NCTR
NATIONAL CENTER for TRANSIT RESEARCH

Estimating Ridership of Rural Demand-Response Transit Services for the General Public

August 2016

prepared for

US DOT

prepared by

Jeremy Mattson

North Dakota State University
Upper Great Plains Transportation Institute
Small Urban and Rural Transit Center

Estimating Ridership of Rural Demand-Response Transit Services for the General Public

Prepared for:

U.S. Department of Transportation

Prepared by:

Jeremy Mattson

Small Urban and Rural Transit Center
Upper Great Plains Transportation Institute
North Dakota State University
Fargo, ND

Final Report

21177060-NCTR-NDSU08

August 2016

Acknowledgements

Disclaimer

The contents presented in this report are the sole responsibility of the Upper Great Plains Transportation Institute and the authors.

ABSTRACT

The objective of this study is to develop a model for estimating demand for rural demand-response transit services for the general public. Lack of data for demand-response service characteristics and geographic coverage has limited the development of such models. This study developed two models for estimating demand. The first used data from the 2013 rural National Transit Database, and the second used more detailed service data collected from surveys of transit agencies. Results showed that in addition to total population, demographic characteristics are important. Ridership was found to significantly increase when the percentage of the population comprised of older adults or people without access to a vehicle increased. Both models showed a negative effect of fares on ridership. The second model analyzed the impacts of service span and reservation requirements on ridership. Results showed that providing more days of service had an expected positive impact on ridership, while allowing users to reserve rides on shorter notice also had a significant positive effect. Compared to previous research, including a greater number of variables and more specific service information improved the performance of the models.

EXECUTIVE SUMMARY

Few studies have provided a method for estimating ridership for demand-response transit services. Lack of data for demand-response service characteristics and geographic coverage has limited the development of such models. While demand models have been previously estimated, such as in TCRP Report 161 (Vanasse Hangen Brustlin, Inc, et al. 2013), they lack many service characteristics that will likely impact ridership, such as fares, span of service, and reservation requirements.

The objective of this study is to develop a model for estimating demand for rural demand-response transit services for the general public. Specific objectives are to estimate the impacts of service characteristics (such as span of service, service coverage, fares, and reservation requirements) and service-area characteristics (such as population and demographic characteristics) on ridership.

Two models were developed. The first used data from the 2013 rural National Transit Database (NTD) and the American Community Survey (ACS). Because data limitations of the rural NTD restrict the number of variables that can be analyzed, a second model was developed using data collected from surveys of rural transit agencies. Both models estimate demand for non-sponsored demand-response service for the general public. Demand for sponsored, or program-related, trips or ADA paratransit was not considered.

Factors Affecting Ridership

It is expected that ridership will be determined by the demand for services, the level of service provided, and the cost of service. Demand for services can be estimated based on population and demographic characteristics. It is expected that demand will be greater in areas with a higher population and that concentrations of transportation-disadvantaged populations will create greater demand. As noted in previous research, it is expected that demand for rural transit will be greater in areas with a larger population of older adults, people with disabilities, and people without access to a personal vehicle.

While population characteristics can be used to estimate demand, ridership will also be influenced by the characteristics of the service provided and the fare levels. Important service characteristics include the number of days per week and hours per day that service is provided, as well as the advance reservation requirements. It is hypothesized that reservation requirements serve as a deterrent and that requiring reservations further in advance will negatively impact ridership.

Other service characteristics could also affect ridership. Some rural transit agencies operate both fixed-route and demand-response service. In those cases, some of the transit demand is being met by fixed-route services, so demand-response ridership is expected to be lower than it would have been had the fixed-route services not been available. In other cases, there may be overlap in service areas, with more than one provider serving an area. Ridership is expected to be lower for demand-response agencies that are sharing service area with other agencies. It may also be expected that ridership for tribal transit may differ from that of other rural transit systems because of the unique geographic and economic characteristics of Native American reservations.

Due to the inefficiencies of providing service in rural areas, transit agencies in rural, low-density areas with long travel distances may not be able to provide as many trips as agencies serving more concentrated and densely population areas. Agencies that strictly serve a municipality, on the other hand, may be able to provide more trips per capita than those serving larger geographic areas because of shorter trip distances and greater population density.

Methods

First, a model was developed that estimated ridership as a function of service area population, demographic characteristics of the service area, and fare levels, and it took into consideration whether the transit agency provides a fixed-route service, has a service area that overlaps with that of another demand-response transit provider, operates only in a municipality, or is a tribal agency. It also accounted for the region of the country in which the agency operates.

This first model used data from the rural NTD for 731 rural demand-response transit agencies, but since this database does not include important service quality characteristics, those factors could not be included in the model. Previous studies have also not included this information due to lack of data. As noted in the Transit Capacity and Quality of Service Manual (Kittelson and Associates et al. 2013), key measures of demand-response transit quality of service include geographic service coverage, service span (days and hours of service), and response time, which refers to the amount of time users must reserve a ride in advance. These three variables are measures of service availability and are expected to have an impact on ridership.

To model the impacts of these service characteristics on ridership, a second model was developed using data obtained through a survey of rural demand-response transit agencies. Some of these data were obtained through a previous study conducted by Godavarthy et al. (2015). This model used data from 68 rural transit agencies, and it estimated ridership as a function of total service area population; percentages of service area population receiving service 6 or more days per week, 5 days per week, 12 or more hours per day, or fewer than 5 hours per day; advance reservation time; and fare levels.

Results

Similar to previous research, the first model showed that demographic characteristics are important. Ridership was found to significantly increase when the percentage of the population comprised of older adults or people without access to a vehicle increased. This model also showed that demand-response ridership is lower if the agency also operates a fixed-route service or if its service area overlaps that of another transit agency. Agencies serving only a municipality were found to have greater ridership, indicating that systems serving smaller geographic areas with more concentrated demand and shorter trip lengths will have higher levels of ridership, holding population and all other variables constant.

Both models showed a negative effect of fares on ridership. The first model estimated a price elasticity of -0.24, and the second estimated an elasticity of -0.12. Given that the first model used data from a much larger number of transit agencies, its results may be more reliable.

The second model demonstrated the importance of service characteristics, as agencies providing more days of service were found to have higher levels of ridership. This ridership increase could be due to there simply being more hours of service, but there could also be increases in trips per service hour if riders find the higher level of service more reliable and useful.

This model also showed that advance reservation time is important. Agencies that allowed users to reserve rides on shorter notice had higher levels of ridership, and the magnitude of the effect was significant. This could have important implications as agencies attempt to improve service qualities to serve more riders. The results suggest that transportation network companies could be successful in rural areas by providing rides with a short response time. More detailed results are shown below.

Population

A 1% increase in population was found to increase ridership by 0.69-0.83%.

Demographics

The first model showed that areas with a higher percentage of older adults or households without access to a vehicle have higher levels of ridership. If the percentage of the population that is aged 65 or older increases by one percentage point, ridership was found to increase by 8%. If the percentage of the population that lives in a household without a vehicle increases by one percentage point, ridership was found to increase by 21%. The percentage of population with a disability, however, was not found to have a statistically significant impact.

Fares

Fares have a negative impact on ridership. A 1% increase in fares was found to reduce ridership by 0.12-0.24%. The low fare elasticities show that ridership is fairly inelastic to price, as would be expected given that many users of demand-response service have few alternative options. Although it is inelastic, fare levels still have some effect, showing that even though many riders may be transit-dependent, ridership will decrease with increases in fares.

Span of Service

The second model showed that ridership is impacted by the number of days that service is available. As the percentage of service area population with service 5 days per week increases by one percentage point, ridership was found to increase 1.41%; and ridership was found to increase 1.65% as the percentage of service area population with service 6 or 7 days per week increases by one percentage point. The effect of service hours per day was not found to be statistically significant, but that could be because of the limited number of observations.

Advance Reservation Requirement

The second model showed that advance reservation time has a negative impact on ridership. Compared to agencies that require reservation two or more days in advance, ridership is 124% higher for providers that require reservation one day in advance and 201% higher for agencies that allow same-day service.

Other Results

- Agencies that provide both fixed-route and demand-response service have lower levels of demand-response ridership than agencies that provide just demand-response service.
- Agencies that serve areas where more than one transit provider is available have lower levels of ridership.
- Demand-response providers that strictly serve a municipality have higher levels of ridership than those serving a larger geographic area, after accounting for population and other factors.
- There are some regional differences in ridership not accounted for by these variables. Notably, region 5 agencies have higher levels of ridership, and agencies in regions 3 and 4 have lower levels.

It should be noted that predicting ridership for an individual agency given the many community-specific or agency-specific factors that are not accounted for in the data is difficult. To evaluate the accuracy of the first model, which was estimated with 2013 rural NTD data, it was used to predict ridership for 2014, and predicted ridership was then compared to actual ridership. Results showed that there is still significant variation in ridership that is not explained by the model, but the model developed in this study performed better than the simpler model described in TCRP 161.

The report provides instructions on how individual transit agencies or transportation planners can use results from the two models. Guides are given for how to obtain the necessary data and use the formulas, which are shown below.

Model #1 states that ridership can be estimated as follows:

Natural log of ridership =

> 0.83 × natural log of population
>
> + 7.99 × percentage of population aged 65 or older
>
> + 21.15 × percentage of population without access to a vehicle
>
> - 0.65 if the agency also operates a fixed-route service
>
> - 0.41 × percentage of population that has access to other demand-response service
>
> + 0.77 if the agency operates strictly within a municipality
>
> - 0.24 × natural log of the fare
>
> - 0.56 if agency operates in FTA region 3
>
> - 0.81 if agency operates in FTA region 4
>
> + 0.50 if agency operates in FTA region 5

Model #2 states that ridership can be estimated as follows:

Natural log of ridership =

> 0.69 × natural log of population
>
> + 1.65 × percentage of population with service 6 or 7 days per week
>
> + 1.41 × percentage of population with service 5 days per week
>
> + 2.01 if agency allows same-day reservations
>
> + 1.24 if agency requires reservations made one day in advance
>
> - 0.65 if agency operates fixed-route service
>
> - 0.12 × natural log of fare

Conclusions

While previous studies have shown that the populations of older adults, people with disabilities, and people without access to a vehicle significantly impact the demand for demand-response transit services in rural areas, this study showed that a number of other factors are also significant determinants. By conducting an out-of-sample validation, the first of this study's models was found to perform better than a model estimated in a previous study.

This study provides two new tools for estimating ridership, which can be used in conjunction with existing models or peer analysis. The inclusion of a greater number of variables and more specific service information improves the performance of the models. These tools can be used by transit agencies or transportation planners to forecast demand for new demand-response services; estimate the impact of service changes, such as changes in geographic coverage, span of service, fares, reservation requirements; and project future ridership based on projected population and demographic changes.

LIST OF FIGURES

LIST OF TABLES

TABLE OF CONTENTS

1. INTRODUCTION

Few studies have provided a method for estimating ridership for demand-response transit services. Lack of data for demand-response service characteristics and geographic coverage has limited the development of such models. TCRP Report 161 (Vanasse Hangen Brustlin, Inc, et al. 2013) provides methods for forecasting demand for rural passenger transportation, including general demand-response services. The demand model developed in TCRP Report 161 includes some service characteristics, such as size of service area and service miles, but it lacks other service characteristics. Factors such as fares, span of service, reservation requirements, and other service characteristics will likely impact ridership. These data are not included in the rural National Transit Database (NTD), but data collected recently by Godavarthy et al. (2015a, 2015b) from demand-response transit providers in North Dakota and Florida provide detail about these service characteristics as well as geographic coverage. These data, when combined with rural NTD data and population and demographic data from the U.S. Census and the American Community Survey (ACS), allow for the development of a more detailed demand model. Results can be used to forecast demand for new demand-response services; to estimate the impact of service changes, such as changes in geographic coverage, span of service, fares, etc.; and to project future ridership based on projected population and demographic changes.

The objective of this study is to develop a model for estimating demand for rural demand-response transit services for the general public. Specific objectives are to estimate the impacts of service characteristics (such as span of service, service coverage, fares, and reservation requirements) and service-area characteristics (such as population and demographic characteristics) on ridership.

Section 2 of this report provides a review of previous demand models for rural demand-response transit, including TCRP Report 161, other TCRP reports, and other related reports. The third section provides a description of the models developed in this study and the data used. Two models were developed for estimating ridership for rural demand-response transit agencies. The first used data from the rural NTD and the ACS. Because data limitations of the rural NTD restrict the number of variables that can be analyzed, a second model was developed using data collected from surveys of rural transit agencies. This model used data collected from Godavarthy et al. (2015b) along with service data collected from additional transit agencies through a new survey. Both models estimate demand for non-sponsored demand-response service for the general public. Demand for sponsored, or program-related, trips or ADA paratransit was not considered.

The estimated results from the models are presented in Section 4. An out-of-sample validation was also conducted, which compared the performances of the first model and the TCRP 161 model. Instructions for how to use the estimated results are presented in Section 5, and conclusions are discussed in Section 6.

2. PREVIOUS DEMAND MODELS

TCRP Report 161 provides methods for forecasting demand and quantifying need for rural passenger transportation. The report categorizes rural transportation services into four markets: general public rural passenger transportation, passenger transportation specifically related to social services or other programs, travel on fixed-routes in micropolitan areas, and travel on commuter services from rural counties to urban centers. The current study is focused on the first of these markets – transit services for the general public.

The TCRP report provides two methods for estimating demand for rural general public, non-program-related, services. The first of these methods is to rely on peer data. The report recommends calculating passenger trips per capita, passenger trips per vehicle mile, and passenger trips per vehicle hour for other systems operating in similar rural settings within the state or adjacent states. Calculating the mean, median, and ranges for each of these measures will provide guidance for estimating ridership for new services.

The second method provided in TCRP 161 is a demand function the authors developed based on data from the 2009 rural NTD. This model is based on the assumption that older adults, people with mobility limitations, and people without access to a vehicle represent the main users of these services, and that demand can be estimated based on the sizes of these demographic groups. More specifically, the model estimates demand as a function of the population aged 60 or older, the mobility-limited population aged 18-64, and the number of residents from households having no vehicle. The full model, with estimated coefficients, is shown below:

Non-program Demand (trips per year) = (2.20 x Population Age 60+) + (5.21 x Mobility Limited Population Age 18-64) + (1.52 x Residents of Household Having No Vehicle)

This method can be used to estimate demand for service in an area by multiplying the number of people in each demographic group by the proper factor. Demographic data for individual communities are available from the American Community Survey (ACS).

As noted, this model applies to general public, non-program trips. Program trips, on the other hand, are those trips that would not be made without the existence of a specific social-service program or activity. To estimate demand for program, or sponsored, trips, TCRP 161 recommends obtaining the following data from the agencies providing these services:
- Number of program participants
- Number of days per week that the program meets
- The number of weeks per year the program is offered
- The proportion of program participants who attend the program on an average day
- The proportion of program participants who require transportation service

Using these data, demand for program trips can be estimated as follows (TCRP 161):

Number of Program Participants × Program Events per Week × the Proportion of Program Participants who attend the Program on an Average Day × the Proportion of Program Participants that are Transit Dependent or Likely to Use the Transit Service provided/funded by the Agency × the Number of Weeks per Year the Program is Offered × 2 (trips per participant per event)

As noted, TCRP 161 also developed models for small city fixed-route service (for cities with a population less than 50,000) and commuter transit from a rural area to an urban center. Demand for these services is

expected to differ significantly from demand-response service and is impacted by different factors. TCRP 161 estimated demand for small city fixed-route service as a function of revenue hours of service provided, population of service area, and college/university enrollment. Lastly, the proportion of workers using transit for commuter trips was estimated as a function of the number of workers commuting, the commute distance, and whether or not the urban place is a state capital.

Other research has been conducted regarding demand for paratransit services. Goodwill and Joslin (2013) assessed methodologies for paratransit service demand and developed a new tool for forecasting demand for transportation-disadvantaged services. Findings from this study can be useful for analyzing ADA complementary paratransit services and other specialized service markets. This method used trip rates from the 2009 National Household Travel Survey for households with zero vehicles for forecasting paratransit demand. This study defined the general transportation-disadvantaged (TD) population to include people with disabilities, older adults, low-income persons, and children defined as "high-risk" or "at-risk," while the critical need TD population was defined to include individuals who are unable to transport themselves or purchase transportation.

TCRP Reports 119 (Koffman et al. 2007) and 158 (Bradley and Koffman 2012) also developed demand estimation techniques for ADA paratransit. TCRP Report 119 provides a tool for estimating total ADA paratransit demand based on the following variables:
- ADA paratransit service area population
- Base fare for ADA paratransit
- Percentage of applicants for ADA paratransit eligibility found conditionally eligible
- Whether or not trip-by-trip eligibility determination based on conditions of eligibility is used
- Percentage of service area population with household incomes below the poverty line
- The effective window used to determine on-time performance

The model developed in TCRP Report 119 was based on an analysis of data for 28 transit systems believed to offer high-quality ADA paratransit service. This report was followed by TCRP Report 158, which provides more detailed modeling, including a sketch planning model and a regional planning model. The regional planning model includes the effects of changes in demographic and travel variables on ADA paratransit registration, and it can be used in conjunction with an existing regional travel demand model to forecast travel by ADA paratransit-eligibility persons, providing details about the number of trips by mode, trip purpose, and destination. The report analyzed the following factors as potentially impacting the number of people registered for ADA paratransit or the number of ADA paratransit trips per registered person (or both): aging of the population, poverty rates, household size, sensitivity to fares, sensitivity to travel time, on-time pick-ups and drop-offs, telephone hold times, and pedestrian access to activities.

The models developed in these studies are specific to certain markets. The ADA paratransit models are useful for estimating demand for complementary paratransit services in areas with fixed-route transit, but are not applicable to general public demand-response service in rural areas. TCRP Report 161 addressed different types of service that may be offered in rural areas or small cities, including general public demand-response, program-specific demand-response, small-city fixed-route, and commuter services. This study is specific to just one of these markets, the market for non-program demand-response service for the general public in rural areas. Therefore, the non-program demand model from TCRP Report 161 is the most pertinent model. In this model, demand is estimated as a function of the sizes of population groups most likely to use the service. However, there may be other factors that could influence ridership, such as the characteristics of the service being provided. The next section presents a discussion of the factors that could impact ridership and a description of the models used to estimate ridership.

3.   METHODS

3.1   Factors Affecting Ridership

It is expected that ridership will be determined by the demand for services, the level of service provided, and the cost of service. Demand for services can be estimated based on population and demographic characteristics. It is expected that demand will be greater in areas with a higher population and that concentrations of transit-dependent or transportation-disadvantaged populations will create greater demand. As noted in previous research, it is expected that demand for rural transit will be greater in areas with a larger population of older adults, people with disabilities, and people without access to a personal vehicle.

While population characteristics can be used to estimate demand, ridership will also be influenced by the characteristics of the service provided and the fare levels. Important service characteristics could include the number of days per week and hours per day that service is provided, as well as the advance reservation requirements. Span of service refers to the number of days and hours that service is available. Transit agencies with a greater span of service are expected to attract more riders, everything else held equal. It is also hypothesized that reservation requirements serve as a deterrent and that requiring reservations further in advance will negatively impact ridership.

Other service characteristics could also affect ridership. Some rural transit agencies operate both fixed-route and demand-response service. In those cases, some of the transit demand in the areas is being met by fixed-route services, so demand-response ridership is expected to be lower than it would have been had the fixed-route services not been available. In other cases there may be overlap in service areas, with more than one provider serving an area. Ridership is expected to be lower for demand-response agencies that are sharing service area with other agencies. It may also be expected that ridership for tribal transit may differ from that of other rural transit systems because of the unique geographic and economic characteristics of Native American reservations.

Lastly, ridership is expected to be impacted by the level of fares charged. It is hypothesized that ridership will be lower for agencies that charge higher fares, everything else held equal.

3.2   Data and Models

Based on these factors, two models were developed to estimate general public demand-response ridership. The first model was developed using data from the 2013 rural NTD. Data limitations of the rural NTD restrict the number of variables that can be analyzed. To address these data limitations, the second model was developed using data collected from surveys of rural transit agencies.

3.2.1 Model #1 using National Transit Database Data

In the first model, ridership is a function of service area population; demographic characteristics of the service area; whether the transit agency provides a fixed-route service, has a service area that overlaps with that of another demand-response transit provider, operates only in a municipality, or is a tribal agency; the fare level; and the region of the country in which it operates. Population and demographic characteristics are expected to influence demand, fares represent the cost of the service, and the other variables represent characteristics of the service or service area that may impact ridership. The model is estimated as follows:

$$\ln R_i = \alpha_1 \ln POP_i + \alpha_2 PELD_i + \alpha_3 PNV_i + \alpha_4 PDIS_i + \alpha_5 FIXRT_i + \alpha_6 POVR_i + \alpha_7 MUNI_i + \alpha_8 TRIBE_i +$$
$$\alpha_9 \ln FARE_i + \alpha_{10} R1_i + \alpha_{11} R2_i + \alpha_{12} R3_i + \alpha_{13} R4_i + \alpha_{14} R5_i + \alpha_{15} R6_i + \alpha_{16} R7_i + \alpha_{17} R8_i + \alpha_{18} R9_i + \varepsilon_i$$

where R_i = demand-response ridership for transit agency i

POP_i = service area population for agency i

$PELD_i$ = percentage of service area population for agency i that is elderly, defined as aged 65 or older

PNV_i = percentage of service area households for agency i without access to a vehicle

$PDIS_i$ = percentage of service area population for agency i with a disability

$FIXRT_i$ = a dummy variable equal to 1 if agency i also operates fixed-route service, 0 otherwise

$POVR_i$ = percentage of service area population for agency i that also has access to another demand-response provider

$MUNI_i$ = a dummy variable equal to 1 if agency i strictly serves a municipality, 0 if agency serves larger geographic area

$TRIBE_i$ = a dummy variable equal to 1 if agency i is a tribal transit operator, 0 otherwise

$FARE_i$ = fare level for agency i

$R1_i \ldots R9_i$ = dummy variables for FTA regions 1 through 9

Data from the American Community Survey (ACS) 2009-2013 5-year estimates, and the 2013 rural NTD for agencies operating demand-response service were used to estimate the model. Ridership was defined as regular unlinked passenger trips. Sponsored trips were not included. Outlier agencies that provided fewer than 100 trips were excluded. The dependent variable in the model is the natural logarithm of ridership.

Service areas were determined based on responses from the rural NTD, which lists the counties served by each agency. Unfortunately, some larger, multi-county agencies do not have all of their counties listed in the NTD. Due to incomplete service area information, these agencies could not be included in the analysis. The analysis assumes that agencies provide an equal level of service throughout the counties listed, though that may not always be the case.

Service area population is the total population of the counties served by an agency. While rural agencies serve cities with population less than 50,000, they may serve multiple cities or counties with a combined population totaling more than 50,000. Because of imprecise service area information, the total county population had to be used to estimate service area population, which would overestimate service area population if the agency did not serve the entire county. To minimize this measurement problem, agencies with an estimated service area population of more than 100,000 were excluded from the analysis, because these agencies most likely do not serve the entire population. Some agencies are listed in the NTD as serving a municipality, rather than an entire county or multi-county region. For these agencies, population was based strictly on the population of the city served.

Population aged 65 or older, population with a disability, and population in households without access to a vehicle were determined for each agency's service area. The populations of these transportation-disadvantaged groups were converted to percentages of total population. County-level and city-level population data from the 2009-2013 ACS 5-year estimates were used. It is hypothesized that ridership will increase with increases in population and that areas that have higher percentages of older adults, people with disabilities, and households without a vehicle will have greater ridership.

Figure 3.1 is a scatter plot of rural demand-response transit ridership in 2013 and population, excluding a few outliers and those with estimated service area population of more than 100,000. The figure shows little correlation between population and ridership (the correlation coefficient is 0.23), suggesting that a number of other factors play a role in determining ridership, though the lack of correlation could also be partly explained by the imprecise estimates of service area population.
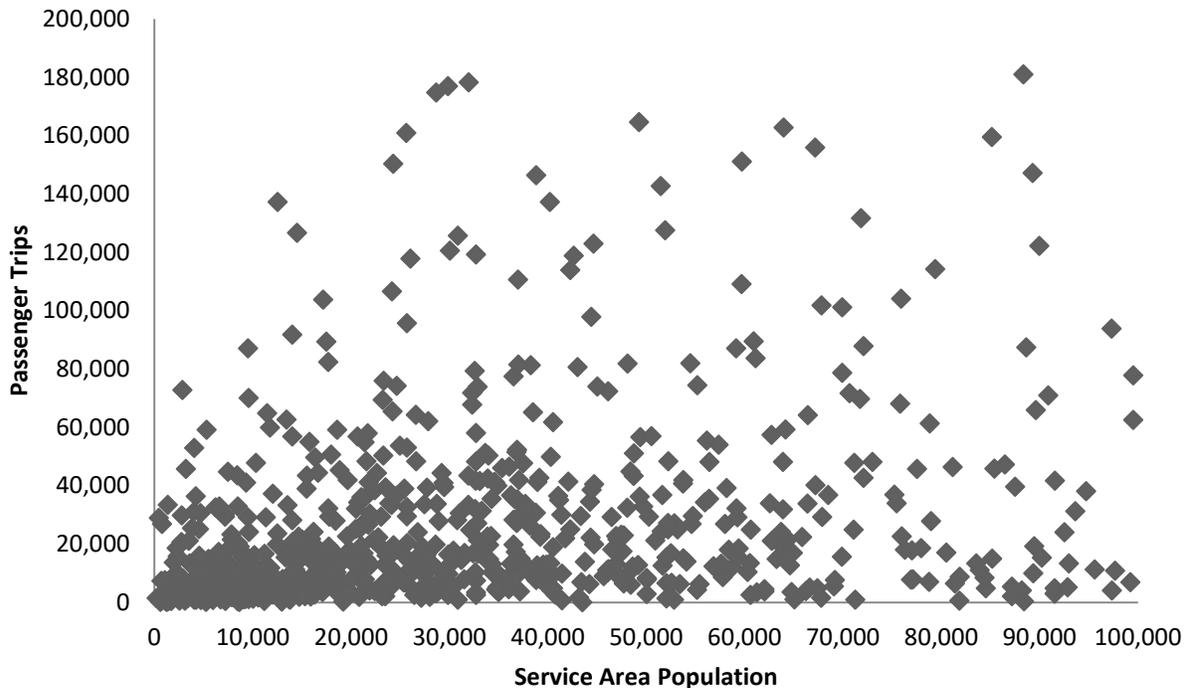

Figure 3.1 Population and Demand-Response Transit Ridership

Some rural agencies provide both fixed-route and demand-response services. If an agency provides both services, then potential demand-response riders could be served, instead, by the fixed-route service. It is expected that agencies that operate both demand-response and fixed-route service will have lower demand-response ridership than agencies that operate only demand-response transit, everything else equal. To account for this factor, a dummy variable was included in the model to represent agencies providing fixed-route service.

Similarly, ridership is expected to be lower if there is another demand-response provider in the service area. Service area data were analyzed to identify instances when more than one agency was listed as serving a particular county. Then, for each agency, the study calculated the percentage of service area population living in a county with access to another demand-response transit provider. It is hypothesized that ridership will decrease if greater percentages of the population have access to other transit services.

Other geographic characteristics can also have an impact on ridership. Due to the inefficiencies of providing service in rural areas, transit agencies in rural, low-density areas with long travel distances may not be able to provide as many trips as agencies serving more concentrated and densely population areas. Agencies that strictly serve a municipality, on the other hand, may be able to provide more trips per capita than those serving larger geographic areas because of shorter trip distances and greater population density. To account for this effect, a dummy variable was included in the model to represent agencies serving only a municipality. This variable equals 1 if the agency serves only a municipality and 0 otherwise.

Ridership may be lower on Native American reservations because of long travel distances and low densities, poor economic conditions, and cultural differences in many rural tribal areas. A dummy variable was included in the model, equal to 1 if the agency is a tribal transit provider and 0 otherwise, to account for this possible effect. Nine dummy variables for FTA regions were also included to account for possible geographic or other differences between regions that are not accounted for by other variables but may influence ridership. The dummy variable was excluded for region 10, as region 10 serves as the base for comparison. A map of FTA regions is shown in Figure 3.2.
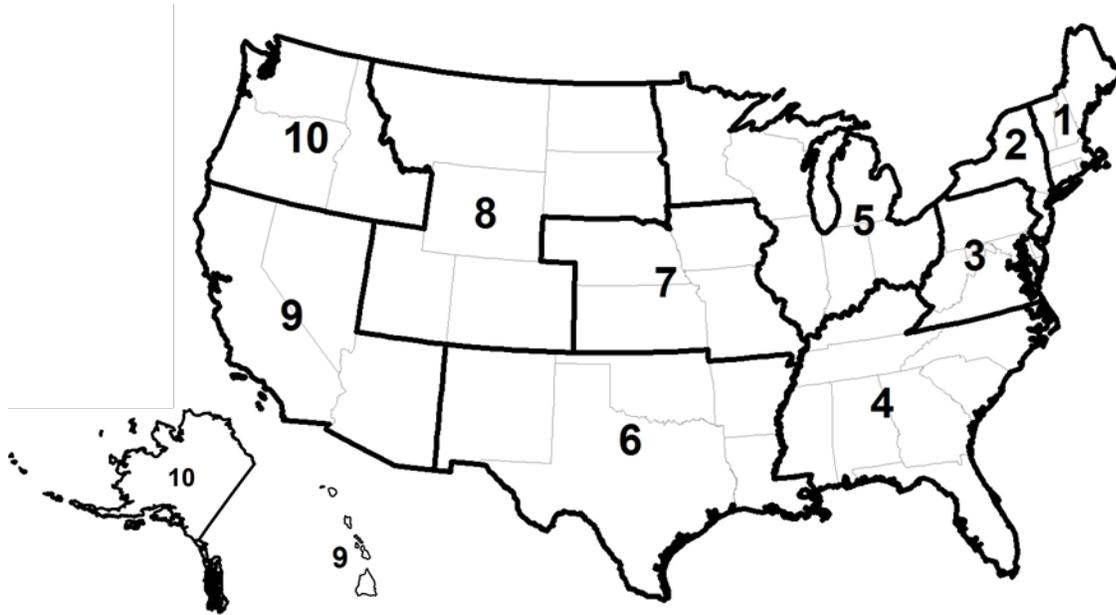
Figure 3.2 FTA Regions

Last, fare levels are expected to have a negative impact on ridership. Average fares charged by each agency were calculated by dividing total fare revenues by the number of trips provided.

Table 3.1 provides summary statistics for the variables used in the model. There were 731 agencies in the 2013 rural NTD for which all of the necessary data were available. Note that many of the larger, multi-county agencies were excluded because of a lack of information about the counties they serve. A few outliers with ridership over 200,000 were also excluded. Table 3.1 describes the data for the 731 agencies used to estimate the model.

Table 3.1 Summary Statistics for Model #1 Variables

	Mean	Median	Standard Deviation	Minimum	Maximum
Passenger trips	26,344	14,976	31,758	180	180,983
Population	30,448	24,609	24,619	204	100,000
Percentage elderly	0.17	0.16	0.04	0.06	0.30
Percentage no vehicle	0.01	0.01	0.01	0	0.11
Percentage disability	0.15	0.14	0.04	0.04	0.32
Fixed route	0.20	0	0.40	0	1
Percentage overlap	0.23	0	0.41	0	1
Municipality	0.16	0	0.37	0	1
Tribal	0.01	0	0.12	0	1
Fare	1.24	0.90	2.41	0	25.89

The data show significant variations between agencies in ridership and service area population. The average demand-response ridership, measured as regular unlinked passenger trips, was 26,344, with a median of 14,976, a standard deviation of 31,758, and a maximum number of 180,983. The average service area population was 30,448, with a standard deviation of 24,619, while the median population was 24,609.

On average, 17% of the service area population is aged 65 or older, 1% lives in a household without access to a vehicle, and 15% has a disability. By comparison, about 13.7% of the U.S. population is aged 65 or older and 12.3% has a disability (Mattson 2015), indicating somewhat higher concentrations of older adults and people with disabilities in the rural areas served by these agencies. Access to vehicles tends to be higher in these areas, though, as a greater percentage of urban residents live in households without a vehicle (Mattson 2015). As shown in Table 3.1, there is some variation in the demographic characteristics of the areas served by these agencies. The percentages of the population consisting of older adults, people in households with no vehicle, and people with a disability range from 6% to 30%, 0% to 11%, and 4% to 32%, respectively.

About 20% of the agencies operate fixed-route service in addition to demand-response service; 16% serve a municipality; and 1% are tribal transit operators. The average fare revenue collected per trip provided was $1.24. The rural NTD does not provide information on fare revenues by mode, so fares had to be estimated by dividing total fare revenues for all modes by total trips for all modes. Since demand-response is the primary mode for most agencies, the estimates are still useful.

3.2.2 Model #2 using Data Collected from Agencies

Missing from previous models are the impacts of service qualities on ridership. While the population of the service area and the demographic characteristics are important determinants of ridership, the number of days per week and hours per day service is available as well as how convenient it is for users to reserve a ride are also important. Previous models have not included these factors because of a lack of data in the NTD or elsewhere.

As noted in the Transit Capacity and Quality of Service Manual (Kittelson and Associates et al. 2013), key measures of demand-response transit quality of service include geographic service coverage, service span (days and hours of service), and response time, which refers to the amount of time users must

reserve a ride in advance. These three variables are measures of service availability and are expected to have an impact on ridership.

With regard to geographic coverage data, in the rural NTD, agencies are asked to identify the county or counties in which they provide service. Some of the multi-county providers do not list all of their counties, so the data are incomplete. The data are also imprecise because an agency might not provide service within some areas of the county, or the level of service provided may differ within the county. Some areas might receive daily service, while others have service twice a week. Service span data are not available in the rural NTD. More detailed service coverage data and information about service span would be useful for modeling ridership. Further, information on advance reservation time could also be useful, as requirements for reserving rides further in advance could negatively impact ridership.

To address these data limitations, Godavarthy et al. (2015a, 2015b) developed online survey tools for collecting more detailed service information from demand-response transit agencies. They used these survey instruments to collect detailed information regarding service area, service span, service type, eligibility requirements, and advance reservation requirements from demand-response transit agencies in North Dakota and Florida. If transit operators provided different levels of service within their service areas (for example, 7 days a week in one area and 5 days a week in another), the survey collected that information.

The information collected from these surveys could be useful for developing a demand-response transit ridership model that estimates ridership as a function of service area population, span of service, advance reservation time, and other factors. While Godavarthy et al. (2015a) collected data for agencies in North Dakota and Florida, data from additional agencies across the country would be needed for the model. Therefore, a similar but simpler survey was sent to rural transit agencies across the country.

This survey asked transit agencies to first identify the county or counties in which they provide service. Then they were asked if there are some areas within the county/counties listed in which their agency does not provide service, if some areas receive more days or hours of service than other areas, or if all areas within the county/counties receive the same level of service. Respondents were then asked how many days per week and hours per day their agency provides service. If different levels of service are provided, or if some areas are not served, respondents were asked to identify specific communities or areas and identify span of service for each.

Lastly, the survey collected information about service eligibility and minimum advance reservation time. Service eligibility options included the general public, people with disabilities, senior citizens, and other. Because this study focuses on general public service, the intent of this question was to ensure that respondents provide demand-response service to the general public. Those providing ADA paratransit or limited eligibility service would be excluded from the analysis. Regarding minimum advance reservation time, the survey options are shown in Table 3.2.

Table 3.2 Minimum Advance Reservation Time
Levels of Service

Options
Up to ½ hour
More than ½ hour and up to 2 hours
More than 2 hours, but still same day
24 hours in advance, or prior service day
48 hours, or 2 days, in advance
More than 48 hours in advance, and up to 1 week
More than 1 week in advance, and up to 2 weeks
More than 2 weeks

The model estimates ridership as a function of total service area population; percentages of service area population receiving service 6 or more days per week, 5 days per week, 12 or more hours per day, or fewer than 5 hours per day; advance reservation time; and fare levels. Population is an indicator of overall demand, fare represents cost, and other variables are measures of level of service. Similar to the previous model, a dummy variable is included to indicate if the agency also operates a fixed-route service. The model is estimated as follows:

$$\ln R_i = \beta_1 \ln POP_i + \beta_2 PctDays6_i + \beta_3 PctDays5_i + \beta_4 PctHrs12_i + \beta_5 PctHrs5_i + \beta_6 RES1_i + \beta_7 RES2_i + \beta_8 FIXRT_i + \beta_9 \ln FARE_i + \varepsilon_i$$

where $PctDays6_i$ = percentage of service area population for agency i that receives service 6 or 7 days per week

$PctDays5_i$ = percentage of service area population for agency i that receives service 5 days per week

$PctHrs12_i$ = percentage of service area population for agency i that receives service 12 or more hours per service day

$PctHrs5_i$ = percentage of service area population for agency i that receives service less than 5 hours per service day

$RES1_i$ = a dummy variable equal to 1 if agency i allows users to reserve same-day rides

$RES1_i$ = a dummy variable equal to 1 if agency i requires users to reserve rides at least one-day in advance

R_i, POP_i, $FIXRT_i$, and $FARE_i$ were previously defined. Similar to the previous model, it is expected that service area population will have a positive effect on ridership, fare levels will have a negative impact, and providing fixed-route service will have a negative effect on demand-response ridership.

The model considers three levels of service days per week – 6 or 7 days, 5 days, less than 5 days – and three levels of service hours – 12 or more hours, 5 to 11.9 hours, and less than 5 hours. It is expected that as the percentage of service area with 5 days of service increases, relative to areas with less than 5 days, then ridership will increase. Furthermore, increases in service area with 6 or 7 days of service should also have positive impacts on ridership, and it is expected that this effect should be greater. Agencies that provide 6 or 7 days of service throughout their service area are expected to have the highest ridership, everything else held constant. Similarly, agencies that provide 12 or more hours of service to a greater percentage of their service area are expected to have higher ridership, because they are able to serve a

greater variety of trip purposes, and those providing fewer than 5 hours of service per day are expected to have lower ridership.

Three levels of minimum advance reservation time are considered in the model: same day, 24 hours in advance (or prior service day), and two or more days in advance. Dummy variables are included for the first of these two, while a reservation time of two or more days in advance is used as the reference. It is hypothesized that reducing the amount of time that rides must be reserved in advance will have a positive effect on ridership, so both dummy variables are expected to be positive, and the effect of same-day reservations should be greater.

The model could be revised to include more detailed levels of service regarding days per week, hours per day, or reservation requirements. For example, more dummy variables could be included to account for the more specific advance reservation times shown in Table 3.2. However, this study limited the number of levels modeled because of the limited number of agencies from which data were available. A larger dataset would allow for more detailed analysis.

Between the data collected by Godavarthy et al. (2015a) and the data collected in this study through the survey of rural transit agencies, usable data were available for 68 rural transit agencies that provide demand-response, non-sponsored trips to the general public. Data were available for agencies from 24 different states representing different parts of the county, with the largest number of responses from North Dakota. This includes 20 agencies from North Dakota, six from Florida, six from Nebraska, four from Minnesota, and the remaining 32 dispersed across the country. Summary statistics for the 68 transit agencies are shown in Table 3.3

Table 3.3 Summary Statistics for Model #2 Variables

	Mean	Median	Standard Deviation	Minimum	Maximum
Passenger trips	31,103	19,490	45,351	63	343,937
Population	41,302	24,666	48,245	1,119	177,453
Percentage with 6 or 7 days	0.29	0	0.43	0	1.00
Percentage with 5 days	0.62	1.00	0.45	0	1.00
Percentage with 12 or more hours	0.34	0	0.45	0	1.00
Percentage with less than 5 hours	0.03	0	0.14	0	1.00
Same-day reservation	0.43	0	0.50	0	1.00
Prior-day reservation	0.49	0	0.50	0	1.00
Fixed-route	0.25	0	0.47	0	1.00
Fare	1.54	1.19	1.27	0	6.46

4. RESULTS

4.1 Estimated Results from Model #1

Estimated results from model #1 are presented in Table 4.1. The model estimated results using data from 731 rural demand-response transit agencies. Results show the expected positive impacts of total population, older adult population, and population without a vehicle. The percentage of population with a disability, however, was not found to have a statistically significant impact. Most other variables were statistically significant with the expected signs, with the exception of the tribal transit variable, which was insignificant. Because the dependent variable in the model is the natural log of ridership, estimated coefficients can be interpreted as percentage impacts on ridership. The results can be interpreted as follows:

- Population has a positive effect on ridership. A 1% increase in population leads to a 0.83% increase in ridership.
- Demographics impact ridership. Areas with a higher percentage of older adults or households without access to a vehicle have higher levels of ridership. If the percentage of the population that is aged 65 or older increases by one percentage point, ridership increases by 8%. If the percentage of the population that lives in a household without a vehicle increases by one percentage point, ridership increases by 21%.
- Agencies that provide both fixed-route and demand-response service have lower levels of demand-response ridership than agencies that provide just demand-response service, after accounting for all other variables.
- Agencies that serve areas where more than one transit provider is available have lower levels of ridership.
- Demand-response providers that strictly serve a municipality have higher levels of ridership than those serving a larger geographic area, after accounting for population and other factors.
- Fares have a negative impact on ridership. A 1% increase in fares leads to a 0.24% reduction in ridership.
- There are some regional differences in ridership not accounted for by these variables. Notably, region 5 agencies have higher levels of ridership, and agencies in regions 3 and 4 have lower levels.

The results notably show that while total population is an important predictor of ridership, areas with higher concentrations of older adults and people without access to a vehicle have greater ridership. These groups are among the primary users of demand-response transit service. The fare elasticity of -0.24 shows that ridership is fairly inelastic to price, as would be expected given that many users of demand-response service are transit-dependent. Although it is inelastic, fare levels still have some effect, showing that even though many riders may be transit-dependent, ridership will decrease with increases in fares.

The positive effect of the municipality variable confirms the hypothesis that agencies serving smaller geographic areas with more concentrated demand and shorter trip lengths will have higher levels of ridership, holding population and all other variables constant.

Table 4.1 Results from Model #1

Independent Variable	Estimated coefficient	Standard error	p-value
Ln(Population)	0.83	0.02	0.000
Percentage elderly	7.99	0.99	0.000
Percentage with no vehicle	21.15	5.65	0.000
Percentage with disability	-0.46	1.20	0.703
Fixed-route	-0.65	0.11	0.000
Percentage overlap	-0.41	0.10	0.000
Municipality	0.77	0.10	0.000
Tribal	0.30	0.31	0.333
Ln(Fare)	-0.24	0.04	0.000
Region 1	-0.60	0.33	0.071
Region 2	-0.57	0.42	0.170
Region 3	-0.56	0.25	0.027
Region 4	-0.81	0.19	0.000
Region 5	0.50	0.20	0.012
Region 6	-0.15	0.22	0.480
Region 7	-0.36	0.19	0.057
Region 8	0.09	0.19	0.628
Region 9	0.16	0.25	0.523
n = 731	0.83	0.02	0.000

4.2 Estimated Results from Model #2

Estimated results from model #2 are shown in Table 4.2. This model shows a 1% increase in population increasing ridership by 0.69%. Like the previous model, this model also shows negative effects of fares and operating fixed-route service on demand-response ridership. The magnitude of the fare effect is smaller, though, with a fare elasticity of -0.12, compared to -0.24 in the previous model. The fixed-route effect is similar to the previous model.

The key difference with the second model is its ability to show the effect of service characteristics on ridership. As expected, the model shows that providing 6 or 7 days of service to a larger percentage of the service has a positive effect on ridership. The results also show that providing 5 days of service significantly increases ridership compared to having less than 5 days, but providing 6 or 7 days of service increases ridership further.

While the results show a positive effect of providing 12 or more hours of service and a negative effect of providing less than 5 hours, these results are not statistically significant. It is expected that increasing service hours should have a positive effect on ridership, but these results were not able to demonstrate that effect, perhaps because of the small sample size.

Results show that making it more convenient to reserve a ride has a positive impact on ridership. Agencies providing same-day service had the highest level of ridership, everything else equal. In fact, agencies allowing same-day service were found to provide about 200% more rides than those requiring reservations two or more days in advance. Agencies requiring reservations one day in advance had higher ridership than those requiring reservations further in advance, but they had lower ridership than those providing same-day service.

Table 4.2 Results from Model #2

Independent Variable	Estimated coefficient	Standard error	p-value
Ln(Population)	0.69	0.07	<.0001
Percentage population with 6 or 7 days	1.65	0.80	0.0439
Percentage population with 5 days	1.41	0.69	0.046
Percentage population with 12 or more hours	0.50	0.43	0.2545
Percentage population with less than 5 hours	-0.40	1.20	0.7397
Same-day reservation	2.01	0.55	0.0006
Prior-day reservation	1.24	0.56	0.0321
Fixed-route	-0.65	0.39	0.1013
Ln(Fare)	-0.12	0.07	0.0843
n = 68			

The results can be interpreted as follows:
- Population has a positive effect on ridership. A 1% increase in population leads to a 0.69% increase in ridership.
- Ridership is impacted by the number of days that service is available. As the percentage of service area population with service 5 days per week increases by one percentage point, ridership increases 1.41%; and ridership increases 1.65% as the percentage of service area population with service 6 or 7 days per week increases by one percentage point.
- Advance reservation time has a negative impact on ridership. Compared to agencies that require reservation two or more days in advance, ridership is 124% higher for providers that require reservation one day in advance and 201% higher for agencies that allow same-day service.
- Agencies that provide both fixed-route and demand-response service have lower levels of demand-response ridership than agencies that provide just demand-response service, after accounting for all other variables.
- Fares have a negative impact on ridership. A 1% increase in fares leads to a 0.12% reduction in ridership.

4.3 Out-of-Sample Model Validation

To test the forecasting performance of model #1 and compare it to the model previously developed in TCRP 161, an out-of-sample model validation was performed. To perform the validation, results from the model were used to predict ridership for 2014, and the predicted ridership for each agency was compared to actual ridership reported in the 2014 rural NTD data. Predicted ridership was estimated based on both model #1 estimates provided in Table 4.1 and the TCRP 161 model for non-program demand, as described in Section 2. Excluding agencies with a service area population greater than 100,000, data were available to predict non-program, demand-response ridership for 688 rural transit agencies using the two models. The Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) for the two models are shown in Table 4.3. RMSE and MAE were calculated based on the predictions for all agencies and then again based on the predictions for agencies with population less than 50,000.

Table 4.3 Out-of-Sample Validation

	Model #1	TCRP 161 Model
Population under 100,000 (n=688)		
RMSE	55,579	73,941
MAE	23,506	28,669
Population under 50,000 (n=544)		
RMSE	48,231	71,439
MAE	19,536	26,027

Attempts to predict 2014 ridership showed that there is still significant variation in ridership that is not explained by these models. While the relationships depicted in these models are useful, predicting ridership for an individual agency given the many community-specific or agency-specific factors that are not accounted for in the data is difficult. However, the results showed that the model developed in this study performed better than the simpler model described in TCRP 161. The Mean Absolute Error was 28,669 for the TCRP 161 model and 23,506 for model #1. RMSE and MAE for both models are shown to be smaller when predicting ridership for areas with a smaller population. While these estimates are an improvement over those from previous studies, more specific data are needed to reduce the error further.

By using more specific service data, it is expected that model #2 would perform better than model #1. However, model #1 was estimated using many more observations, which should make it more reliable. Because the service data needed for model #2 was available for only the 68 agencies included in its estimation, it was not possible to conduct an out-of-sample validation. Therefore, the performance of model #1 could not be compared to that of model #2. If the detailed service data needed for model #2 were available for a larger number of agencies, the model could be refined further and its performance could be evaluated.

5. APPLICATIONS

This section provides instructions on how individual transit agencies or transportation planners can use results from the two models. Results can be used to forecast demand for new demand-response services; estimate the impact of service changes, such as changes in geographic coverage, span of service, fares, reservation requirements; and project future ridership based on projected population and demographic changes.

5.1   Applying Model #1

5.1.1 Formula

Model #1 states that ridership can be estimated as follows:

Natural log of ridership =

> $0.83 \times$ natural log of population
>
> $+ 7.99 \times$ percentage of population aged 65 or older
>
> $+ 21.15 \times$ percentage of population without access to a vehicle
>
> $- 0.65$ if the agency also operates a fixed-route service
>
> $- 0.41 \times$ percentage of population that has access to other demand-response service
>
> $+ 0.77$ if the agency operates strictly within a municipality
>
> $- 0.24 \times$ natural log of the fare
>
> $- 0.56$ if agency operates in FTA region 3
>
> $- 0.81$ if agency operates in FTA region 4
>
> $+ 0.50$ if agency operates in FTA region 5

5.1.2 Data

Population and demographic data can be obtained online through the U.S. Census Bureau's American Fact Finder: www.factfinder.census.gov. The Fact Finder provides data from the decennial census, annual census population estimates, and the American Community Survey (ACS), with data provided for different geographic levels, including counties, census tracts or blocks, and cities. TCRP Report 161 provides detailed instructions on how to obtain data through the American Fact Finder.

The model uses demographic data from the ACS. The ACS provides one-year estimates, three-year estimates, and five-year estimates. The three- and five-year estimates, which are based on data collected over multiple years, have lower margins of error and are, therefore, more reliable, especially for smaller geographic areas. Table 5.1 shows the specific ACS tables used to obtain demographic data for model #1.

Table 5.1 American Community Survey Data Tables Used to Estimate Model #1

Variable	ACS Table
Total population	B01001
Population aged 65 or older	B01001
Population without access to a vehicle	B08014
Population with a disability	B18101

The data used in the model for the population without access to a vehicle, obtained from Table B08014, specifically refers to workers aged 16 or older living in a household with no vehicle available. This is not an exact estimate of the population without a vehicle, because it includes only workers, but it is a useful measure nonetheless. TCRP Report 161 calculated this variable differently, using data from Table B08201.

The remaining variables in the equation are specific to the transit agency, such as fares, the FTA region in which it is located, the availability of other demand-response service, and whether it operates fixed-route service, serves just a municipality, or is tribal transit.

5.1.3 Calculations

The formula uses the natural log of population and fares. A number can be converted to its natural log in Excel using the following formula: =LN(number). Inside the parentheses should be the number or cell reference for population or fares.

The populations of older adults and people without vehicles need to be divided by total population to obtain percentages for each demographic group. These numbers must be represented in the equation as decimal numbers ranging from 0 to 1.

The resulting calculation is the estimate for the natural log of ridership. The natural log of ridership can be converted to actual ridership in Excel using the following formula: =EXP(number). The number in the parentheses is the natural log of ridership calculated from the formula, and the result from the Excel formula is the estimate for actual ridership.

5.2 Applying Model #2

5.2.1 Formula

Model #2 states that ridership can be estimated as follows:

Natural log of ridership =

  $0.69 \times$ natural log of population

  $+ 1.65 \times$ percentage of population with service 6 or 7 days per week

  $+ 1.41 \times$ percentage of population with service 5 days per week

  $+ 2.01$ if agency allows same-day reservations

  $+ 1.24$ if agency requires reservations made one day in advance

17

- 0.65 if agency operates fixed-route service

- 0.12 × natural log of fare

5.2.2 Data

As described in Section 5.2.1, population data can be obtained through the American Fact Finder website. Agencies will need to estimate the percentage of their service area population receiving service 6 or 7 days per week and 5 days per week. Detailed population data for cities, census tracts or blocks, and county subdivisions can be obtained through the American Fact Finder.

5.2.3 Calculations

Similar to model #1, this equation requires using the natural log of population and fares. Section 5.1.3 describes how to make this calculation in Excel. The percentages of population with service 6 or 7 days per week and 5 days per week must be represented with decimal numbers ranging from 0 to 1.

Also similar to model #1, the resulting calculation is an estimate of the natural log of ridership. As described in in Section 5.1.3, the natural log of ridership can be converted to actual ridership using the following formula in Excel: =EXP(number).

6.   CONCLUSIONS

Few studies have been conducted to estimate ridership models for rural, demand-response transit for the general public, and existing models use a limited set of variables. This study developed two such models. The first used data from the 2013 rural NTD, and the second used more detailed service data collected from surveys of transit agencies. While previous results have shown that the populations of older adults, people with disabilities, and people without access to a vehicle significantly impact the demand for demand-response transit services in rural areas, this study showed that a number of other factors are also significant determinants. By conducting an out-of-sample validation, the first of this study's models was found to perform better than a model estimated in a previous study.

6.1   Summary of Results and Implications

Similar to previous research, the first model showed that demographic characteristics are important. Ridership was found to significantly increase when the percentage of the population comprised of older adults or people without access to a vehicle increased. This model also showed that demand-response ridership is lower if the agency also operates a fixed-route service or if its service area overlaps that of another transit agency. The positive effect of the municipality variable showed that agencies serving smaller geographic areas with more concentrated demand and shorter trip lengths will have higher levels of ridership, holding population and all other variables constant.

Both models showed a negative effect of fares on ridership. The first model estimated a price elasticity of -0.24, and the second estimated an elasticity of -0.12. Given that the first model used data from a much larger number of transit agencies, its results may be more reliable.

In addition to these variables, it should be expected that quality of service, especially service availability, should have an important impact on ridership. The second model confirmed this hypothesis, as agencies providing more days of service were found to have higher levels of ridership. This ridership increase could be due to there simply being more hours of service, but there could also be increases in trips per service hour if riders find the higher level of service more reliable and useful. More data are needed to determine if there is an actual increase in trips per service hour. The effect of service hours per day was not found to be statistically significant, but that could be because of the limited number of observations.

Lastly, the second model also showed that advance reservation time is important. Agencies that allowed users to reserve rides on shorter notice had higher levels of ridership, and the magnitude of the effect was significant. This could have important implications as agencies attempt to improve service qualities to serve more riders. The results suggest that transportation network companies could be successful in rural areas by providing rides with a short response time.

This study provides two new tools for estimating ridership, which can be used in conjunction with existing models or peer analysis. The inclusion of a greater number of variables and more specific service information improves the performance of the models. These tools can be used by transit agencies or transportation planners to forecast demand for new demand-response services; estimate the impact of service changes, such as changes in geographic coverage, span of service, fares, reservation requirements; and project future ridership based on projected population and demographic changes.

6.2   Limitations and Need for Further Study

While these models made improvements upon previous ones, they are still limited by the data availability. Larger, multi-county agencies could not be included in model #1 because of a lack of complete service area information. Model #2 was developed with more detailed service information, but data were available for a much smaller number of agencies, with many of these agencies from the same state, which could affect the results somewhat. If a dataset were available with detailed service data from a larger number of transit agencies, a model could be developed that could potentially provide more robust results. Such as model could capture the effects of both detailed service characteristics and demographic characteristics.

This study makes improvements in predicting ridership, but there may be many factors specific to each agency and community that are not captured in the model. The out-of-sample validation showed that the first model outperformed a previous model published in TCRP Report 161, with a lower RMSE and MAE. However, it also showed that both models had a significant amount of error and that there is a large degree of variation in ridership that is difficult to model.

The model outputs could be used to identify high productivity systems, which would be systems with ridership higher than the model predicts based on population, demographic, and service characteristics. The attributes of these high-productivity systems could be studied further to better understand the factors fostering success.

REFERENCES

Bradley, Mark, and David Koffman. 2012. *Improving ADA Paratransit Demand Estimation: Regional Modeling.* TCRP Report 158, Transit Cooperative Research Program, Washington, DC: Transportation Research Board.

Godavarthy, Ranjit Prasad, Jeremy Mattson, Del Peterson, and Jill Hough. 2015a. "Developing a Method for Assessing National Demand-Response Transit Level of Service." *Journal of Public Transportation* 18 (4): 1-15.

Godavarthy, Ranjit, Jeremy Mattson, Patrick Nichols, Del Peterson, and Jill Hough. 2015b. "Developing a Method for Assessing National Demand Response Transit Level of Service." National Center for Transit Research.

Goodwill, Jay A., and Ann Joslin. 2013. "Forecasting Paratransit Services Demand - Review and Recommendations." National Center for Transit Research, University of South Florida.

Kittelson & Associates, Inc., Parsons Brinckerhoff, KFH Group, Inc., Texas A&M Transportation Institute, and ARUP. 2013. *TCRP Report 165: Transit Capacity and Quality of Service Manual Third Edition.* Transit Cooperative Research Program, Washington, DC: Transportation Research Board.

Koffman, David, David Lewis, David Chia, Jon Burkhardt, and Mark Bradley. 2007. *Improving ADA Complementary Paratransit Demand Estimation.* TCRP Report 119, Transit Cooperative Research Program, Washington, DC: Transportation Research Board.

Mattson, Jeremy. 2015. "Rural Transit Fact Book 2015." Small Urban and Rural Transit Center, Upper Great Plains Transportation Institute, North Dakota State University.

Vanasse Hangen Bruslin, Inc., LSC Transportation Consultants, Inc., and Erickson Consulting, LLC. 2013. *Methods for Forecasting Demand and Quantifying Need for Rural Passenger Transportation: Final Workbook.* TCRP Report 161, Transit Cooperative Research Program, Washington, DC: Transporation Research Board.