Funding the Public Bus:
Does the amount of money agencies spend on the basic bus affect ridership?

Abstract

With my increasing awareness of the drop in ridership at the Denver Regional Transportation District, I thought it would be interesting to look nationally at what is happening and what policy drivers influence ridership across the nation. This analysis looks at how investment in the basic public bus might influence ridership. The public bus is the least popular of the modes of public transportation, but it also provides crucial service to areas and people that the more popular rail does not serve. I analyze data from the National Transit Database (NTD), which provides national data on an incredibly broad spectrum of variables. For the purposes of this study, I examine the bus, the commuter bus, and bus rapid transit which the NTD designates as three separate categories. I will only look at total capital use to analyze the most complete data. One obvious issue is that some agencies will have more money available to them than others and larger populations to serve. To control for this, I utilize population and area data from the census. Other variables that I can control for provided by the database are vehicles operated at maximum service, type of service (directly operated vs. purchase transportation), fixed guideway, agency type, agency geographic area type, and total agency capital.

I hypothesized that I would find that increased investment in the public bus will mean increased ridership, but there are so many other factors involved in ridership that it will be difficult to completely attribute that ridership to those dollars invested by the provider. This project will be helpful to develop an understanding of the complex issues involved in public transportation provision and the barriers facing bus transit. Providing transportation is such a complex issue and there are so many aspirational visions that planners have for our cities. Planners and policymakers expect that public transportation will solve (at least in part) so many issues that our communities face. By understanding how investment in the public bus performs, I will be able to draft better policy for that future.

Introduction

Public transportation, globally speaking, has allowed people to move great distances without needing to own a personal vehicle, and many cities plan around large transportation investments, usually a train line (subway, commuter rail, or light rail). However, the train is forced to follow a specific route, takes a large investment to get started, and does not always go where transportation is needed. The humble bus fills that role, moving people in more flexible and less costly ways around the city and providing a service in areas that the train investment has not or cannot be made. While the bus cannot take the place of a car in terms of flexibility or timeliness, many people are dependent on the bus service that many agencies provide to get around. However, some agencies are losing ridership, even as people are becoming more aware of the impacts of personal vehicle ownership on climate change, personal health, and even personal spending.

I will look nationally at reasons for dropping ridership as they relate to financial policies. Schank and Lewis (2013) point to loss of federal funding creating difficulties for state transportation agencies going forward. A couple studies looked at the loss of federal funding as an opportunity to improve policy in Kansas (Lorenz, 2011; Lorenz & Douglas, 2013). Connolly et al suggested eight steps intended to change the way that transit authorities operate in the face of funding loss (2012). Another study looked at why public support for transit does not match ridership and found that there might be two possible explanations: support of tax does not relate to increased transit spending, and therefore might not match support for transit, and support for spending is more closely correlated with concerns about social issues rather than thinking of transit as a method of personal transportation (Manville & Cummins, 2015). Ederer et al., utilizing NTD data, created agency clusters with similar characteristics related to ridership and compared changes within those clusters to isolate factors that would affect ridership; their results showed that causes are not uniform across clusters based on population, service levels, or right of way type (2019). An Australian study based out of Queensland focused on the effects of transit quality of service on bus ridership in suburbs and found that there is a strong relationship between the explanatory variables (service intensity, service span, travel time ratio, and topographic grade factor) and ridership, especially service intensity, suggesting that improving service intensity will increase ridership (Kashfi et al., 2015). A 2005 study looked at the impact of transit fare increase on ridership and revenue in the New York City Transportation system, and found that while a bus ridership decrease did occur, it was smaller than projected (Hickey, 2005). A second study in the UK also looked at the impact of fares (and other factors) on ridership. The intent of the study was to provide a guide on the factors affecting transit demand and found that fare elasticity, quality of service, and income and car ownership all had significant enough effects on demand to report on (Paulley et al., 2006). Alam et al. performed a national study at the metropolitan statistical area level examining internal (system specific such as fare, operating hours, etc.) and external (population density, median household income, etc.) system factors effects on ridership and found that both internal and external system factors affected ridership; the only internal factors that did not affect ridership were transit orientation pattern and the presence of rail service (2018). One study in Germany looked at the relationship between German policies surrounding costs for public transportation provisions and personal vehicle ownership and the increase in public transit demand (Buehler & Pucher, 2011). Another German study found that even though public transportation usage had stagnated in Germany, transit demand is more responsive to external gas prices than to system policies like reduced fares ((Frondel & Vance, 2011).

In this analysis, I will look at how investment in the basic public bus (rubber wheels on a road, powered by an engine) might influence ridership. I will analyze data from the National Transit Database (NTD), a database curated by the Federal Transit Authority (FTA), and collected from every agency providing public transportation across the nation (Federal Transit Administration, 2018). The NTD provides data on an incredibly broad spectrum of variables associated with public transportation. For the purposes of this study, I will examine agencies that provide the bus, the commuter bus, and bus rapid transit, which the NTD designates as three separate categories. The NTD provides capital use broken out by bus as well as by expenditure type (buildings, guideways, vehicles, maintenance, etc.) However, some agencies provide more data than others. Therefore, I will rely on total capital use to analyze a larger number of agencies. One obvious issue with this analysis is that some agencies will have more money available to them than others and larger populations to serve. The intent is therefore to normalize total capital use by passenger vehicle revenue mileage which will allow me to keep Unlinked Passenger Trips as part of the outcome variable. The NTD provides capital use both in total and by infrastructure sector investment, which will allow for a more nuanced investigation of investment factors. Other variables provided by the database that I can control for are: vehicles operated at maximum service, type of service (directly operated vs. purchase transportation), fixed guideway, agency type, agency geographic area type, and total agency capital.

Method

I collected NTD data for 2018 and examined the various tables. Utilizing Microsoft Access, I created a query narrowing down the agencies that provided bus services (since not all agencies in the United States do) and removed those agencies that did not spend any money on their bus service. This process reduced the total agencies under study by x% (or 20 agencies). In the resulting table, I combined bus data or removed duplicate data to provide a single incidence for each agency (rather than having multiple instances for some agencies). I am collecting population and area data from the Census and the Census's geospatial service, TigerWEB, and will join that with the current table to provide complete population, service area, and population density information (U.S. Census Bureau, 2010; US Census Bureau, 2018). Not every agency that reports to the FTA provides this data. Using census data will provide a proxy for this data, even if it does not exactly match the service area for every agency.

I have run several tests on the data to look at how the various variables relate to each other. There is a high correlation not only between an agency's expenditure on bus service, but an agency's total capital as well. There is less of a correlation between an urban vs. non-urban locale for an agency and its ridership, but there are further tests to try on other categorical variables in the dataset. Utilizing many of the variables in the dataset, I ran a Variance Inflation Factor test to check for collinearity on the variables in the dataset and after determining the final variables, ran two sets of correlation, testing not only total Bus Capital funds, but also dollars per mile and dollars per vehicle with and without the MTA New York City Transit system. This allowed me to look at not only the effect that the MTA has on data analysis, but also to look at what other variables may have an effect outside of the constraints of the linear regression.

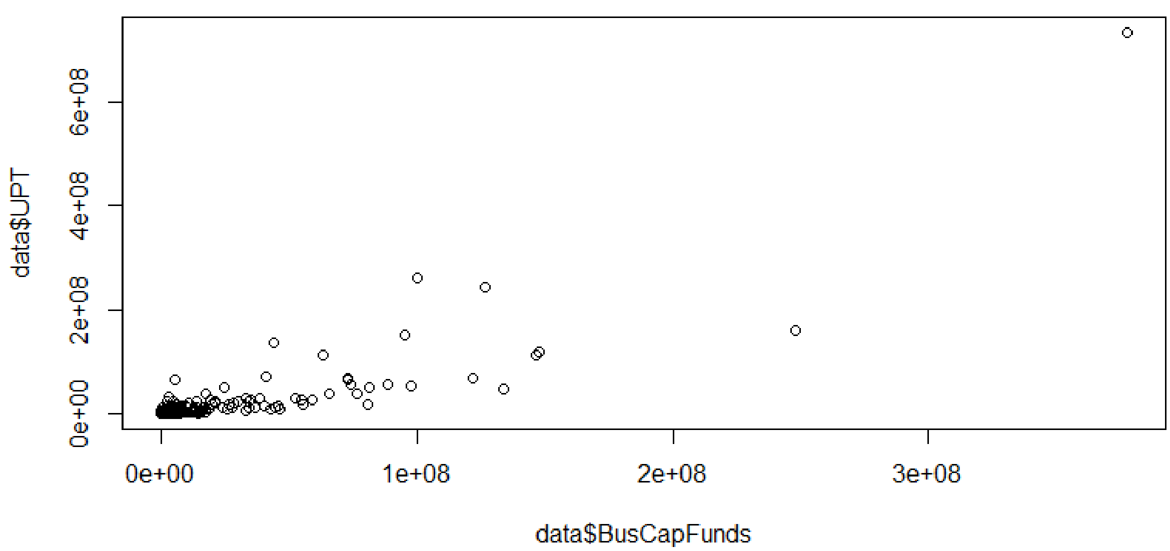
Analysis and Results

To begin with, I ran a correlation test between the two main variables of interest: Unlinked Passenger Trips (UPT) and the Bus Capital Funds. The intention of this is to determine whether there is a relationship between capital expenditures and ridership in an agency. The null hypothesis is that there is no relationship between the two

Table 1: Pearson's r test results from correlation

Data: Unlinked Passenger Trips and Bus Capital Funds
t = 43.593, df = 712, p-value < 2.2e-16
95 percent confidence interval: 0.8337297 0.8734038
Sample estimates: cor 0.8548109

Figure 1: Scatter plot of UPT and Bus Capital Funds.



variables. The alternate hypothesis for this test is that there is a relationship between Unlinked

Passenger Trips and the amount an agency spends on its buses. According to the Pearson's r test results, there is a strong positive correlation between UPT and bus capital funds, r (712) = .85, $p<0.001$. Thus, there is a strong and significant relationship between UPT and bus capital funds.

This relationship raises questions regarding how strong of a relationship there is between UPT and capital funds for bus when controlling for other factors like population and population density. Metropolitan areas with higher populations are going to have larger tax bases and therefore are going to be able to spend more money on their transportation systems.

Next, I performed an initial regression with the prevalent variables contained within the dataset to determine not only which reported variables might or might not affect the outcome variable. Refer to Appendix A for the complete regression results. The significant variables are square miles (area of the municipality) with a p-value of 0.148, the type of service provided by the agency, the vehicles operated at maximum service, revenue miles, and capital funds expended by the agency on bus all with p-values smaller than 0.001. It is interesting that in this regression, the mode type is not a statistically significant (bus rapid transit, commuter bus, and bus). Because this is an analysis of agencies that provide bus service and not an analysis of expenditures by individual mode types, I combined agency services to account for total expenditure on bus service, no matter the type. In future analysis, I intend to look at whether capital expenditures affect ridership by mode type.

Table 2: VIF for variables in initial regression.

	GVIF	Df	GVIFDf))
Agency Locale	1.468	1	1.212
Organization Type	3.198	11	1.054
Population	9.303	1	3.050
Density	2.206	1	1.485
Area in Square Miles	7.482	1	2.735
Mode	1.598	2	1.124
Type of Service	1.469	2	1.101
Vehicles Operated at Maximum Service	26.444	1	5.142
Revenue Miles	21.317	1	4.617
Dollars Per Mile	3.435	1	1.853
Dollars Per Vehicle	3.605	1	1.899
Bus Capital Funds	5.348	1	2.313
Agency Funds	1.065	1	1.032

Table 2 provides the results of the Variance Inflation Factor test for multicollinearity. Any values above five or ten represent a variable that is in some way colinear with another variable in the regression (James et al., 2017). By removing at least one of the collinear variables, I can improve the regression model by removing redundant variables and reperforming the linear regression. I can identify the best variables to remove by comparing the insignificant variables from the regression and the variables that have the highest collinearity. Theory suggests that variables like population, density, and the size of the municipality in square miles will also have an effect on the outcome variable but because density is a calculation of population per area, I can utilize it to approximate one or the other; population has the higher collinearity factor, so I will remove population, but keep area and density. Vehicles operated at maximum service and revenue miles are highly collinear, but also highly significant according to the initial regression. Therefore, I will keep revenue miles and remove both VOMS and dollars per mile (a calculation of bus capital funds and revenue miles). I believe that by removing agency capital I will remove the collinearity of the bus capital funds, which is one of the main predictor variables of interest. The other predictor variables of interest are Dollars Per Mile and Dollars Per Vehicle, which will all be in separate regressions, so I am not concerned about any even minor collinearity they are showing. Therefore, the final predictor variables are agency locale, density, area in square miles, type of service, revenue miles, and bus capital funds. Table 3 below contains summary data for the variables.

With the variables determined, I ran two sets of regressions, one with the full dataset, and one with the MTA New York City Transit system removed from the analysis. Even though the Los Angeles County Metropolitan Transportation Authority (the nearest neighbor to the MTA New York City transit system) serves a similar population density (LA: 6999 persons/sq. mi. and NY: 5319 persons/sq. mi.) and population size (LA: 12,150,996 and NY: 18,351,295), the MTA New York City Transit System has nearly three times the number of Unlinked Passenger Trips per Year of LA (734,660,697 to LA's 260,902,211). This means that the New York Metro System might behave differently in a regression and skew the results. By comparing the results before and after the removal of the New York Metro System, we can see which model performs better and will be a better predictor of policy results. See Tables 4 and 5 for a side by side comparison of the results of the regressions. See Appendices B and C for the VIF test results for these regressions. As none of the VIF results did not exceed 5, none of the variables are collinear.

To begin with, I look at the R^2 values for both regressions. For the full dataset, the R^2 values show that the linear regression model fits very well in both cases (with and without MTA). However, the R^2 values are higher in all cases for those linear models on the dataset without the MTA, meaning the model fits much better when the MTA has been removed from the equation; R^2 (706, N=714) = 0.804, 0.752, and 0.751 with MTA, and R^2 (705, N=713) = 0.834, 0.818, and 0.817 without MTA. Next, while all of the models are statistically significant, the F-statistics values are again higher in the model where the MTA has been removed from the dataset; F(df=7,706, N=714) = 414.2, 305.1, and 304.4, p<0.01 with MTA, and F(df=7,705, N-713) = 506.7, 453.1, and 450.8, p<0.01 without MTA. Since both models significantly fit the data presented, it is interesting to see such a large improvement in the model fit with just the removal of the one datapoint. In a previous iteration of this project, I did temporarily look at removing not only the MTA, but other "outliers" as well, and while there was improvement in those cases, the fit did not improve so much that it made removing over 20 datapoints worth it. Additionally, as mentioned previously, the Denver Regional Transportation District System is of personal interest to me; by removing the "outliers," I removed RTD from the analysis, which seemed to be counterproductive to this analysis.

Next, I move onto the effect of the variables themselves. In both regressions, the total amount of capital expended on bus is significant in terms of ridership. The t-value is well above the cutoff of 1.96 for both. However, values that we expect to affect ridership like area size and population density are not significant or only significant at a 95% confidence level. It is especially interesting that population density is only significant once MTA is removed; whether this is due to the population data use (Census vs Service Area population) and therefore creating a false density measure is not clear. That will be something I will have to analyze in the future; service area population is provided as part of the NTD database, but not for every agency and no easily acquired dataset can serve as a proxy for that data.

The other values of interest are the dollars per mile and dollars per vehicle. However, when looking at the regressions, these values are not significant at any level and they do not pass the t-test (value higher than 1.96) nor is the confidence level higher than 90%. In the one case that the confidence level is above 90%, the t-value is not high enough to indicate significance. This indicates to me that there is no hard and fast rule for total dollar amounts of spending by an agency, i.e. if an agency were to spend so many dollars per mile or per operating vehicle, they would get so many riders. However, because all else being equal, there is a positive relationship between expenditures and ridership, there may be another relationship between the allocation of those capital funds and bus ridership.

Conclusion

Based on the results of the regressions, the money an agency spends on bus has some effect on ridership, controlling for population density, municipality size, and other service factors. In the future I would like to look more in depth at the mode types, although for this regression mode type did not seem to be a factor in ridership levels. As mentioned earlier, to study by agency, I combined the mode data; I wonder if that will change if I reexamine the data with the modes kept separate. I also intend to look at the information provided for bus service looking at where the dollars each agency spends are allocated. It is possible that where an agency spends its money is just as important as how much.

This analysis does not happen in a vacuum. With the threats and changes that the current pandemic poses to public transportation, it is even more important to see what makes an effective public transportation system and to consider how and why public transportation will play a role in the future. Looking at how effective agency dollars is just one piece of that larger puzzle.

Table 3: Data Summary Statistics

	Obs	Mean	Prop.	Std. Dev.	Min	Max
Outcome Variable						
UPT						
Annual Total Unlinked Passenger Trips	714	6015458		33684560	543	734660697
Predictor Variables						
Bus Capital Funds						
The total funds expended by a capital agency on the bus.	714	$6,468,798		$23,602,631	$525	$377,617,697
Dollars Per Mile						
The number of dollars an agency spends in capital per revenue mile	714	$2.09		$3.22	$.003	$33.68
Dollars Per Vehicle						
The number of dollars an agency spends in capital per vehicle operated at maximum service	714	$77,676.40		$112,083.40	$93.30	$1,034,153.30
Square Miles						
Area of the municipality, census designated place, or urbanized area in which the agency is located.	714	369.6		695.5	0.0	3,450.0
Revenue Miles						
The number of miles that a vehicle is scheduled to or actually travels while in revenue service.	714	2,614,969		7,421,893	2,541	98,385,713
Population Density						
Persons per square mile of the municipality, census designated place, or urbanized area served by the agency.	714	2,232		1,406.18	10	9,188
Type of Service						
The method by which an agency provides bus transportation to those it serves.	714					
Directly Operated			74.9%			
Purchased Transportation			23.1%			
Agency Locale						
Agency location in an urban or non-urban area.	714					
Urban			72.3%			
Non-Urban			27.7%			

Sources: (Federal Transit Administration, 2018; Hlavac, 2018)

Table 4: Linear Regression Results (Full Dataset)

	Dependent variable:		
	UPT		
	(1)	(2)	(3)
Bus Capital Funds	0.629***		
	t = 13.833		
Dollars Per Mile		223,171.900	
		t = 1.123	
Dollars Per Vehicle			-0.126
			t = -0.022
Area in Square Miles	-1,382.977	-1,651.038	-1,646.095
	t = -1.353	t = -1.434	t = -1.428
Revenue Miles	2.435***	4.194***	4.196***
	t = 15.664	t = 41.692	t = 41.646
Population Density	217.502	156.433	173.566
	t = 0.450	t = 0.287	t = 0.318
Type of Service – Directly Operated	30,863,194.000***	39,453,654.000***	39,410,353.000***
	t = 7.085	t = 8.122	t = 8.105
Type of Service – Purchased Transportation	30,721,373.000***	39,237,387.000***	39,389,517.000***
	t = 6.879	t = 7.875	t = 7.899
Agency Locale - Urban	-3,460,040.000**	-3,384,647.000**	-3,268,375.000**
	t = -2.563	t = -2.221	t = -2.130
Constant	-32,118,139.000***	-41,341,097.000***	-40,996,115.000***
	t = -6.973	t = -8.033	t = -7.957
Observations	714	714	714
R^2	0.804	0.752	0.751
Adjusted R^2	0.802	0.749	0.749
Residual Std. Error (df = 706)	14,979,293.000	16,872,554.000	16,887,603.000
F Statistic (df = 7; 706)	414.218***	305.111***	304.387***
Note:			*p<0.1; **p<0.05; ***p<0.01

(Hlavac, 2018)

Table 5: Linear Regression Results (New York Metro System Removed)			
	Dependent variable:		
	UPT		
	(1)	(2)	(3)
Bus Capital Funds	0.223***		
	t = 8.472		
Dollars Per Mile		180,978.800*	
		t = 1.816	
Dollars Per Vehicle			1.642
			t = 0.566
Area in Square Miles	-819.042	-861.990	-843.423
	t = -1.486	t = -1.493	t = -1.457
Revenue Miles	2.343***	2.875***	2.874***
	t = 27.939	t = 49.528	t = 49.376
Population Density	575.186**	574.846**	584.159**
	t = 2.206	t = 2.104	t = 2.134
Type of Service – Directly Operated	16,128,121.000***	17,679,617.000***	17,660,444.000***
	t = 6.789	t = 7.127	t = 7.103
Type of Service – Purchased Transportation	15,378,477.000***	16,765,351.000***	16,839,243.000***
	t = 6.313	t = 6.587	t = 6.602
Agency Locale - Urban	-1,557,373.000**	-1,447,773.000*	-1,410,384.000*
	t = -2.135	t = -1.892	t = -1.829
Constant	-17,642,393.000***	-19,557,055.000***	-19,362,593.000***
	t = -7.034	t = -7.455	t = -7.366
Observations	713	713	713
R^2	0.834	0.818	0.817
Adjusted R^2	0.833	0.816	0.816
Residual Std. Error (df = 705)	8,076,533.000	8,457,924.000	8,475,763.000
F Statistic (df = 7; 705)	506.664***	453.122***	450.793***
Note:			*p<0.1; **p<0.05; ***p<0.01

(Hlavac, 2018)

Bibliography

1. Alam, B. M., Nixon, H., & Zhang, Q. (2018). Factors Affecting Travel Demand by Bus: An Empirical Analysis at U.S. Metropolitan Statistical Area Level. *Transportation Research Record*, *2672*(8), 817–826. https://doi.org/10.1177/0361198118798714

2. Buehler, R., & Pucher, J. (2011). Making public transport financially sustainable. *Transport Policy*, *18*(1), 126–138. https://doi.org/10.1016/j.tranpol.2010.07.002

3. Connolly, K., Wolfgram, L., & Santos, J. (2012). Steps to Leading Change: A Transportation Agency's Business Plan Approach. *Transportation Research Record*, *2274*(1), 15–22. https://doi.org/10.3141/2274-02

4. Ederer, D., Berrebi, S., Diffee, C., Gibbs, T., & Watkins, K. E. (2019). Comparing Transit Agency Peer Groups using Cluster Analysis. *Transportation Research Record*, *2673*(11), 505–516. https://doi.org/10.1177/0361198119852074

5. Federal Transit Administration. (2018). *NTD Data Products* [Text]. Federal Transit Administration. https://cms7.fta.dot.gov/ntd/ntd-data

6. Frondel, M., & Vance, C. (2011). Rarely enjoyed? A count data analysis of ridership in Germany's public transport. *Transport Policy*, *18*(2), 425–433. https://doi.org/10.1016/j.tranpol.2010.09.009

7. Hickey, R. L. (2005). Impact of Transit Fare Increase on Ridership and Revenue: Metropolitan Transportation Authority, New York City. *Transportation Research Record*, *1927*(1), 239–248. https://doi.org/10.1177/0361198105192700127

8. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2017). Collinearity. In *An Introduction to Statistical Learning with Applications in R* (pp. 99–102). Springer Science.

9. Kashfi, S. A., Bunker, J. M., & Yigitcanlar, T. (2015). Effects of Transit Quality of Service Characteristics on Daily Bus Ridership. *Transportation Research Record*, *2535*(1), 55–64. https://doi.org/10.3141/2535-06

10. Lorenz, J. (2011). Building Support for New Transportation Funding and Financing Program: Linking Investments to Stakeholder Priorities Through Collaboration and Economic Impact Analysis. *Transportation Research Record*, *2245*(1), 8–16. https://doi.org/10.3141/2245-02

11. Lorenz, J., & Douglas, L. (2013). Less Money, Better Policy: Possible. *Transportation Research Record*, *2345*(1), 1–8. https://doi.org/10.3141/2345-01

12. Manville, M., & Cummins, B. (2015). Why do voters support public transportation? Public choices and private behavior. *Transportation; New York*, *42*(2), 303–332. http://dx.doi.org.aurarialibrary.idm.oclc.org/10.1007/s11116-014-9545-2

13. Paulley, N., Balcombe, R., Mackett, R., Titheridge, H., Preston, J., Wardman, M., Shires, J., & White, P. (2006). The demand for public transport: The effects of fares, quality of service, income and car ownership. *Transport Policy*, *13*(4), 295–306. https://doi.org/10.1016/j.tranpol.2005.12.004

14. Schank, J., & Lewis, P. (2013). Federal Role in State Transportation Finance. *Transportation Research Record*, *2345*(1), 9–16. https://doi.org/10.3141/2345-02

15. U.S. Census Bureau. (2010). *TIGERweb*. https://tigerweb.geo.census.gov/tigerwebmain/TIGERweb_main.html

16. US Census Bureau. (2018, October 15). *About* [Government]. The United States Census Bureau. https://www.census.gov/programs-surveys/metro-micro/about.html