TransInfo
Tier I University Transportation Center

FINAL REPORT
Behavior Reactions to Transit Network Disruptions: A Case Study on Washington Metro SafeTrack Project

Date of report: December 2018

Zhuo Yang, Graduate Research Assistant, George Mason University
Hamza Masud, Graduate Research Assistant, George Mason University
Guanqi Liu, Graduate Research Assistant, George Mason University
Shanjiang Zhu, PhD, Assistant Professor, George Mason University

Prepared by:
Side and Reva Dewberry Department of Civil, Environmental, and Infrastructure Engineering, George Mason University
4400 University Drive, MS 6C1, Fairfax, VA, 22030

Prepared for:
Transportation Informatics Tier I University Transportation Center
204 Ketter Hall
University at Buffalo
Buffalo, NY 14260

<table>
<tr><td>1. Report No.</td><td>2. Government Accession No.</td><td>3. Recipient's Catalog No.</td></tr>
<tr><td colspan="2" rowspan="2">4. Title and Subtitle

Behavior Reactions to Transit Network Disruptions: A Case Study on Washington Metro SafeTrack Project</td><td>5. Report Date
December 2018</td></tr>
<tr><td>6. Performing Organization Code</td></tr>
<tr><td colspan="2">7. Author(s)

Zhuo Yang, Hamza Masud, Guanqi Liu and Shanjiang Zhu,</td><td>8. Performing Organization Report No.</td></tr>
<tr><td colspan="2" rowspan="2">9. Performing Organization Name and Address

CEIE, George Mason University, 4400 University Drive, MS6C1, Fairfax, VA, 22030</td><td>10. Work Unit No. (TRAIS</td></tr>
<tr><td>11. Contract or Grant No.
DTRT13-G-UTC48</td></tr>
<tr><td colspan="2" rowspan="2">12. Sponsoring Agency Name and Address
US Department of Transportation
Office of the
UTC Program, RDT-30
1200 New Jersey Ave., SE
Washington, DC 20590</td><td>13. Type of Report and Period Covered
Final 06/2016 – 12/2018</td></tr>
<tr><td>14. Sponsoring Agency Code</td></tr>
<tr><td colspan="3">15. Supplementary Notes</td></tr>
<tr><td colspan="3">16. Abstract

The recent network disruptions in the Washington Metro system showed the new reality associated with aging transit infrastructure and highlighted the potential severity of such disruptions. However, relevant studies in the literature are limited and agencies need more empirical evidence to help them better planning and implementing maintenance work. To fill this research gap, this study analyzed both aggregated ridership data of the Washington Metro system collected from the National Transit Database, and the individual travel survey data collected from a National Science Foundation project to provide more empirical evidence on behavioral reactions to transit network disruptions. Particularly, this study highlighted the long-term impact of such disruptions on transit ridership, and the issues related to using stated preference or attitude survey data alone for planning purposes. Findings from this study would help transit agencies to better plan for future maintenance needs.</td></tr>
<tr><td colspan="2">17. Key Words

Washington Metro Safetrack Project, Transit Network Disruption, Travel Behavior</td><td>18. Distribution Statement

No restrictions. This document is available from the National Technical Information Service, Springfield, VA 22161</td></tr>
</table>

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	Complete	

Acknowledgements

This study is funded by the TransInfo University Transportation Center. The travel survey data used in this study was collected with the support of National Science Foundation Grant # #1649189: RAPID: Transit Network Disruption, Service Reliability, and Travel Behavior. The authors would also like to thank Dr. Lei Zhang and his research team at the University of Maryland for helping with the data collection and providing valuable comments and suggestions.

Disclaimer

The contents of this report 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's University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

Contents

1. Introduction

The transportation networks in the United States have been gradually developed since the early 1800s (Garrison & Levinson, 2014) and have been a great asset for the U.S. economy and quality of life. Over the years, travelers have witnessed the birth, rise, and maturing of many transportation systems, including inter-city railroads, highways, and public transit. As the systems matures and their period of rapid expansion ends, the management agencies have to deal with a reduced pace of investment and an aging infrastructure (Garrison & Levinson, 2014). Today, the aging infrastructure has become a severe problem in the U.S. and the transit system that serves many urban commuters and plays a key role for congestion mitigation in big cities, is in a particularly dire condition. The American Society of Civil Engineers' Infrastructure Report Card has released a comprehensive overview of the nation's infrastructure condition in 2017 (ASCE, 2017), and gives a D+ to the overall condition of the U.S. infrastructure system. Among the 16 categories of infrastructure, transit ranked the worst and received a D- with $90 billion backlogs of needs. The Report Card further stated: "the symptoms of overdue maintenance and underinvestment have never been clearer".

This chronic lack of funding has a significant impact on the reliability of transit services in many U.S. cities. According to the Mercury News in California, Ganfni (2016) reported that the Bay Area Rapid Transit (BART) would need months to get a power issue resolved. Barker (2018) reported through the Baltimore Sun News about an unexpected full shutdown of the entire DC metro system due to malfunctioning of power cables. One study (Perk, Flynn, & Volinski, 2008) based on on-board surveys conducted on 8 different transit systems in Florida showed that on average 17% of transit riders were dissatisfied or very dissatisfied about the service reliability and this number increased to 47.8% among commuters. Unexpected transit disruptions have been complained by lots of people through social media and have brought the issue to a national level (Harnack, 2016).

Although transit service disruptions are not uncommon, the recent network disruptions in the Washington Metro system showed the new reality associated with aging transit infrastructure and highlighted the potential severity of such disruptions as we continue with the current practice in transit funding and service provision. Due to the urgent need for accelerated track maintenance work among other safety enhancements, the Washington D.C. metro system (the Metro) initiated the SafeTrack project, which included a series of projects (dubbed as "surges") that required either single-track operations or complete shutdown of a line segment for extended time periods (ranging from a few days to 4 weeks) from Jun 2016 to Jun 2017 (WMATA 2016).

During SafeTrack, the Metro provided shuttle services that helped bridge metro riders between the affected metro stations. However, riders had to walk out of one metro station, take one of the shuttle buses, and get off at the other metro station, and continue with another metro train for the rest of the trip. This extra effort represented a significant burden for metro riders. In addition, the headways for the affected metro lines, even outside of the segment that was under work, were significantly increased during the SafeTrack, which also contributed to longer waiting time at the stations and thus longer travel time for the trip.

Facing this unprecedented transit network disruption in a major metropolitan area, some media outlets predicted that there would be a "transportation cataclysm" for the region (Washington Post, 2016). It is true that travelers had to adapt their travel plans to the new reality for an extended period of time, and sometimes in a major way. In response to major network disruptions, travelers may choose to 1) change departure time to account for the extra delay; 2) change mode; 3) change destination; 4) cancel the trip (including telecommuting), or 5) make no changes but accept the longer travel time and late arrival. The behavioral changes of metro riders may further affect traffic patterns on other systems, and the ripple effect spreads. However, there was no consensus either in the literature or in the media on how and to what extent the SafeTrack would affect metro riders and travelers of other modes in the region on an inter-connected, multi-modal transportation system.

In the literature, there have been very few studies on network disruptions in general and even fewer studies on transit network disruptions. Among the few exceptions, most studies focused on transit strikes and the majority of the strikes lasted only for a few days. To fill this research gap, this study analyzed both aggregated ridership data of the Washington Metro system collected from the National Transit Database, and the individual travel survey data collected from a National Science Foundation project (*Zhu et al. 2017*) to provide more empirical evidence on behavioral reactions to transit network disruptions.

2. Literature Review

2.1 Transit Network Disruptions

Transit network disruptions could be result of accidents, system failures, maintenance needs, and man-made or natural disasters. The impact of each disruption varies both in geographic and time dimensions (Chen, Zhang, He, Xiong, & Li, 2014). Replacing a breakdown bus may only take half an hour. However, it is much harder to restore a metro rail service when something goes wrong. For example, a simple runaway event in London created chaos among travelers early morning on August 13, 2010. A public inquiry was made due to a five-hour breakdown of Urban Transit Rail System, of Singapore, that disrupted the services for thousands of commuters on December 15, 2011. Unlike the traffic on road networks, it is almost impossible to reroute metro services (Lo & Hall, 2006). Bridging affected metro stations through a parallel bus service is a widely used practice to maintain the metro service (van Exel & Rietveld, 2001). Bikesharing was also used as an alternative means of transportation for travelers impacted by metro rail disruptions in the Washington DC area (Kaviti et al. 2018). However, significant delays could be added due to the transfers, and the limited capacity of buses compared to metro trains. These delays could cause repercussion on the entire network as travelers may miss their connections (Blumstein & Miller, 1983).

For an extended event, travelers are usually better informed and can adjust their travel behavior accordingly. For example, during the 13-day long transit strikes in New York City in 1967, 10% travelers canceled trips, 16.7% switched to carpool, and 50% drove alone. In a 1995 transit strike in Netherland, 30% travelers switched to driving, and another 10% canceled trips. Moreover, longer transit service disruptions could also have long-term effects on transit ridership. For example, the 1981 and 1986 Orange County transit strike in California reduced 15% to 20% of transit trips even after the strike (Di, Liu, Zhu, & Levinson, 2016). The New York City transit strike also caused 2.1-2.6% reduction in transit ridership. Zhu et al. (Zhu & Levinson, 2011) provided a detailed review of this topic. In an earlier study, Van (van Exel & Rietveld, 2001) summarized 13 major strikes throughout the world.

Though the aforementioned studies show the significance of transit service disruptions on travel behavior and transportation system performance, several critical research needs remain. Many previous studies relied on stated-preference, which may not capture the true travel behavior. Moreover, no study has investigated the learning and adaption process during the service disruption, which impedes us from modeling the re-equilibration process during such an event. This study will address those issues using a panel data collected both before and after the transit service disruptions.

The number of studies in the literature that analyzed extensive transit network disruptions is limited. Therefore, this study extends the literature review to cover studies on surface network disruptions. Research methods and data collection techniques may inform the current study.

2.2 Surface Network Disruptions

Similar to the transit network, road network could also be disrupted for an extended period of time due to maintenance, disasters, special events, etc. Given the much larger mileage of road network compared to the transit network, road network disruptions are more common and attracted more research interest in the literature.

Some road network disruptions affected one or a few key links on the network. In those cases, changing routes were the most natural and common behavioral reaction. Depending on the severity of the network disruptions, other behavioral changes, including changing departure time, mode, or destination, could also be observed.

For example, (Hunt, Brownlee, & Stefan, 2002) studied the 14-month long closure of the Center Street Bridge in the City of Calgary, Canada, using traffic counts and telephone surveys data. It was noticed that 4.4% less daily trips had been made and the morning peak hour had been shifted 15 minutes forward.

Another high-profile example of road network disruptions is the I-35W Bridge collapse in the Twin Cities, Minnesota, which took out an 8-lane Interstate bridge on a very busy commuting corridor for a year. (Zhu & Levinson, 2011) studied the traffic and behavioral effect using a variety of data, including traffic volumes and speed from sensors, paper-based surveys, and GPS-based surveys. The study found that the most common reactions were either changing the route, or changing departure time, while the overall travel demand in the region did not show a big reduction. However, travel time became longer for the trips commuting to downtown or the University of Minnesota. More interestingly, after the new I-35W bridge was restored, the traffic patterns did not go back to the patterns before the bridge pattern, clearly showing an asymmetric behavior shift (Di et al., 2016; Zhu, Levinson, & Liu, 2016).

Unlike road closures or bridge collapse that strike one or a few links on the network, natural disasters such as an earthquake could affect a large area and break many links simultaneously. The behavioral reactions to such events could be different to those of road closures or bridge collapse. Following the Northridge earthquake occurred in 1994, (Giuliano & Golob, 1998) collected traffic volumes of freeways and arterial roads with the help of the Los Angeles Department of Transportation (LDOT). They found that the traffic decreased 59% immediately after an interchange bridge ramp collapsed along I-5. The trips went up to 84% of its pre-earthquake capacity when the bridge had been restored to its 70%.

Regardless of the type of network disruptions, negative impacts were observed, including capacity losses, travel demand reduction, an increase of travel time, worse level of service, etc. Travelers have to adapt their travel routines to the new reality and may choose to change the route, departure time, mode, destination, or a combination of these options to mitigate the impact. If the disruption is predictable, traffic management agencies could develop mitigation plans and help travelers to better deal with the negative impacts of network disruptions. To develop such a plan, traffic management agencies need to better understand the potential impact of network disruptions in a quantitative way. However, such tools are not available due to the lack of empirical evidence and well-calibrated models based on real data.

Among the seven transit service disruptions reviewed in Zhu and Levinson (Zhu & Levinson, 2011), only one reported the impact on driving using traffic volume data observed in the field. However, four of them reported that some transit riders (from 10% to 50%) chose driving as the alternative mode, and two of them reported a shift from riding transit to carpool (16.7% to 28%). More importantly, after the service disruption ended, four studies reported a sustained loss in transit ridership, ranging from 0.3% to 20%. Moreover, no study has developed quantitative models to associate such impact with factors that transit operators and transportation management agencies could control or influence. Without such efforts, transportation agencies could not accurately predict the impact and develop countermeasures.

Kaviti et al (2018) studied the impact of disruptions of metro rail service on service on bikeshare ridership near rail stations affected by the disruptions. The study found that the bikeshare ridership increased during the period of disruption and reverted to normal levels after the repairs were completed.

Despite of these previous efforts in the related field, the studies on behavioral reactions to transit network disruptions were limited. The Washington Metro SafeTrack project provided a unique opportunity to further investigate this important research question.

3. Impact of Network Disruptions on Metro Ridership

3.1. NTD Data

This section analyzes the impact of the Washington Metro Safetrack Project on the metro ridership. The metro ridership data was collected from the National Transit Database (NTD). All US transit agencies who receive funding from the Urbanized Area Formula Program (5307) or Rural Formula Program (5311) are required to report a wide range of performance data to NTD. Two major transit ridership data, the Unlinked Passenger Trips (UPTs) and Passenger-Miles of Travel (PMTs) are reported annually, and the time series dates back to 1991. Since 2002, large transit operators are also required to report up-to-date time series of monthly UPTs, which are used in this study.

Seasonal effects clearly exist in monthly transit ridership data due to various reasons (weather, economic cycle, etc.). Figure 1 showed the monthly ridership of the Washington Metro system collected from the NTD. The snowfall data, which explains the major drops in ridership, is also plotted in the same figure.

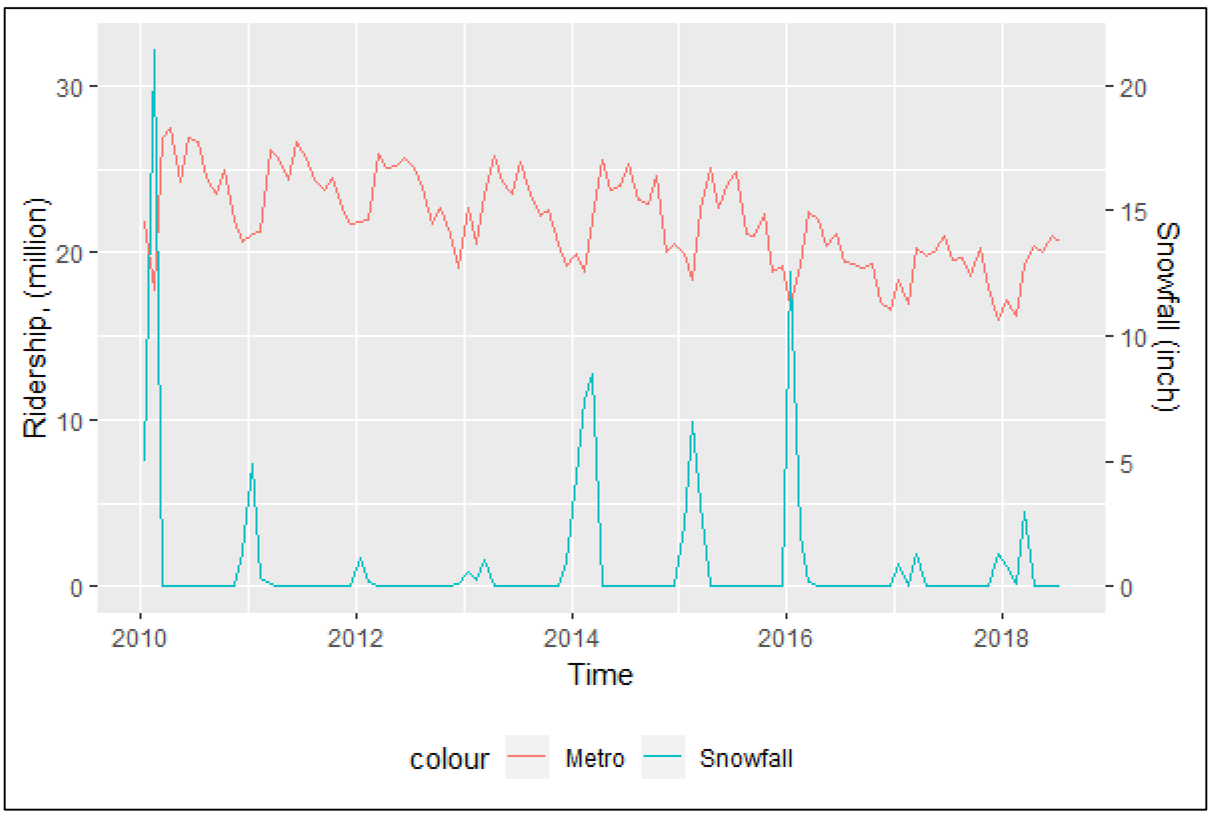

Figure 1: Comparison between Monthly Metro Ridership and Snowfall in DC MSA

3.2 Research Method

Because of such seasonal effects, monthly transit ridership is usually serial correlated, which makes conventional regression models problematic. Therefore, time-series analysis was usually adopted to predict the trend in monthly ridership. However, pure time-series analysis is a data-driven approach and does not take advantage of the existing knowledge on factors that may affect transit ridership. Therefore, this study will adopt the Seasonal Auto-Regressive Integrated Moving Average model (SARIMA) model with covariates, which combines the strength of regression models and time-series analysis, to analyze the monthly transit ridership. The inclusion of covariates, or independent variables, allows the researchers to consider the factors (such as fare, gas price, etc.) that have known the impact on the metro ridership, while the time-series component helps to control the serial correlations in residuals. Many researchers (Chiang, Russell, & Urban, 2011; Timmermann, 2006) showed that this integrated method provides the best individual forecast in many out-of-sample forecasting studies.

The model is described as follows:

$$\boldsymbol{y_t = \beta x_t + n_t} \qquad \textbf{Equation 1}$$

$$\boldsymbol{n_t = \phi_1 n_{t-1} + \cdots + \phi_p n_{t-p} - \theta_1 z_{t-1} - \cdots - \theta_q z_{t-q} + z_t} \qquad \textbf{Equation 2}$$

Where y_t is the monthly ridership at the metro at time t, $\boldsymbol{x_t}$ and $\boldsymbol{\beta}$ are the vectors of independent variables and coefficients, respectively. n_t is the residual for the linear regression. ϕ_s and θ_s are coefficients for the auto regression and moving average processes, while z_t is a white noise process. In addition, a seasonal process could be introduced to further control the seasonal effect.

3.3 Results

Using the metro ridership data from January 2010 to December 2017, a seasonal ARIMAX model could be estimated, which is summarized below in Table 1.

Table 1 Outputs of the SARIMAX Model for Monthly Transit Ridership

	SARIMAX (0,0,1)(1,0,0)[12]	
Variables	**Coefficients**	**S.E.**
ma1	0.5801	0.1022
sar1	0.8354	0.0587
intercept	20,294,226	3162050
Gas Price	962,356.6	272832.6
Metro Fare	-7,035,798	1361799

Precipitation	-34,029	29332.19
Temperature	29,701.09	17568
Snowfall	-105,450	16365.29
Inauguration	962,705.1	389164.5
SafeTrack Days	-71,058.8	11552.92
Post SafeTrack dummy	-958,598	488266.1
Silverline dummy	785,706.5	376947.1
Hurricane dummy	-1,186,407	455262.4
Cherry dummy	366,222.2	258523.9
Number of business days	561,548.2	65128.31
sigma^2 estimated as 5.501e+11: log likelihood=-1433.05		
BIC = 2939.13	MAPE = 2.52%	

MAPE: Mean absolute percentage error
BIC: Bayesian Information Criterion

The model is a seasonal time-series model with covariates, with 1 order of moving average, and 1 order of seasonal autoregression. Overall, the model provides a very good fitting, with a MAPE of 2.52% (see Figure 2). All coefficients are statistically significant, which are described one by one below.

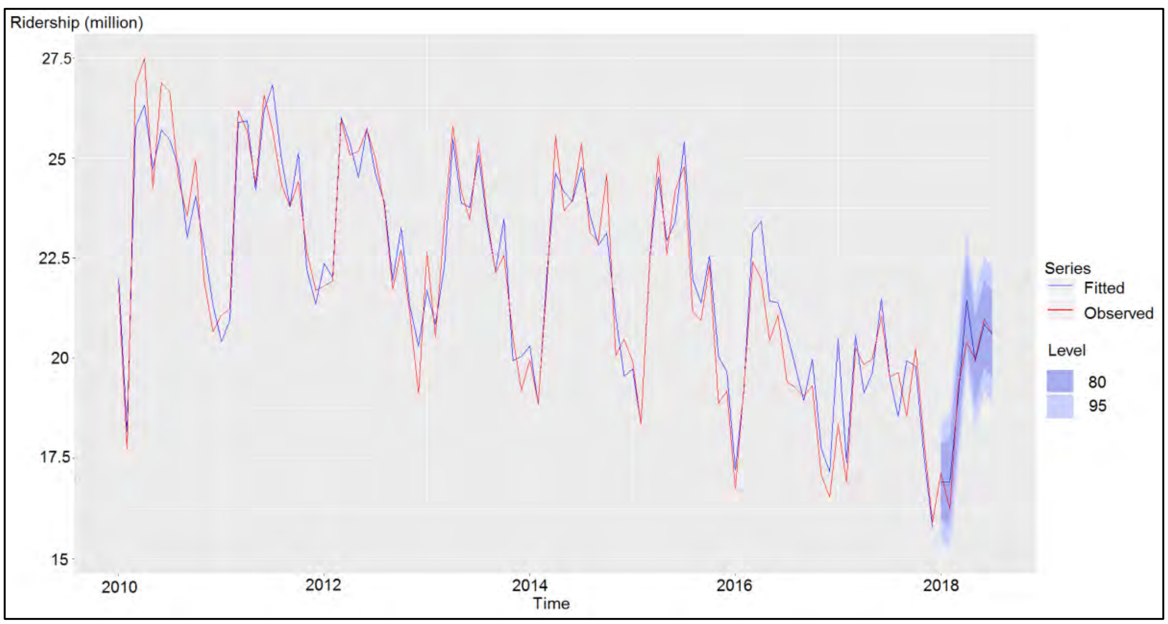

Figure 1: Actual and predicted DC metro ridership under SARIMA(0,0,1)(1,0,0)[12] model

SafeTrack Days

From June 2016 to July 2017, the Washington Metro system implemented a series of accelerated track maintenance projects. During SafeTrack, the Metro either reduced the service through single-track operation or shut down a line segment for a period ranging from several days to four weeks. After July 2017, the Metro conducted four additional track works in a similar way. Anecdotes in media coverage showed that the SafeTrack projects had a negative impact on the metro ridership. The research team tested this hypothesis by introducing the SafeTrack Days variable, which records the number of days in a particular month when SafeTrack projects were ongoing. To keep the tractability, the research team did not further differentiate workday vs. weekend and single-track operation vs. segment shutdown. The model shows that each day of SafeTrack projects reduced the monthly ridership by about 71,000, which is approximately 0.3% of monthly ridership.

Post SafeTrack Dummy

The SafeTrack may also have some long-term impacts on metro ridership according to media coverage and previous research in the literature. A dummy variable is encoded to represent the period after the SafeTrack projects were completed. The model shows that the monthly ridership dropped by about 958,000 after the SafeTrack, which is approximately 4% of the monthly ridership.

Gas Price

The gas price is the monthly average gas price per gallon for the Washington D.C. metropolitan area, provided by the Bureau of Labor Statistics. The model shows that increasing gas price would discourage driving and lead to higher metro ridership, with about 962 thousand more monthly riders for every $1 increase in gas price per gallon.

Metro Fare

The metro system in the Washington D.C. area uses a very complex distance-based fare scheme with an additional distinction between peak and non-peak hours. To keep the model trackable and easy to apply, we used the basic boarding charge (0-3 miles) during the peak hours (when the majority of travel happens) as the independent variable for the fare. During the study period, the Metro implemented several fare changes: from $1.65 in early 2010 to $2.25 as of September 2018, in six increments. The model shows that the fare increases have a negative impact on monthly metro ridership, with about a 7% decrease in monthly ridership for a one-quarter hike in basic boarding charge.

Inauguration

The inauguration is a dummy variable which equals 1 when a presidential inauguration ceremony occurred once every four years. It would bring about 962,000 additional metro trips for that particular month.

Silver Line Dummy

Phase 1 of the Washington Metro Silver Line, which includes 11.7 miles of new tracks and 5 new stations, is the newest addition to the Metro network. It was opened on July 26, 2014, and serves the Tysons Corner area, the new business hub in Northern Virginia. The model shows that it brought about 786,000 new trips to the Metro, which is about 3.3% of the monthly system-wide total.

Hurricane Dummy

The Hurricane Sandy hit the Washington Metro area in October 2012, leading to significant disruptions to the transportation system and a two-day closure of the federal government. It had a negative impact of about 1.2 million metro trips for that month.

Cherry Dummy

The cherry blossom festival is a major tourism attraction for the Washington D.C. Metropolitan area. The exact dates of the festival vary from year to year, depending on the weather. The peak falls either in March or April, so could not be completed captured by monthly patterns. A dummy variable is introduced to indicate the peak of cherry blossom, usually during which most tourists come. The model shows it would bring about 366,000 trips if the peak falls in that particular month.

Snowfall

The Washington Metro area would be hit from time to time by snowstorms, which usually lead to school and business closures and disruptions of traffic. A variable was introduced for the total inches of snowfall in the Washington Metropolitan Area (measured by the weather station at the Reagan National Airport). The model shows the snowfall also has a negative impact of metro ridership.

Number of Business Days

Intuitively, the number of business days in a month has a positive impact on the monthly ridership, with about 561,000 more trips with each additional business days.

Temperature

Temperature also has a positive and statistically significant impact of the monthly ridership. But the level of significance is weaker compared to other variables.

Precipitation

The precipitation has a small impact on the monthly ridership, and it is not statistically as significant as other variables.

Coefficients for ARMA

Using the Bayesian Information Criterion, the best fitting model includes one degree of moving average and one degree of a seasonal autoregression. Both coefficients are significant. Compared to a simple linear regression model, the seasonal time series model with covariates significantly reduced the serial correlations in residuals and improved the modeling fitting (MAPE improved from 3.97% to 2.52%).

The model provides the agencies the capacity to predict future metro ridership under various policy scenarios and social-economic conditions. Figure 3.2 showed the predicted monthly metro ridership and the actual monthly ridership for the first half of 2018. The red line represents the observed monthly metro ridership, while the blue line represented the fitted (01/2010-12/2017) and predicted (01/2018-07/2018) monthly metro ridership accordingly to the seasonal ARIMA model with covariates. For the period of 01/2018-07/2018, the actual monthly ridership falls within the band of confidence intervals, showing strong goodness of fit.

The statistical analysis of the aggregated ridership data provided important insights on the impact of various factors, including the metro safetrack project, on the metro ridership. The next section complements findings from this section by further analyzing the individual responses collected through survey data.

4. Individual Behavior Responses to Transit Disruptions

4.1. Panel Data

Between June 2016 and June 2017, the Washington Metro system either shut down or significantly reduced Metro rail transit services through continuous single track between stations to accommodate 16 separate SafeTrack system maintenance projects (dubbed as “surge”). This event provided a unique opportunity to improve our knowledge of travelers’ behavioural responses to major transit system disruptions and the consequent system mobility, reliability and resiliency impact.

Travel surveys were conducted both before and after each surge to collect panel data of planned and actual travel choices among metro riders. Before each surge, survey questionnaires were distributed at metro stations that would be severely affected by the particular SafeTrack project. Respondents may choose to complete the paper-based survey and mail them back to the research team using the pre-paid envelope, or to complete the web-based survey with the same questions by either visiting the survey website on computer or scanning the QR codes on the questionnaire with smartphone. Questions included respondents’ awareness of the metro shutdown event, the characteristics of their trips, their habits of traffic information acquisition (or lack thereof), the planned changes (or no change) due to the upcoming metro shutdown, and their social demographic information. Respondents were also asked if they would like to complete a follow-up survey after the particular SafeTrack surge, and for their contact information if they were willing to further participate.

A follow-up survey was mailed to the respondents who agreed to complete a follow-up survey. Questions included several alternatives to travel taken by respondents during the SafeTrack surge in reaction to the service disruptions, and the most effective method they eventually chose. Respondents also reported their new travel patterns after the metro service was completed restored. Figure 3 summarized the design and deployment of the survey instruments.

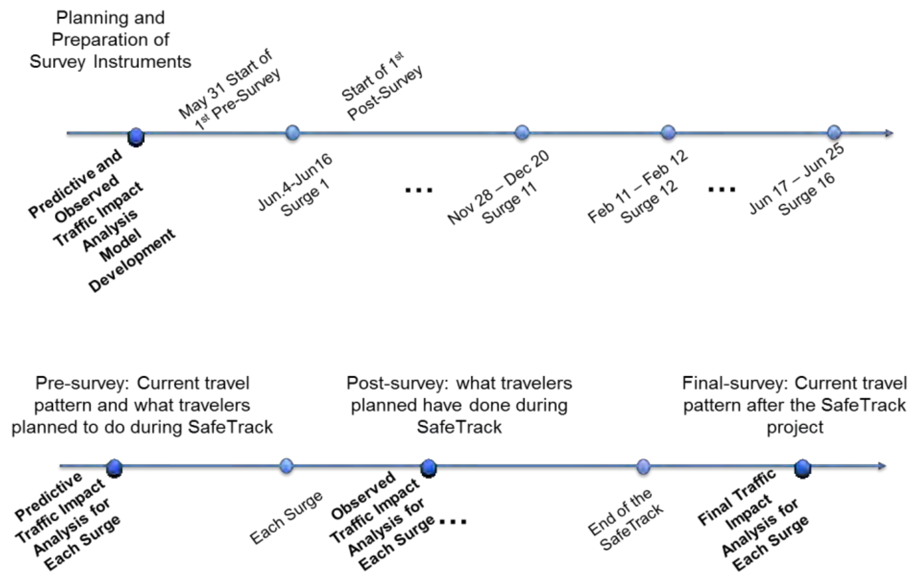

Figure 3: Design and deployment of the survey instruments for the SafeTrack study

4.2 Demographics

A total of 2943 people responded to the pre-event survey distributed before each SafeTrack surge. Table 2 summarizes the social demographic information of the respondents. Among all, 58% are female and 42% are male. The majority of survey respondents are between the age of 25 and 64, which is consistent with the facts that most metro riders during the peak hours are commuters. Similar to other surveys cited in the literature, the education and income distribution among respondents are skewed to the higher end. Moreover, the average income of metro riders is actually higher than the general public. Table 3 further compared the income distributions between respondents of the SafeTrack Survey and the respondents of the 2012 WMATA Metro Rail Passenger Survey (TPB, 2012). The two studies used consistent income brackets except for one point ($19,999 instead of $14,999). Considering that the SafeTrack survey focused on the peak hours (thus mostly commuters) instead of the entire day, the slight higher average income among SafeTrack survey respondents does not represent a significant bias. Thus, this study will only present results based on original survey results without further adjustments based on a re-sampling process.

Table 2: Demographics of Survey Respondents (Samples with no replies to a specific question were excluded when calculating percentages)

Total: 2943	Frequency	Percentage
Gender		
Male	1,216	42.0%
Female	1,678	58.0%
Age		
19-24	165	5.7%
25-34	659	22.8%
35-44	570	19.7%
45-54	652	22.6%
55-64	624	21.6%
65-74	178	6.2%
75+	15	0.5%
Other	24	0.8%
Education Level		
Less than sigh school	12	0.4%
High school graduate	69	2.4%
Some college	78	2.7%
Associate degree	245	8.5%
Bachelor's degree	882	30.6%
Graduate or professional degree	1,594	55.3%
Annual Household Income		
Less than $10,000	45	1.7%
$10,000 - $14,999	24	0.9%
$15,000 - $29,999	67	2.5%
$30,000 - $49,999	208	7.6%
$50,000 - $74,999	413	15.2%
$75,000 - $99,999	365	13.4%
$100,000 - $149,999	655	24.0%
$150,000 - $199,999	460	16.9%
$200,000 or more	489	17.9%

Table 3: Comparison of Income Distribution between 2012 WMATA Metro Rail Passenger Survey and the SafeTrack Study

Annual Household Income	Orange Line	Yellow Line	Green Line	Blue Line	Red Line	Average		SafeTrack Study
Less than $10,000	4.13%	4%	6.61%	3.85%	4.05%	4.36%	Less than $10,000	1.7%
$10,000 -	3.21%	4.00%	5.14%	3.45%	3.6%	3.75%	$10,000 -	0.9%

$19,999							$14,999	
$20,000 - $29,999	4.49%	4.82%	7.18%	4.19%	4.6%	4.87%	$15,000 - $29,999	2.5%
$30,000 - $49,999	12.66%	12.87%	16.36%	12.08%	12.69%	13.04%	$30,000 - $49,999	7.6%
$50,000 - $74,999	17.39%	17.17%	18.64%	17.49%	17.00%	17.30%	$50,000 - $74,999	15.2%
$75,000 - $99,999	14.74%	15.7%	14.5%	15.6%	14.51%	14.78%	$75,000 - $99,999	13.4%
$100,000 - $149,999	23.2%	21.34%	17.26%	22.42%	20.45%	20.79%	$100,000 - $149,999	24.0%
$150,000 - $199,999	12.94%	11.41%	8.65%	11.30%	11.58%	11.24%	$150,000 - $199,999	16.9%
$200,000 or more	11.38%	8.7%	5.66%	9.6%	11.51%	9.87%	$200,000 or more	17.9%

Of the 2943 respondents of the survey conducted before the SafeTrack surges, 609 responded to the follow-up survey that focused on their actual behavioral responses to the SafeTrack project. The survey was anonymous, and the two responses were linked with a unique respondent ID that was assigned to each respondent. This panel survey approach allows researchers to derive valuable empirical evidence of the consistency between transit riders' stated and actual responses to a major network disruption event, which has been largely speculated but rarely empirically evaluated in the literature.

4.3 Comparison between Stated and Actual Responses

A total of 2913 respondents reported their planned travel choices in response to the upcoming SafeTrack survey, 30.4% indicated that they would follow the travel routine; 23.3% planned to change departure time but stay with the metro; 32.1% planned to change to other modes; 6.0% would cancel the trip completely; and 8.1% planned to change their destination.

In comparison, among the 609 respondents who completed the follow-up survey and reported their actual travel choices during the SafeTrack period, only 16.4% stayed with their travel routine. A significantly higher number of respondents changed their departure time (30.7% instead of 23.3%) or cancelled trips (11.5% instead of 6.0%). The percentage of respondents who turn to an alternative traveling mode or destination is similar to that of the pre-survey, respectively 29.9% and 6.4%. The numbers are summarized in Table 4.

Table 4 Planned and Actual Behavioral Reactions to the SafeTrack Project

Change in Travel Behavior	**Planned**	**Actual**
No Change	30.4%	16.4%
Change Departure Time	23.3%	30.7%
Change Mode	32.1%	29.9%
Cancel Trip	6.0%	11.5%
Change Destination	8.1%	6.4%
Total	2913	609

The comparison showed a significant portion of travelers who hadn't planned to change their travel routine were forced to do something to meet their travel needs. And it seems most respondents chose to change departure time instead. However, this comparison, which was based on an average number, could be misleading and under-estimated the metro riders' imperfect travel plan when encountering a major network disruption. Table 5 further compared the planned and actual behavioral reactions of the same respondent. The row represents the planned reaction while the column represents the actual travel choices during the SafeTrack. If a respondent stick to the original travel choices, the number should be added to the corresponding diagonal cells. Most diagonal cells have a value lower than 50%. Although on average the number of metro riders planned to change mode and actually changed mode were similar, only 47.5% were consistent in their planned and actual behavior. More than half of respondents who planned to change mode switched to an alternative adaption plan, while respondents chose other options join the group who changed mode. In short, the consistency between the planned and actual travel plans in response to the metro SafeTrack project was low, which raises serious doubts on the approach to evaluate the potential impacts and develop mitigation plans for future network disruptions based on interviews and stated preference surveys.

Table 5 Confusion Matrix between the Planned and Actual Behavioral Reactions to the SafeTrack Project

	No change	Change departure time	Change destination	Change mode	Cancel trip	Count
No change	**25.5%**	37.3%	8.5%	22.9%	5.9%	168
Change departure time	14.6%	**46.2%**	5.1%	26.6%	7.6%	163
Change destination	14.6%	20.8%	**29.2%**	27.1%	8.3%	53
Change	12.9%	25.7%	3.0%	**47.5%**	10.9%	209

mode						
Cancel trip	22.2%	11.1%	0.0%	8.9%	**57.8%**	46

This study further explored the factors that may help to explain the imperfectness in developing travel plans during a major network disruption. Figure 4 compared the social demographic information of the group who were consistent in their planned and actual behavior, and the group who were not. However, the difference doesn't seem to be relevant to such factors.

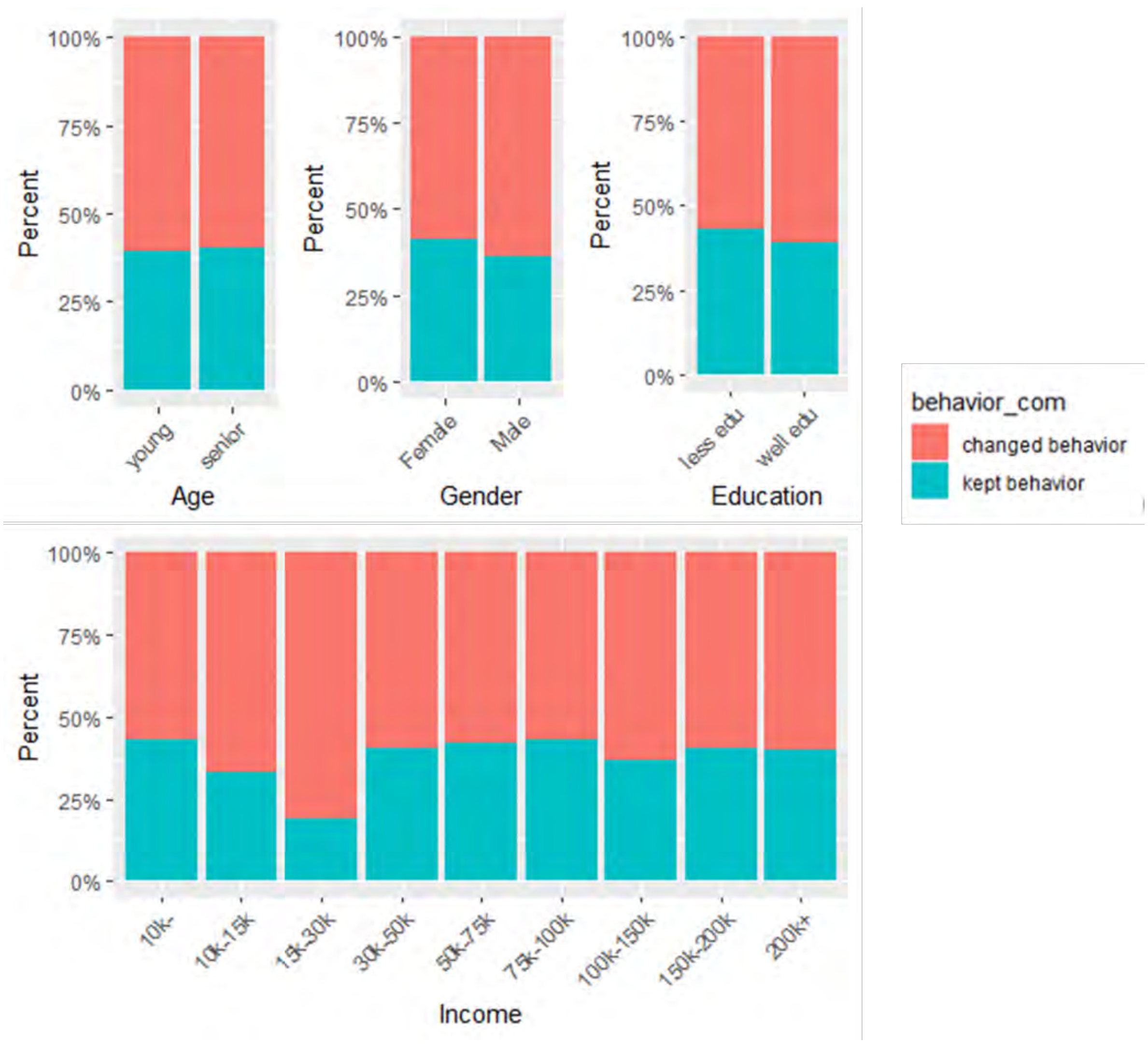

Figure 4 Behavior change across age, gender, and income.

Instead, the difference is more likely to be related to how well travelers were informed about the upcoming network disruptions and the frequency of their metro usage (which may be correlated with their familiarity with the metro system). The group who indicated they were well informed about the SafeTrack project made better traveling plans compared to the group who stated that they were not well informed (see figure 5). The frequent metro riders were more consistent between the planned and actual travel plans compared to the group who used the metro less frequently.

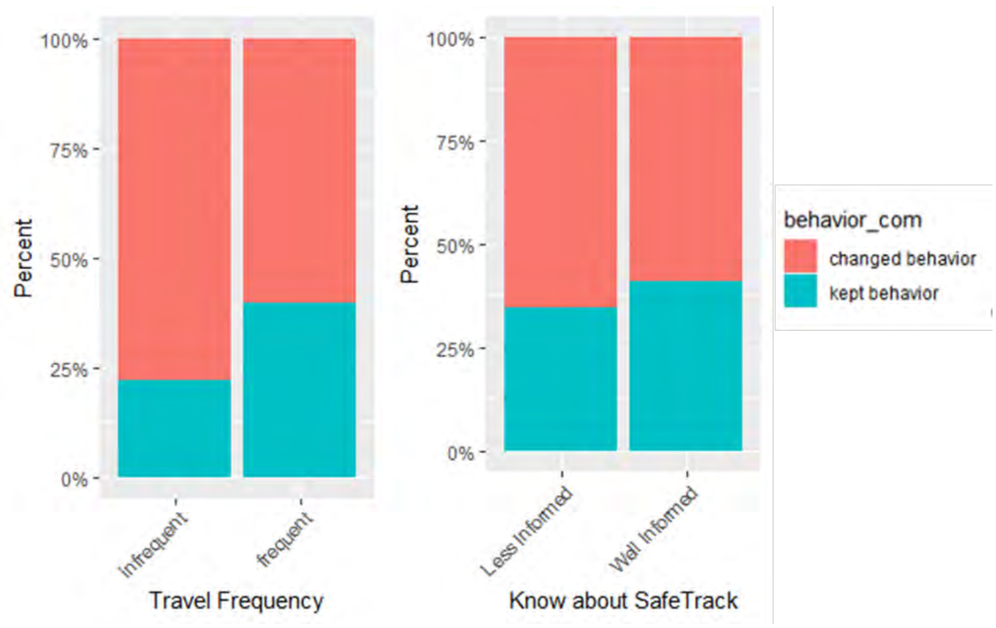

Figure 5: Behavior change by travel frequency and knowledge of the SafeTrack

4.4 Stay or Not Stay

A key question that faced by the transit operators, government agencies who are funding the metro services, and the public who relies on the metro system every day is the SafeTrack project's long-term impact on the metro ridership. Through the statistical analysis on aggregated metro ridership, the impact is detectable. However, it does not provide enough details on why people didn't return to the metro even after the services were restored. This study would provide empirical evidence on this critical question using the panel survey data.

Table 6 showed the actual mode choice among respondents who completed the follow-up surveys. Although about 30% chose other alternatives during the SafeTrack project (some people in the other choice category actually used the metro, but may choose other stations), only a little over 10% of them went back to the metro system even after the SafeTrack was completed. About 5% stayed with driving, either as a driver or as a passenger. About 3.6% chose ride-hailing services, 2.1% non-motorized modes, 3.6% regular buses, and 3.9% others (including MARC and VRE trains).

Table 6: Actual Mode Choice during and after the Metro SafeTrack Project

	During Surges	Immediately after Surges
Carpool	3.1%	2.3%
Drive alone	7.5%	2.8%
Metrorail	59.1%	81.8%
Regular Bus Service	5.0%	3.6%
Taxi	0.0%	0.0%
Uber, Lyft, etc	4.2%	3.6%
Walk or Bike	2.7%	2.1%
Other Mode	18.2%	3.9%
Total	**636**	**617**

This study further explored what factors may explain the decision of whether going back to the metro system after the services were fully restored. Among the total 617 respondents of the follow-up surveys, 138 chose an alternative mode. However, not all of them responded to the final survey were questions about their attitude to the metro system and their perceptions about its services. Only 78 subjects answered all the questions, and 50 of them went back to the metro services after choosing an alternative mode during the SafeTrack. Figure 6 plotted the two groups (travelers went back to the metro are in red, and those who did not are in blue) about demographic variables. It seems older people and people with higher education levels were more likely to go back to the metro system.

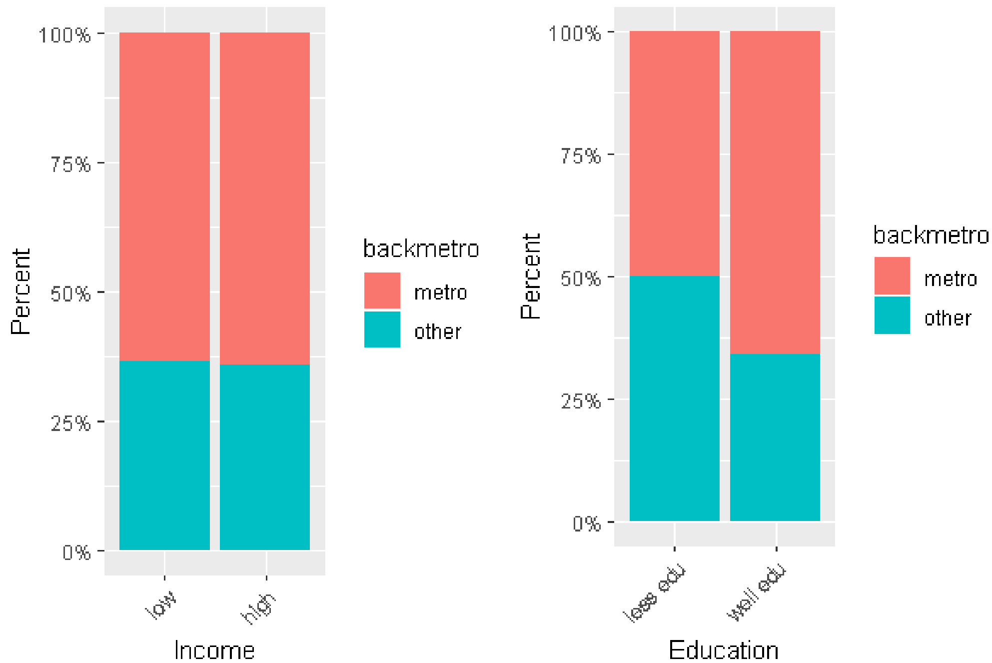

Figure 6: Decisions of whether Switching back to the Metro vs. Demographics

Figure 7 further correlates the decisions of whether switching back to the metro after choosing an alternative mode during the SafeTrack with subjective evaluation of the outcomes of the SafeTrack project and the performance of the metro system. It shows that metro riders who are positive about the reliability of the metro system after the SafeTrack project were more likely to go back to the metro. Similarly, people who are satisfied with the metro schedule and frequency were more likely to switch back to the metro system. Although the sample pool of subjects who responded to all questions was too small to support rigorous statistical analysis, results of the study still highlighted the importance of service reliability for the metro to attract riders.

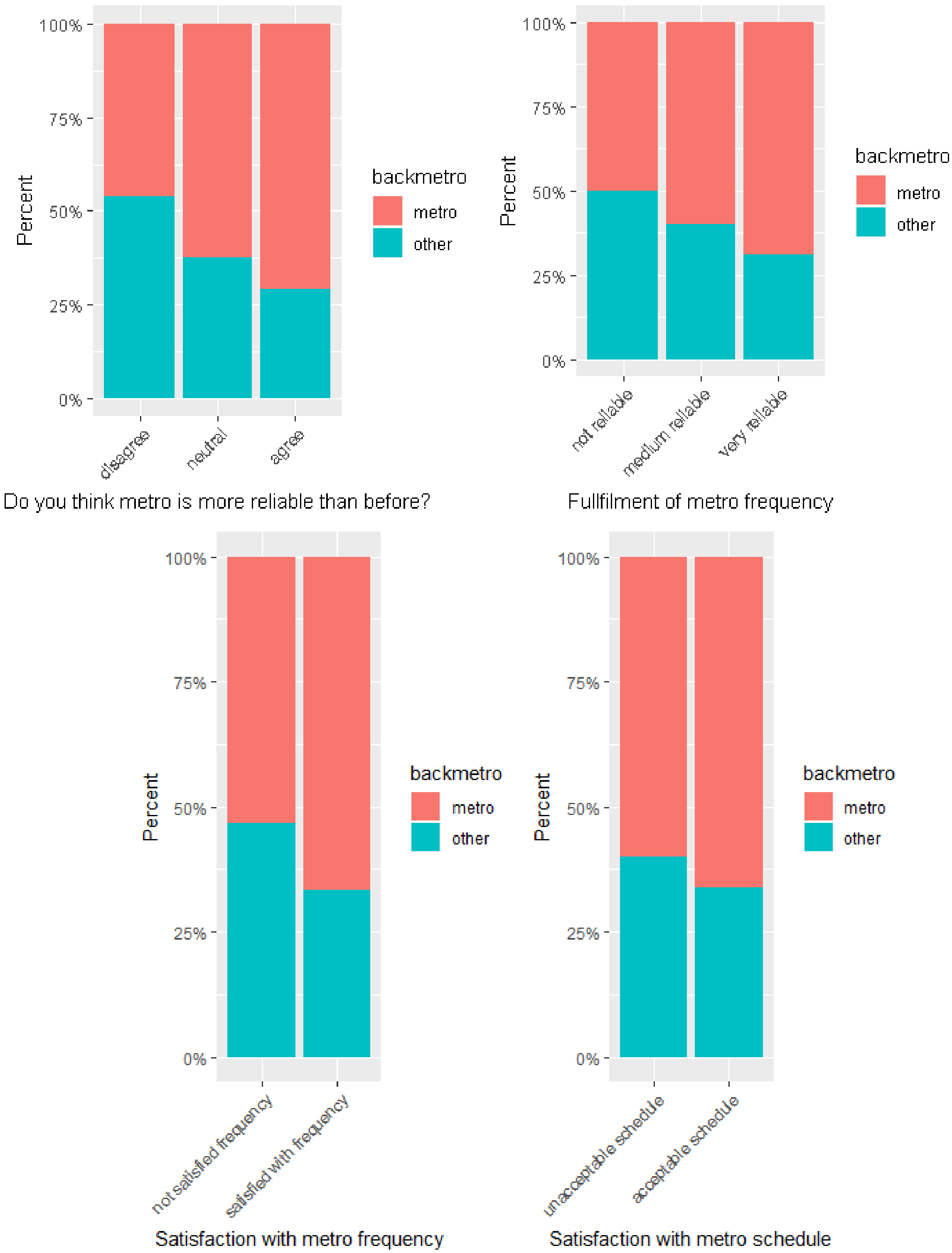

Figure 7: Decisions of whether Switching back to the Metro vs. Subjective Evaluation of the Metro Services.

6. Conclusions

This study developed a SARIMA model with covariates to investigate the impact of the Washington Metro SafeTrack project on the metro ridership while controlling a wide range of confounding factors, including the weather, the special event, the fare level, new service provisions, and gas prices. The method takes advantage of existing knowledge on factors that are likely to affect the monthly transit ridership, while also controls the serial correlations that are present in the residuals. The model provides excellent goodness of fit to the data and performs well in the out-of-sample forecasting.

This study clearly shows that SafeTrack put both short-term and long-term tolls on the metro ridership. During the SafeTrack, each day of SafeTrack reduced the monthly ridership by about 71,000 on average, which was about 0.3% of the monthly ridership. After the SafeTrack project, the monthly ridership dropped by about 958,000 on average compared to the level before the SafeTrack, which is approximately 4% of the monthly ridership. The study period extends to July 2018, which is one-year after the SafeTrack project ended. It has been long enough to consider any tempering effects. The impact of the period that is even longer than 1 year is to be observed in the future.

This study also provided empirical evidence for other factors that are known to be relevant to metro ridership, but there lacks a consensus on the magnitude in the literature. For example, gas price would have a negative impact on driving and thus a positive impact on metro ridership. Metro riders are sensitive to fare levels, with a quarter increase on the basic fare level leading to a 7% reduction in monthly ridership. Weather is also critical for metro ridership. The impact of a snow storm is very significant, while rain has less impact, but is still statistically significant.

In order to develop good models for metro ridership, some local knowledge is also necessary since many special events may significant affect the metro ridership. Examples in the D.C. area include the Cherry Blossom Festival and the Presidential Inauguration Ceremony.

The analysis of monthly transit ridership clearly showed some series correlations. Without controlling them, the model will be biased. This analysis also supports conclusions from earlier studies that the combined model (SARIMA with covariates) provides the best goodness of fit compared to either ARIMA or the regression models.

This study evaluated the stated and actual behavioral reactions to the SafeTrack project using a panel survey data. It found that facing major transit network disruptions, the most common reactions were doing nothing, changing departure time, or changing mode. Relatively fewer system users would cancel trips (mostly through telecommuting), but the number is still not negligible.

A comparison between the stated and actual behavioral reactions to the SafeTrack project showed that there were major differences between what transit riders planned and what they actually did in response to the transit network disruptions, and less than half of people showed behavior consistent with their stated intensions. This difference raised serious doubts on the practice of developing plans in preparation for future network disruptions based on stated preference data. The data suggested that a better level of familiarity with the system and more knowledge about the upcoming network disruption could help improve the consistency in behavior. More studies are needed in the future to establish that tie quantitatively.

This study also evaluated the likelihood of transit riders who chose an alternative mode during the network disruptions to switch back to the metro system. The results showed that people who are positive about the reliability of the system after the maintenance project, and those who are more satisfied with the system schedule and frequency were more likely to switch back to the metro system. However, the sample size was too small to test this hypothesis statistically.

This study provides important empirical evidence on the impact of transit network disruptions on transit ridership. It provides quantitative results that will help transit operators and traffic management agencies prepare for future events. Future research will develop more robust models that would support accuracy predictions of ridership and behavior reactions to transit network disruptions.

7. References

ASCE, A. (2017). ASCE's 2017 Infrastructure Report Card. Retrieved October 28, 2018, from https://www.infrastructurereportcard.org/cat-item/transit/

Barker, J. (2018, March 15). Entire D.C. Metro to shut down Wednesday at least 29 hours for inspections. Retrieved October 28, 2018, from http://www.baltimoresun.com/news/maryland/bs-md-dc-subway-shutdown-20160315-story.html

Blumstein, A., & Miller, H. D. (1983). Making do: The effects of a mass transit strike on travel behavior. *Transportation*, *11*(4), 361–382. https://doi.org/10.1007/BF00150724

Chen, X. (Michael), Zhang, L., He, X., Xiong, C., & Li, Z. (2014). Surrogate-Based Optimization of Expensive-to-Evaluate Objective for Optimal Highway Toll Charges in Transportation Network. *Computer-Aided Civil and Infrastructure Engineering*, *29*(5), 359–381. https://doi.org/10.1111/mice.12058

Chiang, W.-C., Russell, R. A., & Urban, T. L. (2011). Forecasting ridership for a metropolitan transit authority. *Transportation Research Part A: Policy and Practice*, *45*(7), 696–705. https://doi.org/10.1016/j.tra.2011.04.016

Di, X., Liu, H. X., Zhu, S., & Levinson, D. M. (2016). Indifference bands for boundedly rational route switching. *Transportation*, 1–26. https://doi.org/10.1007/s11116-016-9699-1

van Exel, N. J. A., & Rietveld, P. (2001). Public transport strikes and traveller behaviour. *Transport Policy*, *8*(4), 237–246. https://doi.org/10.1016/S0967-070X(01)00022-1

Ganfni, M. (2016, March 18). BART meltdown: Fix could take months. Retrieved October 28, 2018, from https://www.mercurynews.com/2016/03/17/bart-meltdown-fix-could-take-months/

Garrison, W. L., & Levinson, D. M. (2014). *The Transportation Experience: Policy, Planning, and Deployment*. Oxford University Press.

Giuliano, G., & Golob, J. (1998). Impacts of the Northridge Earthquake on Transit and Highway Use, 20.

Harnack, L. (2016, May 19). The Impact of Aging Infrastructure. Retrieved October 28, 2018, from http://www.masstransitmag.com/article/12210099/the-impact-of-aging-infrastructure

Hunt, J., Brownlee, A., & Stefan, K. (2002). Responses to Centre Street Bridge Closure: Where the "Disappearing" Travelers Went. *Transportation Research Record: Journal of the Transportation Research Board*, *1807*, 51–58. https://doi.org/10.3141/1807-07

Kaviti, S., Venigalla, M. M., Zhu, S., Lucas, K., & Brodie, S. (2018). Impact of pricing and transit disruptions on bikeshare ridership and revenue. Transportation. https://doi.org/10.1007/s11116-018-9904-5

Lo, S.-C., & Hall, R. W. (2006). Effects of the Los Angeles transit strike on highway congestion. *Transportation Research Part A: Policy and Practice*, *40*(10), 903–917. https://doi.org/10.1016/j.tra.2006.03.001

Perk, V., Flynn, J., & Volinski, J. M. (2008). Transit Ridership, Reliability and Retention. Retrieved from https://trid.trb.org/view/875466

Timmermann, A. (2006). Chapter 4 Forecast Combinations. In G. Elliott, C. W. J. Granger, & A. Timmermann (Eds.), *Handbook of Economic Forecasting* (Vol. 1, pp. 135–196). Elsevier. https://doi.org/10.1016/S1574-0706(05)01004-9

Washington Post, B. (2016). SafeTrack at one month: A disaster, mitigated - The Washington Post. Retrieved November 12, 2018, from https://www.washingtonpost.com/opinions/safetrack-at-one-month-a-disaster-mitigated/2016/07/01/529b8aa0-3f92-11e6-84e8-1580c7db5275_story.html?utm_term=.9b7ad16300fc

WMATA. (2016b). SafeTrack | WMATA. Retrieved October 28, 2018, from https://www.wmata.com/service/SafeTrack.cfm

Zhu, S., & Levinson, D. (2011). *Disruptions to Transportation Networks: A Review*. St. Louis, United States: Federal Reserve Bank of St Louis. Retrieved from http://search.proquest.com/docview/1698805255

Zhu, S., Masud, H., Xiong, C., Yang, Z., Pan, Y., & Zhang, L. (2017). Travel Behavior Reactions to Transit Service Disruptions. *Transportation Research Record: Journal of the Transportation Research Board*, *2649*, 79–88