Geographic and Demographic Methodology for Peer Group Classification of Rural Demand Responsive Transportation

Kai Monast
Institute for Transportation Research and Education
North Carolina State University
Centennial Campus Box 8601
Raleigh NC 27695
Phone: 919-515-8768
Fax: 919-515-8898
Email: kcmonast@ncsu.edu

Darcy Zorio
Institute for Transportation Research and Education
North Carolina State University
Centennial Campus Box 8601
Raleigh NC 27695
Phone: 919-515-8624
Fax: 919-515-8898
Email: dbdowns@ncsu.edu

Thomas J. Cook
Institute for Transportation Research and Education
North Carolina State University
Centennial Campus Box 8601
Raleigh NC 27695
Phone: 919-515-8622
Fax: 919-515-8898
Email: tjc@ncsu.edu

Submission Date: 07/30/10
Word Count: 5,163
ABSTRACT

This research proposes the formation of rural demand response transportation peer groups based on the challenges transportation systems face from uncontrollable geographic and demographic factors. Using geographic and demographic factors that are outside of the control of the transportation systems establishes peer groups whose members share similar challenges in providing transportation services. By accounting for uncontrollable factors, the differences in performance among transportation systems are more directly due to variances in controllable factors.

The study area includes all 80 rural demand response transportation systems in North Carolina. Four factors are used to sort transportation systems into peer groups: range of service area elevation, highway density, population density, and the ratio of rural population to the total service area population. Scores of 1-5 are assigned to transportation systems for each factor to cluster transportation systems with similar characteristics. These scores are added to create a single score reflecting the geographic and demographic profile of each transportation system. Transportation systems are then sorted into peer groups based on their profile scores.

This methodology classifies transportation systems into peer groups where every transportation system in the group has a similar opportunity to perform as well as the highest performing member of its group. By accounting for uncontrollable factors, the differences between the transportation systems are primarily due to controllable factors, allowing more effective performance comparisons among transportation systems.
INTRODUCTION

Given the wide range of operating characteristics, many demand responsive transportation systems claim they are unique and cannot be effectively compared to other systems. This claim has resulted in problems when attempting to create effective peer groups.

Most efforts at developing peer groups have been based on factors that are under the control of or can be influenced by a transportation system. Purchasing new vehicles or implementing a new technology, for instance, may correlate to improved ridership, fares, and funding assistance. However, significant geographic and demographic factors are not controllable by the transportation system.

This research proposes to create demand responsive transportation peer groups based on uncontrollable (exogenous) geographic and demographic factors. Using uncontrollable geographic and demographic factors establishes peer groups whose members share similar opportunities to succeed. The degree to which a transportation system has succeeded in relation to its peers is then due to controllable factors, including fleet size, presence of technology, local funding, skills and experience of staff, etc.

Peer groups are determined using four factors for the 80 North Carolina demand response public transportation systems. The two geographic factors are range of elevation and highway miles per square mile. The two demographic factors are population density and rural population ratio. Transportation systems are assigned a score for each of the four factors.

The scores are then added to determine the degree to which geographic and demographic factors influence a transportation system’s potential operating performance. Five peer groups are established by clustering the total scores. Since the peer groups are based on uncontrollable geographic and demographic characteristics, transportation systems can be measured against their peers using performance statistics and other controllable variables.

LITERATURE REVIEW

Ripplinger (2010) uses a hierarchical cluster analysis to classify rural and small urban transit agencies into peer groupings (1). The author states that data from the Rural National Transit Database (NTD) is useful because of its consistency and uniformity, but notes that some of the data are not reliable. The study classified agencies into four groups based on service type. Systems in each group were clustered using annual vehicle-miles, annual vehicle-hours and fleet size and evaluated using operating statistics for each of these clusters. Regarding the use of operational variables to create peer groups, the study concludes that collecting service area data may result in “the greatest improvement in terms of constructing rural transit agency peer groups.” The author continues by stating that service area data will allow peer groups to be created based on uncontrollable factors.

Vaziri and Deacon (1984) provide an in depth methodology for the creation of peer groups and their potential use in rating transit performance (2). The study asserts that exogenous variables (those that cannot be controlled by the transit agency) “have no lesser impact on performance” than endogenous variables (those that can be controlled by the transit agency). The authors concluded that “peer groups should be based solely on those uncontrollable market and environmental variables that significantly influence transit performance.” Vaziri and Deacon selected six exogenous variables to be used in their study by conducting a factor analysis and using their professional judgment. This method resulted in transit systems grouped by poverty levels, size, youth population, education, automobile availability and density of development. Performance was compared within peer groups based on the assumption that “the subject system
could achieve performance levels demonstrated by others in its peer group if the proper policy
decisions were made.” The authors concluded that it is appropriate to compare transit systems
using peer groups and that these peer groups should be based on uncontrollable variables.

TCRP 136: Guidebook for Rural Demand Response Transportation: Measuring,
Assessing and Improving Performance (2009) discussed the formation of peer groups and the use
of performance measures specific to rural demand response transportation (DRT) (3). The
research acknowledged that rural DRT is diverse because of the variety of organizational
structures, service area sizes, fleet sizes and funding streams. The report describes 28 factors that
influence DRT performance and classifies each one as controllable, partially controllable or
uncontrollable based on the system’s ability to influence each factor. Many different
methodologies were discussed to assess DRT performance, including peer group comparisons.
The report created a typology of rural DRT systems and discussed various factors that affect
performance such as ridership market served, service area or operating environment and use of
advanced technology. For the research portion of the report, three service area-based criteria
(municipal DRT systems, county DRT systems and multi-county DRT systems) were used to
present performance data of the systems that participated in the study.

METHODOLOGY
The purpose of this research was to develop an adaptive methodology for determining peer
groups based on uncontrollable geographic and demographic variables. Grouping rural
transportation systems using uncontrollable geographic and demographic factors will result in
performance differences between the systems that are due to factors under the transportation
system’s control. Once systems are grouped based on these uncontrollable factors, differences in
performance among peers are explained by policies, scheduling practices, technology, and many
other such controllable factors. Transportation systems that desire to improve can compare
themselves to high performing peers and identify differences in controllable factors.

A four step process is used to determine the peer groups, as listed below:
1. Identify Study Area;
2. Calculate Factors;
3. Analyze Data;

1. Identify Study Area
The first step in determining new peer groups using the method is to determine the study area.
The study area used in this research includes all 80 rural demand response transportation systems
in North Carolina.

The rural demand response transportation systems in North Carolina are spread across a
large geographic area including mountainous regions with peaks over 6,000 feet above sea level
and flat coastal areas with island communities only accessible by boat. Some service areas are
completely rural, while others have dense urban population centers. Fixed route transportation is
available in some service areas, while in others the primary source of transportation is human
service agency-centered demand response transportation.

Table 1 shows some of the variations of the transportation systems in the study area.
TABLE 1. Study Area Characteristics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Service Area Size (miles²)</td>
<td>609</td>
<td>380</td>
<td>199</td>
<td>2,314</td>
</tr>
<tr>
<td>Fleet Size</td>
<td>19</td>
<td>11</td>
<td>2</td>
<td>69</td>
</tr>
<tr>
<td>Annual Passenger Trips</td>
<td>89,450</td>
<td>145,960</td>
<td>9,068</td>
<td>1,203,377</td>
</tr>
<tr>
<td>Annual Service Miles</td>
<td>626,795</td>
<td>630,565</td>
<td>93,084</td>
<td>4,091,492</td>
</tr>
<tr>
<td>Annual Service Hours</td>
<td>33,633</td>
<td>40,691</td>
<td>3,448</td>
<td>316,092</td>
</tr>
<tr>
<td>Passengers per Service Mile</td>
<td>0.16</td>
<td>0.20</td>
<td>0.04</td>
<td>1.84</td>
</tr>
<tr>
<td>Passengers per Service Hour</td>
<td>2.96</td>
<td>2.78</td>
<td>1.17</td>
<td>24.76</td>
</tr>
</tbody>
</table>

Source: 2009 NCDOT Operating Statistics

2. Calculate Factors

Many geographic and demographic factors were considered. Geographic factors address the relative difficulties of providing service posed by the service area. Variables used to create these factors may include service area size, speed, trip length, elevation, intersections, highway miles, road miles, or a combination of two or more variables. These data are available from the US Census, USGS, private companies, and state and local sources.

Demographic factors address difficulties in providing service related to the dispersion of trip origins and destinations. Variables used to create these factors may include population, urban population, rural population, elderly population, youth population, disabled population, households without a vehicle, population below the poverty level or median income, or a combination of two or more variables. Many demographic variables are available from the US Census, other Federal sources including NTD, and state and local sources.

The geographic and demographic data listed above and their combinations were considered as possible factors. Proposed factors were tested using the following criteria to determine their appropriateness for creating peer groups:

1. Are the data readily available, accurate, and complete? Previous peer group studies have pointed out potential issues with NTD statistics (1). Issues with data reliability exist in other data sources as well. For instance, geographic information such as road miles may not include privately owned neighborhood roads.

2. Do the factors address significant differences among the transportation systems in the study area? Factors that do not reflect differences are ill-suited for establishing peer groups.

3. Does a single geographic factor and single demographic factor adequately capture the diversity of the study area? If not, it is necessary to select additional factors.

After applying the selection criteria to potential factors, four factors are used in this analysis. Range of Elevation is calculated by subtracting the lowest elevation point from the highest elevation point in the service area (data source: NCDOT). Range of Elevation is a geographic factor that indicates the potential difficulty in operating transportation service due to lower operating speeds resulting from long driveways, steep inclines, curvy roads, etc. Transportation systems with a large range of elevation encounter the greatest degree of challenges in service provision.

However, the elevation factor alone does not adequately account for potential geographic difficulties in providing service experienced by coastal communities and isolated areas with small ranges in elevation. Therefore, it was necessary to select an additional geographic measure.
Highway Density is calculated by dividing the total length in miles of State and Federal highways by the service area size (data sources: NCDOT, Census 2000). Highway Density is a geographic factor that indicates the potential mobility network constraints, as highways tend to increase mobility options by offering a greater number of routes. It would be preferable to use all road miles in the service area for this calculation. However, including local roads in the analysis normally requires collection of local data whose quality and completeness can be questionable. Highway miles are readily available from state and Federal sources, and are accurate and complete.

The Highway Density factor accounts for coastal areas with expansive water bodies to circumnavigate and for rural areas that have few transportation network options due to lack of highways. Transportation systems with lower Highway Density encounter greater challenges in service provision.

Together, Range of Elevation and Highway Density provide an effective measure of the difficulty of providing service posed by the geographic constraints of the service area.

Population Density is calculated by dividing the total population by the service area size (data source: Census 2000). Population Density indicates the relative proximity of trip origins. Transportation systems with lower Population Density will be more likely to have longer trip lengths, which will be more difficult to serve.

However, Population Density alone does not account for the potential dispersion of destinations. Some rural transportation systems have urban clusters and/or urbanized areas within their service area that are able to provide access to basic employment, medical, shopping, and educational facilities. Other rural transportation systems may have similar overall population densities, but do not include urban clusters or urbanized areas—forcing the transportation system to travel outside their service area to provide access to these basic employment, medical, shopping, and educational facilities.

Rural Population Ratio addresses this concern by dividing the rural population by the total population of the service area using Census 2000 data and definitions. Rural Population Ratio indicates the demand for trips outside of the service area, as rural areas will have less services available within the area. Leaving the service area to provide trips can be costly, time consuming, and inefficient for transportation systems. A service area with no urban population (Rural Population Ratio of 1.0) does not include an urban cluster or urbanized area within its boundaries. Transportation systems with higher Rural Population Ratios will have a higher likelihood of costly and time consuming out of service area trips.

Together, Population Density and Rural Population Ratio measure the difficulties in providing service posed by the demographic constraints of the service area.

Rural transportation systems in North Carolina are allocated funding for and expected to provide general public service, which is why persons with disabilities, households without vehicles, population below the poverty level, and other similar transportation disadvantaged factors are not included.

The presence of a separate transportation system operating fixed-route within a rural transportation system’s service area was also considered for use as a factor. These fixed-route trips are not provided by the rural transportation system, but could impact a rural system’s demand for service. After analysis, it was determined that the presence of a fixed route system was inversely related to the Rural Population Ratio. Therefore, inclusion of a fixed route factor was not necessary.
Table 2 below shows the mean, mode, minimum, maximum, and standard deviation for the four selected factors, as applied to the North Carolina study area.

**TABLE 2. Descriptive Statistics of Factors**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Mode</th>
<th>Standard Deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range of Elevation</td>
<td>1451.0</td>
<td>560.5</td>
<td>1723.0</td>
<td>35.0</td>
<td>5556.0</td>
</tr>
<tr>
<td>Highway Density</td>
<td>1.7</td>
<td>1.8</td>
<td>0.6</td>
<td>0.5</td>
<td>3.1</td>
</tr>
<tr>
<td>Population Density</td>
<td>180.8</td>
<td>97.9</td>
<td>221.0</td>
<td>9.5</td>
<td>1321.5</td>
</tr>
<tr>
<td>Rural Population Percent</td>
<td>0.62</td>
<td>0.65</td>
<td>0.28</td>
<td>0.04</td>
<td>1.0</td>
</tr>
</tbody>
</table>

3. Analyze Data

Numerous methods are available to cluster transportation systems for each factor. Since this research is intended to be easily replicated, it was necessary to select a common and easy to use clustering method. The selected method uses Jenks Natural Breaks, also known as goodness of variance fit (GVF). This method incorporates common statistical formulas that measure variance in an iterative process. This process optimizes homogeneity within the peer groups (4, 5).

The Jenks process is incorporated into GIS software as a standard classification method. Therefore, all transportation system data and factor data are linked to GIS software for analysis and visual representation.

Two of the selected factors, range of elevation and population density, have broad ranges of values and significant differences between the mean and the mode, indicating that there are outliers in the datasets (see Table 2). It is difficult to classify datasets with many outliers into a small number of categories that will adequately describe the members of the group. The researchers determined through trial and error that those two factors require a minimum of five categories to sufficiently create groups of transportation systems that are similar. For consistency, five categories are used for all four factors.

The groups are assigned a score from 1-5 for each factor to cluster similar transportation systems. For *Range of Elevation, Population Density* and *Rural Population Ratio*, the transportation systems in the group with the highest values receive a score of five and the transportation systems in the group with the lowest values receive a score of one. For *Highway Density*, the transportation systems in the group with the highest values receive a score of one and the transportation systems in the group with the lowest values receive a score of five.

Figures 1-4 display the clusters for each factor, along with the group size (n) and number of transportation systems in each range.
FIGURE 1. Range of Elevation

FIGURE 2. Highway Density
4. Create Peer Groups

The methodology could conclude with four separate grouping schemes based on individual factors. However, combining the four schemes to create one overall grouping method is more valuable for researchers and practitioners. It was determined that there should be between 5 and 25 transportation systems in each peer group to best classify similar transportation systems while having a manageable total number of groups. Four peer groups were attempted, but that method did not provide a fine enough distinction between peer groups and one group’s size did not meet the goal of 5-25 systems in each group. Therefore, five peer groups are used to categorize the 80 transportation systems in the study area.

Consideration was given to assigning more importance to certain factors. Equal weights are used because it is not possible to scientifically determine whether a factor should be weighted higher or lower than other factors.
To establish the peer groups, a transportation system’s assigned score for each factor is summed to determine the overall geographic and demographic profile. The Jenks Natural Breaks method is repeated to classify the systems into five peer groups based on their profiles. The resulting peer groups consist of transportation systems that share similar combined geographic and demographic profiles. Members of each peer group are considered to experience a similar degree of difficulty in providing transportation service in their service area. Figure 5 displays the proposed peer groups.

FIGURE 5. Proposed Peer Groups

Selected descriptive statistics for the proposed peer groups are displayed in Table 3. The descriptive statistics have large variations within each peer group, which would be expected because the peer groups are determined based on uncontrollable geographic and demographic factors instead of factors controllable by the transportation system.

TABLE 3. Peer Group Descriptive Statistics for Selected Characteristics

<table>
<thead>
<tr>
<th>Peer Group</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group Size (n=80*)</td>
<td>6*</td>
<td>14</td>
<td>21</td>
<td>25</td>
<td>14</td>
</tr>
<tr>
<td>Daily Service Miles</td>
<td>8,357</td>
<td>2,370</td>
<td>2,591</td>
<td>1,709</td>
<td>1,113</td>
</tr>
<tr>
<td>Average</td>
<td>6,084</td>
<td>1,165</td>
<td>1,673</td>
<td>1,054</td>
<td>475</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1,900</td>
<td>381</td>
<td>704</td>
<td>358</td>
<td>505</td>
</tr>
<tr>
<td>Minimum</td>
<td>15,737</td>
<td>5,634</td>
<td>6,574</td>
<td>5,004</td>
<td>2,258</td>
</tr>
<tr>
<td>Maximum</td>
<td>12</td>
<td>22</td>
<td>24</td>
<td>17</td>
<td>11</td>
</tr>
<tr>
<td>Fleet Size</td>
<td>32</td>
<td>22</td>
<td>24</td>
<td>17</td>
<td>11</td>
</tr>
<tr>
<td>Average</td>
<td>10.8</td>
<td>5.9</td>
<td>15.2</td>
<td>9.3</td>
<td>2.9</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>21</td>
<td>13</td>
<td>7</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Minimum</td>
<td>42</td>
<td>34</td>
<td>69</td>
<td>37</td>
<td>17</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.13</td>
<td>0.16</td>
<td>0.12</td>
<td>0.19</td>
<td>0.17</td>
</tr>
<tr>
<td>Passengers per Service Mile</td>
<td>0.08</td>
<td>0.08</td>
<td>0.04</td>
<td>0.35</td>
<td>0.10</td>
</tr>
<tr>
<td>Average</td>
<td>0.04</td>
<td>0.08</td>
<td>0.07</td>
<td>0.06</td>
<td>0.05</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.26</td>
<td>0.36</td>
<td>0.23</td>
<td>1.84</td>
<td>0.39</td>
</tr>
<tr>
<td>Minimum</td>
<td>2.14</td>
<td>2.94</td>
<td>2.31</td>
<td>3.58</td>
<td>3.18</td>
</tr>
<tr>
<td>Maximum</td>
<td>3.96</td>
<td>10.38</td>
<td>3.58</td>
<td>24.76</td>
<td>6.75</td>
</tr>
</tbody>
</table>

* One transportation system in Group 1 does not report performance data
ADAPTING THE METHODOLOGY

The methodology may be adapted to determine peer groups for other study areas. It is believed that the four factors are appropriate for establishing peer groups in other areas, but it may be necessary to adjust the number of groups within each factor. Increase the number of groups used in each factor if there is a large variation among the transportation systems’ characteristics. Decrease the number of groups if there is little variation.

It may be necessary to adjust the weights or eliminate one of the factors. Researchers are encouraged to use professional judgment when determining factor weights.

For study areas where some transportation systems have long distances to major medical, educational, and employment centers, it may be necessary to include a factor that will account for these distances. One potential factor is miles to such facilities. Those data may need to be calculated manually.

Finally, care should be taken when selecting multiple factors, to avoid those that have a close correlation.

CONCLUSIONS

This research proposes categorizing rural demand responsive transportation systems based on uncontrollable geographic and demographic profiles. This process can provide operators and researchers with peer groups that experience similar levels of geographic and demographic challenges within their service areas.

This scheme groups transportation systems so each system in a peer group has a similar opportunity to perform as well as the highest performing member of its group. By accounting for uncontrollable factors, the differences between the transportation systems’ performance will be primarily due to controllable factors. The transportation system director, staff, and/or governing board can work together to adjust the factors under their control to improve performance.

In conclusion, this peer group methodology for classifying rural demand response transportation systems based on uncontrollable factors can be applied to other study areas because of the general availability of the data and simplicity of the process.

FUTURE RESEARCH

Once a peer grouping methodology has been established, the next step is to identify the appropriate measures used to compare transportation systems within a peer group. Passengers per service hour is an industry-wide factor of performance that could be used. Other measures may be based on financial performance, such as Federal and State subsidy per passenger trip.

Using the performance of the highest-achieving system in a peer group as a target, the lower performing peers could identify and address controllable factors that may be hampering their performance. If improvements in performance were realized by these efforts it would strengthen the legitimacy of these results.

Finally, this methodology should be applied to other study areas to test its validity and adaptability.

REFERENCES


**ACKNOWLEDGEMENTS**

Special thanks to Debra Collins, Judson Lawrie, Jeremy Scott, and Bastian Schroeder at the Institute for Transportation Research and Education for contributing knowledge and expertise to this paper.