in 1995) mountainous country (46% mountain, 43% forest, 6% lakes, 3% agriculture, 2% built) with a long coastline (25,148 km continental coastline) situated on the Arctic Circle, comparable in size to Poland, The Ivory Coast, Malaysia, or New Mexico. Most people in Norway live in cities on the coastline. Approximately 80% live in urban areas. Oslo, the capital, has a well-developed public transit system with a mix of buses, rail, and boats. All route and fare trials included in this meta-analysis took place in smaller cities (population < 150,000), where the public transit system consists of almost exclusively buses. Some of the areas have rail and/or boat service as well, but these travel modes make up only a very small part of the service. Rail’s share of passenger transport in Figure 1 stems mostly from regional and intercity train services not included in this meta-analysis, but one of the frequency trials involved a tram (trolley, streetcar) service.

FIGURE 1.
Annual inland motorized passenger transport volume in Norway

Note: Includes only journeys starting and ending in Norway.
Method

The 101 customer surveys were fairly identical and included questions about the quality of public transport, travel behavior, and demographics. The questions measuring the direct impact of the price reduction or frequency increase were: “Are you aware of the recent changes in the ticket prices (or frequency of services)?” and, if so, “By what mode of transport would you have conducted this specific journey if the recent changes had not taken place?” The possible answers offered for the last question included all public modes of transport, private motorized modes of transport, non-motorized modes of transport, and the option of not traveling at all. The outcome variable included in this analysis is the proportion of respondents reporting that they were aware of the changes in the prices (or services) and that they would not have used public transport for this specific journey if the changes had not taken place.

Of the 101 projects, 12 were omitted due to very small survey sample (< 30 respondents) and another 12 were omitted because the projects were either quite unique (e.g., concerned boats rather than buses) or could not be exclusively categorized (e.g., included changes in prices as well as services). Among the remaining 77 projects, 25 were price reduction trials, 12 were increased frequency trials, and 40 were other route trials (mostly new routes). Finally, because some projects were merely continuations of earlier projects in the trial scheme and implied no further changes to the price or service, the final sample of price and frequency trials to be included in the meta-analysis consists of 15 price reduction trials and 9 increased frequency trials. All 24 trials took place in small cities (< 150,000 inhabitants), and all routes surveyed were general local bus routes (except one frequency trial that involved a tram service).

First, we calculated the average effect of each of the two groups of projects and compared the two to see if one type of project had a larger effect than the other. Then, we investigated the relationship between the outcome (i.e., proportion of journeys generated by the project) and the size of the change in price or frequency. The independent variables, price decrease and frequency increase, were measured as the price reduction in percent of the original fare and the increase in departures in percent of the original number of departures on a route.

Results

Information on the size of the price reduction or frequency increase, sample size of the customer surveys, and effect size discovered in the customer surveys are presented in Table 1 (price reduction projects) and Table 2 (service frequency increase projects). Effect sizes are presented as the number (labeled “raw effect”) of respondents reporting that they would not have used public transport for the current journey if the project (reduced price or increased frequency) had not taken place and the proportion (m) this number of respondents represents relative to the total number of people interviewed.
Effect of Price Reduction and Increased Service Frequency on Public Transport Travel

TABLE 1.
Description of Projects Involving Price Reduction

<table>
<thead>
<tr>
<th>Project ID</th>
<th>Region</th>
<th>Price Reduction</th>
<th>Sample Size (n)</th>
<th>Raw Effect*</th>
<th>Proportion**</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-020</td>
<td>Østfold</td>
<td>30.00%</td>
<td>186</td>
<td>20</td>
<td>0.11</td>
</tr>
<tr>
<td>2-005</td>
<td>Akershus</td>
<td>33.33%</td>
<td>112</td>
<td>41</td>
<td>0.37</td>
</tr>
<tr>
<td>4-001A</td>
<td>Hedmark</td>
<td>25.00%</td>
<td>114</td>
<td>37</td>
<td>0.32</td>
</tr>
<tr>
<td>4-001B</td>
<td>Hedmark</td>
<td>25.00%</td>
<td>106</td>
<td>26</td>
<td>0.25</td>
</tr>
<tr>
<td>7-001</td>
<td>Vestfold</td>
<td>37.50%</td>
<td>270</td>
<td>29</td>
<td>0.11</td>
</tr>
<tr>
<td>9-001</td>
<td>Aust-Agder</td>
<td>18.60%</td>
<td>153</td>
<td>70</td>
<td>0.46</td>
</tr>
<tr>
<td>10-001</td>
<td>Vest-Agder</td>
<td>64.44%</td>
<td>514</td>
<td>238</td>
<td>0.46</td>
</tr>
<tr>
<td>10-002</td>
<td>Vest-Agder</td>
<td>22.22%</td>
<td>421</td>
<td>162</td>
<td>0.38</td>
</tr>
<tr>
<td>11-007</td>
<td>Rogaland</td>
<td>50.00%</td>
<td>404</td>
<td>89</td>
<td>0.22</td>
</tr>
<tr>
<td>11-009</td>
<td>Rogaland</td>
<td>36.36%</td>
<td>89</td>
<td>22</td>
<td>0.25</td>
</tr>
<tr>
<td>15-001</td>
<td>Møre og Romsdal</td>
<td>35.00%</td>
<td>805</td>
<td>230</td>
<td>0.29</td>
</tr>
<tr>
<td>15-002</td>
<td>Møre og Romsdal</td>
<td>34.21%</td>
<td>1125</td>
<td>324</td>
<td>0.29</td>
</tr>
<tr>
<td>15-015</td>
<td>Møre og Romsdal</td>
<td>35.00%</td>
<td>1955</td>
<td>707</td>
<td>0.36</td>
</tr>
<tr>
<td>18-001A</td>
<td>Nordland</td>
<td>34.48%</td>
<td>90</td>
<td>26</td>
<td>0.29</td>
</tr>
<tr>
<td>18-001B</td>
<td>Nordland</td>
<td>48.84%</td>
<td>90</td>
<td>46</td>
<td>0.51</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>6434</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Number of respondents reporting that journey was generated by price reduction.
**Raw effect divided by sample size.

TABLE 2.
Description of Projects Involving Increase in Service Frequency

<table>
<thead>
<tr>
<th>Project ID</th>
<th>Region</th>
<th>Frequency Increase</th>
<th>Sample Size (n)</th>
<th>Raw Effect*</th>
<th>Proportion**</th>
</tr>
</thead>
<tbody>
<tr>
<td>7-005</td>
<td>Vestfold</td>
<td>67.83%</td>
<td>308</td>
<td>250</td>
<td>0.81</td>
</tr>
<tr>
<td>7-006</td>
<td>Vestfold</td>
<td>12.50%</td>
<td>235</td>
<td>105</td>
<td>0.45</td>
</tr>
<tr>
<td>10-018</td>
<td>Vest-Agder</td>
<td>42.58%</td>
<td>207</td>
<td>63</td>
<td>0.30</td>
</tr>
<tr>
<td>10-033</td>
<td>Vest-Agder</td>
<td>40.54%</td>
<td>166</td>
<td>77</td>
<td>0.46</td>
</tr>
<tr>
<td>10-036</td>
<td>Vest-Agder</td>
<td>44.44%</td>
<td>126</td>
<td>52</td>
<td>0.41</td>
</tr>
<tr>
<td>11-004</td>
<td>Rogaland</td>
<td>23.56%</td>
<td>143</td>
<td>39</td>
<td>0.27</td>
</tr>
<tr>
<td>11-027</td>
<td>Rogaland</td>
<td>45.56%</td>
<td>112</td>
<td>34</td>
<td>0.30</td>
</tr>
<tr>
<td>16-004</td>
<td>Sør-Trøndelag</td>
<td>26.98%</td>
<td>3494</td>
<td>329</td>
<td>0.09</td>
</tr>
<tr>
<td>16-009**</td>
<td>Sør-Trøndelag</td>
<td>9.32%</td>
<td>154</td>
<td>45</td>
<td>0.29</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>4945</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Number of respondents reporting that journey was generated by increase in service frequency.
**Raw effect divided by sample size
***This trial involved tram (trolley, streetcar), whereas other trials involved buses.

We did not expect the effect sizes in the projects to be similar, because the projects differed with regard to the size of the price reduction or increase in service frequency implemented. As such, the effect sizes are not representations of a general fixed effect.
and, therefore, this is a meta-analysis of random effects (Shadish and Haddock 1994). In a fixed effects model, the effect sizes in the set of studies are all estimates of a common (i.e., fixed) effect in the population. Any difference between a study effect and the common effect would be due to sampling error, as a given study uses a sample of subjects from the population. The term “random effects” reflects the assumption that the effect sizes in a random effects meta-analysis are sampled from a larger population of effect sizes (Raudenbush 1994). We nevertheless started our analyses by investigating the homogeneity of the effect sizes.

Because the variance ($\sigma^2$) of a proportion ($m$) is determined in part by the magnitude of the proportion (see Equation 1 and Figure 2), proportions are not equally detectable, and the difference between proportions is not an appropriate measure of effect size (Cohen 1988). For example, at $n = 250$ and $\alpha = 0.05$, the power to detect the difference between 0.1 and 0.2 is 0.89, whereas the power to detect the difference between 0.5 and 0.6 is 0.62 (Lenth 2004). Thus, we could not use proportions (or elasticities based on proportions) in our meta-analyses. Instead, we transformed the proportions to radians $\phi$ (see Equation 2, where $n =$ sample size, and Figure 3) and calculated the effect size $h$ representing the difference between a proportion and zero (see Equation 3) (Cohen 1988). The variance of the effect size $h$ is not influenced by the magnitude of the effect size (see Equation 4 and Figure 4).

$$\sigma^2 = \frac{m(1-m)}{n}$$ (1)

**FIGURE 2.**
Variance of proportions

![Variance of proportions graph](image)
\[ \phi_m = 2 \times \arcsin \sqrt{m} \]

\[ \phi_0 = 2 \times \arcsin \left( \frac{1}{4n} \right) \]  

(2)

**FIGURE 3.**
Proportions and corresponding radians

\[ h = \phi_m - \phi_0 = \left( 2 \times \arcsin \sqrt{m} \right) - \left( 2 \times \arcsin \left( \frac{1}{4n} \right) \right) \]  

(3)

\[ v_j = \frac{1}{n_j} \]  

(4)

**FIGURE 4.**
Variance of Cohen’s h
When calculating the homogeneity test statistic \( Q \) (see Equation 5), the projects were weighted based on the within-study variance \( \nu_i \) (see Equation 4) of the effect size \( h_i \) (Shadish and Haddock 1994). Formulas for weights \( w_i \) are given in Equation 6 (where \( k = \) number of studies).

\[
Q = \sum_{i=1}^{k} w_i h_i^2 - \left( \frac{\sum_{i=1}^{k} w_i h_i}{\sum_{i=1}^{k} w_i} \right)^2
\]  

\[
w_i = \frac{1}{\nu_i}
\]

We used Equation 5 to calculate both overall homogeneity \( Q_T \) and the within-group \( Q_{Wj} \) homogeneity for both types of projects. The input and computations regarding the price reduction projects are given in Table 1 (input) and Table 3 (computations), whereas the input and computations regarding the increased service frequency projects are given in Table 2 (input) and Table 4 (computations). To calculate the overall homogeneity \( Q_T \), we simply combined the numbers from Table 1 and Table 2 (input), and Table 3 and Table 4 (computations).

**TABLE 3.**

<table>
<thead>
<tr>
<th>Project ID</th>
<th>Region</th>
<th>( h )</th>
<th>( \nu )</th>
<th>( w )</th>
<th>( w_i h_i )</th>
<th>( w_i h_i^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-020</td>
<td>Østfold</td>
<td>0.67</td>
<td>0.0054</td>
<td>186</td>
<td>123.78</td>
<td>82.38</td>
</tr>
<tr>
<td>2-005</td>
<td>Akershus</td>
<td>1.30</td>
<td>0.0089</td>
<td>112</td>
<td>145.06</td>
<td>187.87</td>
</tr>
<tr>
<td>4-001A</td>
<td>Hedmark</td>
<td>1.21</td>
<td>0.0088</td>
<td>114</td>
<td>137.70</td>
<td>166.33</td>
</tr>
<tr>
<td>4-001B</td>
<td>Hedmark</td>
<td>1.03</td>
<td>0.0094</td>
<td>106</td>
<td>109.34</td>
<td>112.79</td>
</tr>
<tr>
<td>7-001</td>
<td>Vestfold</td>
<td>0.67</td>
<td>0.0037</td>
<td>270</td>
<td>179.81</td>
<td>119.74</td>
</tr>
<tr>
<td>9-001</td>
<td>Aust-Agder</td>
<td>1.48</td>
<td>0.0065</td>
<td>153</td>
<td>226.82</td>
<td>336.25</td>
</tr>
<tr>
<td>10-001</td>
<td>Vest-Agder</td>
<td>1.50</td>
<td>0.0019</td>
<td>514</td>
<td>768.85</td>
<td>1150.07</td>
</tr>
<tr>
<td>10-002</td>
<td>Vest-Agder</td>
<td>1.34</td>
<td>0.0024</td>
<td>421</td>
<td>562.93</td>
<td>752.70</td>
</tr>
<tr>
<td>11-007</td>
<td>Rogaland</td>
<td>0.98</td>
<td>0.0025</td>
<td>404</td>
<td>394.26</td>
<td>384.75</td>
</tr>
<tr>
<td>11-009</td>
<td>Rogaland</td>
<td>1.04</td>
<td>0.0112</td>
<td>89</td>
<td>92.12</td>
<td>95.35</td>
</tr>
<tr>
<td>15-001</td>
<td>Møre og Romsdal</td>
<td>1.13</td>
<td>0.0012</td>
<td>805</td>
<td>907.45</td>
<td>1022.93</td>
</tr>
<tr>
<td>15-002</td>
<td>Møre og Romsdal</td>
<td>1.13</td>
<td>0.0009</td>
<td>1125</td>
<td>1274.06</td>
<td>1442.86</td>
</tr>
<tr>
<td>15-015</td>
<td>Møre og Romsdal</td>
<td>1.29</td>
<td>0.0005</td>
<td>1955</td>
<td>2522.25</td>
<td>3254.10</td>
</tr>
<tr>
<td>18-001A</td>
<td>Nordland</td>
<td>1.13</td>
<td>0.0111</td>
<td>90</td>
<td>101.64</td>
<td>114.79</td>
</tr>
<tr>
<td>18-001B</td>
<td>Nordland</td>
<td>1.59</td>
<td>0.0111</td>
<td>90</td>
<td>142.87</td>
<td>226.80</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td>6434</td>
<td>7688.94</td>
<td>9449.72</td>
</tr>
</tbody>
</table>
TABLE 4.
Computational Details for Projects Involving Increase in Service Frequency

<table>
<thead>
<tr>
<th>Project ID</th>
<th>Region</th>
<th>$h_i$</th>
<th>$v_i$</th>
<th>$w_i$</th>
<th>$w_i^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>7-005</td>
<td>Vestfold</td>
<td>2.24</td>
<td>0.0032</td>
<td>308</td>
<td>926.88</td>
</tr>
<tr>
<td>7-006</td>
<td>Vestfold</td>
<td>1.46</td>
<td>0.0043</td>
<td>235</td>
<td>546.25</td>
</tr>
<tr>
<td>10-018</td>
<td>Vest-Agder</td>
<td>1.17</td>
<td>0.0048</td>
<td>207</td>
<td>241.43</td>
</tr>
<tr>
<td>10-033</td>
<td>Vest-Agder</td>
<td>1.50</td>
<td>0.0060</td>
<td>166</td>
<td>265.64</td>
</tr>
<tr>
<td>10-036</td>
<td>Vest-Agder</td>
<td>1.39</td>
<td>0.0079</td>
<td>126</td>
<td>164.66</td>
</tr>
<tr>
<td>11-004</td>
<td>Rogaland</td>
<td>1.10</td>
<td>0.0070</td>
<td>143</td>
<td>156.65</td>
</tr>
<tr>
<td>11-027</td>
<td>Rogaland</td>
<td>1.16</td>
<td>0.0089</td>
<td>112</td>
<td>130.21</td>
</tr>
<tr>
<td>16-004</td>
<td>Sør-Trøndelag</td>
<td>0.62</td>
<td>0.0003</td>
<td>3494</td>
<td>2178.98</td>
</tr>
<tr>
<td>16-009</td>
<td>Sør-Trøndelag</td>
<td>1.14</td>
<td>0.0065</td>
<td>154</td>
<td>175.40</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td>4945</td>
<td>4340.42</td>
</tr>
</tbody>
</table>

The overall within-group homogeneity ($Q_w$) is the sum of the individual within-group homogeneity statistics (see Equation 7, where $l = \text{number of groups}$). The between-group homogeneity ($Q_B$) is the difference between the overall homogeneity statistic ($Q_T$) and the overall within-group homogeneity statistic ($Q_w$) (see Equation 8).

$$Q_w = \sum_{j=1}^{l} Q_{w_j}$$  \hspace{1cm} (7)

$$Q_B = Q_T - Q_w$$  \hspace{1cm} (8)

Calculations of homogeneity test statistics:

$$Q_{W_{\text{price}}} = 9449.72 - \frac{7688.94^2}{6434} = 261.07$$

$$Q_{W_{\text{frequency}}} = 4829.22 - \frac{4340.42^2}{4945} = 1019.47$$

$$Q_T = (9449.72 + 4829.22) - \frac{(7688.94 + 4340.42)^2}{(6434 + 4945)} = 1562.06$$

$$Q_w = 261.07 + 1019.47 = 1280.54$$

$$Q_B = 1562.06 - 1280.54 = 281.52$$

If the homogeneity statistic ($Q$) is larger than the upper-tail critical value of chi-square at $k - 1$ degrees of freedom, the observed variance in study effect sizes is significantly
greater than what can be expected by chance. The results of the homogeneity test are presented in Table 5. The test revealed that overall the effect sizes were not homogenous. The effect sizes in the price reduction projects were not homogenous, nor were the effect sizes in the increased frequency projects. The test also indicated that there was a significant difference between the two groups of projects, but finding that the effect sizes within the groups are not homogeneous calls for a random effects analysis.

### TABLE 5. Homogeneity

<table>
<thead>
<tr>
<th></th>
<th>Q</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price reduction projects</td>
<td>261.07</td>
<td>14</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Increased frequency projects</td>
<td>1019.47</td>
<td>8</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Overall within groups</td>
<td>1280.54</td>
<td>22</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Between groups</td>
<td>281.52</td>
<td>1</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Overall</td>
<td>1562.06</td>
<td>23</td>
<td>&lt;.01</td>
</tr>
</tbody>
</table>

Calculating the homogeneity test statistic had a purpose beyond investigating the homogeneity of the effects of the price reduction and frequency increase. In a random effects model the total variance ($\sigma^2_{\text{total}}$) of the effect of an individual study reflects both the within-study variance ($\sigma^2_{\text{within}}$) and the between-studies variance ($\sigma^2_{\text{between}}$). The relationship between the total variance ($\sigma^2_{\text{total}}$), the within-study variance ($\sigma^2_{\text{within}}$), and the between-studies variance ($\sigma^2_{\text{between}}$) is displayed in Equation 9. The within-group homogeneity test statistic ($Q_{\text{within}}$) was used to estimate the between-studies variance ($\sigma^2_{\text{between}}$) for each of the two groups of projects (see Equation 10). When calculating the average effect size ($h_i$) for each of the two groups of projects we weighed each individual effect size by its total variance rather than the within-study variance used in the homogeneity test. See Equation 11 regarding the average effect size ($h_i$) and Equation 12 regarding weights ($w_i^*$).

\[
v_i^* = \sigma^2_{\text{between}} + \sigma^2_{\text{within}}
\]

\[
\hat{\sigma}^2_{\theta} = \left[ Q - (k - 1) \right] / c
\]

\[
c = \sum_{i=1}^{k} w_i - \left[ \frac{\sum_{i=1}^{k} w_i^2}{\sum_{i=1}^{k} w_i^*} \right]
\]

\[
h_i = \frac{\sum_{i=1}^{k} w_i^* h_i}{\sum_{i=1}^{k} w_i^*}
\]

\[
w_i^* = \frac{1}{v_i}
\]
Effect of Price Reduction and Increased Service Frequency on Public Transport Travel

By completing the equations, we found the between-studies variance (see Table 6) and average effect size (see Table 7) for both types of projects. Computational details are presented in Table 8 (price reduction projects) and Table 9 (increased service frequency projects). Equation 13 gives the formula for calculating the total variance ($\sigma^2$) of the average effect size ($\hat{h}$), which was needed to calculate the confidence interval of the average effect size (see Equation 14) and for significance testing (see Equation 15 regarding a z-test of the average effect size). Equation 16 gives the formula for transforming the effect size ($\hat{h}$) back to a proportion ($m$).

### Table 6. Between-studies Variance

<table>
<thead>
<tr>
<th></th>
<th>$c$</th>
<th>$\sigma^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price reduction projects</td>
<td>5418.75</td>
<td>0.0456</td>
</tr>
<tr>
<td>Increased frequency projects</td>
<td>2416.97</td>
<td>0.4185</td>
</tr>
</tbody>
</table>

$$\sigma^2 = \frac{1}{k} \sum_{i=1}^{k} \left(1/\sigma^2_i\right)$$  \hspace{1cm} (13)

Confidence interval = $\hat{h} \pm C_{\alpha} \sqrt{\sigma^2}$  \hspace{1cm} (14)

$\alpha = .05$ gives $C_{.05} = 1.96$

$$z = \frac{|\hat{h}|}{\sqrt{\sigma^2}}$$  \hspace{1cm} (15)

$$m = \left(\sin\left(\frac{\phi_m}{2}\right)\right)^2$$  \hspace{1cm} (16)

$$\phi_m = \hat{h} + \phi_0 = \hat{h} + \left(2 \times \arcsin\left(\frac{1}{4n}\right)\right)$$

### Table 7. Average Effects

<table>
<thead>
<tr>
<th></th>
<th>$m$</th>
<th>$\hat{h}$</th>
<th>$\sigma^2$</th>
<th>$z$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(95% CI)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price reduction projects</td>
<td>0.30</td>
<td>1.16</td>
<td>0.0034</td>
<td>19.94</td>
<td>(&lt;.01)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.05-1.28)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Increased frequency projects</td>
<td>0.37</td>
<td>1.31</td>
<td>0.0471</td>
<td>6.03</td>
<td>(&lt;.01)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.88-1.73)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* CI = Confidence interval
The difference between the average effect for the price reduction projects ($m_i = 0.30$) and the average effect for the increased frequency projects ($m_i = 0.37$) constituted a group difference of $h = 0.15$ that was highly significant ($z = 7.72, p < .01$). Thus, we can conclude that the increased frequency projects on average generated a larger proportion of journeys than the price reduction projects (i.e., 37% vs. 30%, see Figure 5). See Equation 17 for calculating the difference ($h$) between two proportions and Equation 18 for significance testing ($z$-test of the difference between two proportions). Note that the difference in effect size between the two groups of projects is not adjusted for difference in the size of price reduction or frequency increase. The impact of the size of the price reduction or frequency increase on journeys generated is evaluated later. There is, however, no indication that the difference in average effect

<table>
<thead>
<tr>
<th>Project ID</th>
<th>Region</th>
<th>$w^2_i$</th>
<th>$v_i^*$</th>
<th>$w_i^*$</th>
<th>$w_i^*/h_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-020</td>
<td>Østfold</td>
<td>34596</td>
<td>0.0510</td>
<td>19.62</td>
<td>13.06</td>
</tr>
<tr>
<td>2-005</td>
<td>Akershus</td>
<td>12544</td>
<td>0.0545</td>
<td>18.34</td>
<td>23.75</td>
</tr>
<tr>
<td>4-001A</td>
<td>Hedmark</td>
<td>12996</td>
<td>0.0544</td>
<td>18.39</td>
<td>22.22</td>
</tr>
<tr>
<td>4-001B</td>
<td>Hedmark</td>
<td>11236</td>
<td>0.0550</td>
<td>18.17</td>
<td>18.75</td>
</tr>
<tr>
<td>7-001</td>
<td>Vestfold</td>
<td>72900</td>
<td>0.0493</td>
<td>20.28</td>
<td>13.51</td>
</tr>
<tr>
<td>9-001</td>
<td>Aust-Agder</td>
<td>23409</td>
<td>0.0521</td>
<td>19.18</td>
<td>28.44</td>
</tr>
<tr>
<td>10-001</td>
<td>Vest-Agder</td>
<td>264196</td>
<td>0.0475</td>
<td>21.03</td>
<td>31.46</td>
</tr>
<tr>
<td>10-002</td>
<td>Vest-Agder</td>
<td>177241</td>
<td>0.0480</td>
<td>20.85</td>
<td>27.87</td>
</tr>
<tr>
<td>11-007</td>
<td>Rogaland</td>
<td>163216</td>
<td>0.0481</td>
<td>20.80</td>
<td>20.30</td>
</tr>
<tr>
<td>11-009</td>
<td>Rogaland</td>
<td>7921</td>
<td>0.0568</td>
<td>17.60</td>
<td>18.21</td>
</tr>
<tr>
<td>15-001</td>
<td>Møre og Romsdal</td>
<td>648025</td>
<td>0.0468</td>
<td>21.35</td>
<td>24.07</td>
</tr>
<tr>
<td>15-002</td>
<td>Møre og Romsdal</td>
<td>1265625</td>
<td>0.0465</td>
<td>21.51</td>
<td>24.36</td>
</tr>
<tr>
<td>15-015</td>
<td>Møre og Romsdal</td>
<td>3822025</td>
<td>0.0461</td>
<td>21.69</td>
<td>27.98</td>
</tr>
<tr>
<td>18-001A</td>
<td>Nordland</td>
<td>8100</td>
<td>0.0567</td>
<td>17.63</td>
<td>19.92</td>
</tr>
<tr>
<td>18-001B</td>
<td>Nordland</td>
<td>8100</td>
<td>0.0567</td>
<td>17.63</td>
<td>27.99</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>6532130</td>
<td></td>
<td>294.09</td>
<td>341.89</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Project ID</th>
<th>Region</th>
<th>$w^2_i$</th>
<th>$v_i^*$</th>
<th>$w_i^*$</th>
<th>$w_i^*/h_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>7-005</td>
<td>Vestfold</td>
<td>94864</td>
<td>0.4217</td>
<td>2.37</td>
<td>5.32</td>
</tr>
<tr>
<td>7-006</td>
<td>Vestfold</td>
<td>55225</td>
<td>0.4227</td>
<td>2.37</td>
<td>3.46</td>
</tr>
<tr>
<td>10-018</td>
<td>Vest-Agder</td>
<td>42849</td>
<td>0.4233</td>
<td>2.36</td>
<td>2.76</td>
</tr>
<tr>
<td>10-033</td>
<td>Vest-Agder</td>
<td>27556</td>
<td>0.4245</td>
<td>2.36</td>
<td>3.52</td>
</tr>
<tr>
<td>10-036</td>
<td>Vest-Agder</td>
<td>15876</td>
<td>0.4264</td>
<td>2.35</td>
<td>3.26</td>
</tr>
<tr>
<td>11-004</td>
<td>Rogaland</td>
<td>20449</td>
<td>0.4255</td>
<td>2.35</td>
<td>2.57</td>
</tr>
<tr>
<td>11-027</td>
<td>Rogaland</td>
<td>12544</td>
<td>0.4274</td>
<td>2.34</td>
<td>2.72</td>
</tr>
<tr>
<td>16-004</td>
<td>Sør-Trøndelag</td>
<td>12208036</td>
<td>0.4188</td>
<td>2.39</td>
<td>1.49</td>
</tr>
<tr>
<td>16-009</td>
<td>Sør-Trøndelag</td>
<td>23716</td>
<td>0.4250</td>
<td>2.35</td>
<td>2.68</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>12501115</td>
<td></td>
<td>21.23</td>
<td>27.78</td>
</tr>
</tbody>
</table>
size between the two types of projects is due to difference in the size of the change in price versus size of the change in frequency (e.g., comparing a large change in frequency with a small change in price). For both groups of projects the average price reduction or frequency increase was 35%. The range of the change was slightly larger for the frequency increase projects (min = 9%, max = 68%, see Table 2) compared to the price reduction projects (min = 19%, max = 64%, see Table 1).

Because the homogeneity test revealed that the effect sizes of neither the increased frequency projects nor the price reduction projects were homogenous, further investigation of the variance of the effect sizes was called for (Shadish and Haddock 1994). Raudenbush (1994) recommends an iterative full information maximum likelihood approach for estimating both fixed effects and the between-studies variance, but according to Raudenbush (1994) the maximum likelihood approach may be less suitable for small samples. Because our two samples consisted of only 15 price reduction projects and 9 increased frequency projects, the alternative method of moments described by Raudenbush (1994) may be more suitable. We investigated the effect of the size of the change in price or frequency using both methods. Since the method of moments approach gave almost identical results as the maximum likelihood approach, and both pointed towards the same conclusion, we report only the results from the maximum likelihood approach here. We conducted separate regression analyses on the price reduction projects and the increased frequency projects with Cohen’s $h$ (see Table 3 and Table 4) as the dependent variable and price reduction (see Table 1) or frequency increase (see Table 2) as the independent variable.
We first conducted an ordinary least squares regression with the objective of finding an initial estimate of the between-studies variance ($\hat{\sigma}_\theta^2$). We calculated the initial estimate of the between-studies variance ($\hat{\sigma}_\theta^2$) by inserting the residual mean square (mean square error or $MSE$) from the regression into the formula presented in Equation 19.

Then, we conducted a weighted least squares regression using weights ($w_i^*$) calculated from the formula in Equation 12 based on the total variance ($v_i^*$) calculated from the formula in Equation 9. We used the residuals ($res_i$ in Equation 20) of the first weighted least squares regression to re-estimate the between-studies variance ($\hat{\sigma}_\theta^2$) using the formula presented in Equation 20. The re-estimate of the between-studies variance ($\hat{\sigma}_\theta^2$) was then used to calculate weights (see Equation 12) for another weighted least squares regression. The process was repeated (iterated) until the estimates were essentially unchanged (i.e., converged). In our analyses, the estimates converged after the second iteration for both groups of projects, showing a strong ($st. beta = .60$) effect of frequency increase and a small effect ($st. beta = .15$) effect of price reduction. The results are presented in Table 10.

\[
\hat{\sigma}_\theta^2 = MSE - \bar{v}
\]  

(19)

\[
\bar{v} = \sum_{i=1}^{k} v_i / k
\]

\[
\hat{\sigma}_\theta^2 = \frac{\sum w_i^* (res_i^2 - v_i)}{\sum w_i^*}
\]  

(20)

**TABLE 10. Maximum Likelihood Estimates**

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\sigma}_\theta^2$</th>
<th>MSE</th>
<th>Beta</th>
<th>Std. error *</th>
<th>Std. beta</th>
<th>$T^*$</th>
<th>$P^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price reduction</td>
<td>0.0614</td>
<td>1.1516</td>
<td>0.3508</td>
<td>0.6226</td>
<td>0.62</td>
<td>0.56</td>
<td>0.58</td>
</tr>
<tr>
<td>Increased frequency</td>
<td>0.1072</td>
<td>1.2598</td>
<td>1.4382</td>
<td>0.7192</td>
<td>0.63</td>
<td>2.00</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Note: The results (except between-studies variance $\hat{\sigma}_\theta^2$) are from a weighted least squares regression. The dependent variable is the effect size $h$ (see tables 3 and 4) and the independent variable is price reduction (see Table 1) or frequency increase (see Table 2).

* Uncorrected standard error yields incorrect significance statistics. The correction makes use of the residual mean square (MSE). See Table 11 for corrected statistics.

According to Hedges (1994) and Raudenbush (1994) the standard error used to calculate the significance of the weighted least square effect size must be corrected in meta-analyses. The formula for the corrected standard error is presented in Equation 21 and the formula for the significance test (z-test) of the effect size (i.e., unstandardized regression coefficient) is presented in Equation 22. The corrected standard errors, z-values, and significance statistics are presented in Table 11.
Effect of Price Reduction and Increased Service Frequency on Public Transport Travel

\[ S_{\text{cor}} = S / \sqrt{MSE} \]  

(21)

Where \( S \) and \( MSE \) are the standard error and residual mean square from the regression analysis.

\[ Z = |\beta| / S_{\text{cor}} \]  

(22)

Where \( \beta \) is the unstandardized regression coefficient from the regression analysis.

**TABLE 11. Regression Coefficients and Corrected Statistics**

<table>
<thead>
<tr>
<th></th>
<th>Beta</th>
<th>( S_{\text{cor}} )</th>
<th>Std. beta</th>
<th>( z )</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price reduction</td>
<td>0.3508</td>
<td>0.5802</td>
<td>.15</td>
<td>0.60</td>
<td>.55</td>
</tr>
<tr>
<td>Increased frequency</td>
<td>1.4382</td>
<td>0.6407</td>
<td>.60</td>
<td>2.24</td>
<td>.03</td>
</tr>
</tbody>
</table>

Note: The results are from a weighted least squares regression. The dependent variable is the effect size \( h \) (see Table 3 and Table 4) and the independent variable is price reduction (see Table 1) or frequency increase (see Table 2). The corrected standard error relates only to the significance statistics (\( z \) and \( p \)). The effect size estimates are unchanged.

The strong effect of increased frequency was clearly significant \((p = .03)\). The small effect of price reduction was not statistically significant \((p = .55)\). The results of the regression analyses are illustrated in Figure 6.

**FIGURE 6. Effect of price reduction and frequency increase on generation of journeys**

Note: Simple regression maximum likelihood estimates

**Discussion**

Finding that increased frequency, on average, generates more journeys than reduced price is in line with the conclusions of some reviews and meta-analyses (e.g., Preston
2014; Holmgren 2007) and replicates the results from a longitudinal study in England (Dargay and Hanly 2002). It is also in line with Bresson et al. (2003), who suggest that structural differences in the French and English samples may account for the differences between England and France in their study. The data from France were gathered in urban areas only, whereas in England the data represent the entire country, including non-urban areas. The areas represented in our sample are mostly urban according to Norwegian standards, but more sparsely-inhabited when compared to French and English cities. Thus, our sample probably resembles the English sample more than the French sample with regard to the population density of the areas included. Also, the results from the French sample are less clear when looking at a longer period in time and including other indicators. The difference between price elasticity and service elasticity disappears when expanding the time period from 1987–1995 to 1975–1995 (Bresson et al. 2004). When using seat kilometers rather than vehicle kilometers, as a measure of service frequency the relationship shifts completely, so that service elasticity is larger than price elasticity (Bresson et al. 2004).

Finding that there is a strong relationship between the size of the frequency increase and the number of journeys generated is logical. If more departures mean more journeys, then even more departures should mean even more journeys, up to a point where lack of departures is no longer a barrier for choosing public transport. These results from Norway suggest people had a need for transport that could be satisfied with public transport. Although we do not believe the demand for transport is unlimited, the results support the “demand follows supply” hypothesis. If public transport service was increased, people used public transport more. If public transport service was increased more, people used it even more. Chen, Varley, and Chan (2011) came to a similar conclusion when investigating data on public transport between New York City and New Jersey from the period 1996 to 2009.

Not finding a significant relationship between the size of the price reduction and the number of journeys generated is puzzling. One possible explanation is that the price level of public transport in Norway was close to a level where price was no longer a barrier (i.e., very low price elasticity). Thus, a small price reduction may have been enough to reach that level, and any further reduction in prices may not serve any purpose. However, the changes in demand reported in this study does not suggest price elasticity is low. If we calculated a weighted (by sample size) average price elasticity based on the data from this study, it would be -.94, which is more extreme than the -.3 often used as a rule of thumb or the -.4 identified as average short term (1–2 years) price elasticity in another meta-analysis (Paulley et al. 2006). An alternative explanation is that any price reduction might have been considered favorably by part of the population, regardless of the size of the price reduction. This part of the population may report that choosing public transport was due to this positive event (i.e., the price reduction), a statement that may be correct or that may represent a positive attitude toward the event rather than a fact about their decision to travel by public transport. In either case, it may be the price reduction as an event and not the new price level that creates the extra journeys and/or the positive attitude. However, we did not find a fixed effect of the price reduction trials. The meta-analysis indicates great variation in the effects.
The heterogeneous effects of the price reduction projects in this meta-analysis suggest there is great uncertainty related to the outcome of price reductions. As such, this study joins the ranks of previous reviews and meta-analyses concluding there is large and unexplained variation in the effect of price changes.

The studies included in this meta-analysis have some methodological limitations. The data are from cross-sectional studies, in which travelers were asked what they have done if the changes in price or frequency had not taken place. A longitudinal study with measures taken before and after the change in price or frequency would have stronger validity. Although the participants de facto were traveling with public transport after the change in price or frequency, their former and alternative travel behavior was measured subjectively through self-reports. Future research should measure changes in travel behaviour longitudinally and with more objective measures (e.g., observation or documentation). Future research should also include information on the level of the price and service frequency before the change, as well as other information on the routes (e.g., route length, population density and socioeconomic characteristics of the area) that may explain the differences in the results.

Acknowledgment

This research was financially supported by the Research Council of Norway and the Norwegian government’s Ministry of Transport and Communication.

References


**About the Author**

**INGE BRECHAN** *(inge.brechan@inn.no)* is a senior lecturer in psychology at Inland Norway University of Applied Sciences. This research was conducted while he was senior research psychologist at the Institute of Transport Economics in Oslo. His research interests include transport psychology, consumer psychology, health psychology, and environmental psychology.