FINAL REPORT

DEVELOPING REFINED ESTIMATES OF INTERCITY BUS RIDERSHIP

DTFH6114P00097 4.23.2015



PREPARED FOR: FEDERAL HIGHWAY ADMINISTRATION

SUBMITTED BY: RSG 55 RAILROAD ROW WHITE RIVER JUNCTION, VT 802.295.4999 WWW.RSGINC.COM

DEVELOPING REFINED ESTIMATES OF INTERCITY BUS RIDERSHIP

PREPARED FOR: FEDERAL HIGHWAY ADMINISTRATION

CONTENTS

1.0		INTRODUCTION1	
2.0		RESEARCH APPROACH2	
	2.1	Data Sources2	
	2.2	Analytical Concept and Steps	
	2.3	Research Products	
3.0		COMPILE EXISTING DATA	
	3.1	Overview	
	3.2	Schedule Data Overview	
	3.3	NEC Schedule Data and Survey Data5	
		Sampling and Survey Administration	
		Questionnaire Design	
		Estimating Ridership	
		Processing Survey Data	
		Geography11	
		Estimated Weekly Trip Volumes	
	3.4	Defining the Top 200 Markets 12	
	3.5	Schedule Data for the Top 200 Markets	
	3.6	Markets by Region	
		Region 1: North East	
		Region 2: South	
		Region 3: Florida	

		Region 4: Midwest	. 30
		Region 5: Texas	. 34
		Region 6: North West	. 37
		Region 7: South West	. 39
		Region 8: Inter-regional	. 43
4.0		ESTIMATED BUS RIDERSHIP FOR THE TOP 200 MARKETS	46
	4.1	Estimation Data sources	46
		Schedule Database	. 46
		Northeast Origin-Destination Study	. 46
		Core-Based statistcal Areas	. 47
	4.2	Estimation Approach	47
		Estimation Steps for One Bus Trip	. 47
		Estimation for The Entire Schedule Database	. 48
		Discussion of the Estimation Approach	. 49
	4.3	Applying the Estimation Approach	49
		Scoring Approach	. 49
		Estimating the Target Load Factor	. 50
	4.4	Validation	50
	4.5	Results: Bus Ridership in the Top 200 Markets	51
5.0		FUTURE YEAR (2040) BUS RIDERSHIP ESTIMATES	63
6.0		RECOMMEND APPROACH FOR FURTHER REFINEMENT AND UPDATING	66
	6.1	Repeatable Analytical Process	66
	6.2	Recommendations for Refinement	66

List of Figures

FIGURE 1: SUBMARKET MAP	11
FIGURE 2: HISTOGRAM OF TOP 200 MARKETS' DEMAND	16
FIGURE 3: TOP 200 MARKETS BY REGION	17
FIGURE 4: TOP 200 MARKETS' SCHEDULE DATA	18
FIGURE 5: POPULATION IN THE NORTH EAST REGION	22
FIGURE 6: LONG DISTANCE TRIPS IN THE NORTH EAST REGION	22
FIGURE 7: SCHEDULED BUS SERVICES IN THE NORTH EAST REGION	23
FIGURE 8: POPULATION IN THE SOUTH REGION	26
FIGURE 9: LONG DISTANCE TRIPS IN THE SOUTH REGION	26
FIGURE 10: SCHEDULED BUS SERVICES IN THE SOUTH REGION	27
FIGURE 11: POPULATION IN THE FLORIDA REGION	29
FIGURE 12: LONG DISTANCE TRIPS IN THE FLORIDA REGION	29
FIGURE 13: SCHEDULED BUS SERVICES IN THE FLORIDA REGION	30
FIGURE 14: POPULATION IN THE MIDWEST REGION	31
FIGURE 15: LONG DISTANCE TRIPS IN THE MIDWEST REGION	32
FIGURE 16: SCHEDULED BUS SERVICES IN THE MIDWEST REGION	32
FIGURE 17: POPULATION IN THE TEXAS REGION	35
FIGURE 18: LONG DISTANCE TRIPS IN THE TEXAS REGION	35
FIGURE 19: SCHEDULED BUS SERVICES IN THE TEXAS REGION	36
FIGURE 20: POPULATION IN THE NORTH WEST REGION	38
FIGURE 21: LONG DISTANCE TRIPS IN THE NORTH WEST REGION	38
FIGURE 22: SCHEDULED BUS SERVICES IN THE NORTH EAST REGION	39
FIGURE 23: POPULATION IN THE SOUTH WEST REGION	40
FIGURE 24: LONG DISTANCE TRIPS IN THE SOUTH WEST REGION	41
FIGURE 25: SCHEDULED BUS SERVICES IN THE SOUTH WEST REGION	41
FIGURE 26: LONG DISTANCE TRIPS AMONG REGIONS	44
FIGURE 27: SCHEDULED BUS SERVICES AMONG REGIONS	44
FIGURE 28: DISTANCE SCORE INTERPOLATION CURVE	50
FIGURE 29: VALIDATING THE ESTIMATION METHOD AGAINST THE NEC STUDY RESULTS	51
FIGURE 30: HISTOGRAM OF ESTIMATED WEEKLY BUS RIDERSHIP	51
FIGURE 31: ESTIMATED 2014 WEEKLY BUS RIDERSHIP IN TOP 200 MARKETS	52
FIGURE 32 ESTIMATED 2014 WEEKLY BUS RIDERSHIP IN THE NORTH EAST REGION	55
FIGURE 33: ESTIMATED 2014 WEEKLY BUS RIDERSHIP IN THE SOUTH REGION	56
FIGURE 34: ESTIMATED 2014 WEEKLY BUS RIDERSHIP IN THE FLORIDA REGION	57
FIGURE 35: ESTIMATED 2014 WEEKLY BUS RIDERSHIP IN THE MIDWEST REGION	58
FIGURE 36: ESTIMATED 2014 WEEKLY BUS RIDERSHIP IN THE TEXAS REGION	59
FIGURE 37: ESTIMATED 2014 WEEKLY BUS RIDERSHIP IN THE NORTH WEST REGION	60
FIGURE 38: ESTIMATED 2014 WEEKLY BUS RIDERSHIP IN THE SOUTH WEST REGION	61
FIGURE 39: ESTIMATED 2014 WEEKLY BUS RIDERSHIP AMONG REGIONS	62

List of Tables

TABLE 1: CITIES INCLUDED IN STUDY	7
TABLE 2: SAMPLED TRIPS BY SITE AND DATE	8
TABLE 3: SAMPLED TRIPS BY CARRIER	9
TABLE 4: COMPLETED SURVEYS BY METHOD	9
TABLE 5: BOARDING AND ALIGHTING MATRIX	12
TABLE 6: TOP 10 ORIGIN-DESTINATION PAIRS	12
TABLE 7: NATIONAL LONG DISTANCE TRIPS BY MODE	14
TABLE 8 LONG DISTANCE TRIP MODE SHARE BY DISTANCE BAND	14
TABLE 9: COMPARING TOP 200 MARKET SIZE AGAINST THE NATIONAL TOTAL	15
TABLE 10 SUMMARY OF TOP 200 MARKETS	15
TABLE 11: DISTRIBUTION OF TOP 200 MARKET DEMAND	16
TABLE 12 DISTRIBUTION OF TOP 200 MARKETS' SUPPLY	18
TABLE 13 TOP BUS CARRIERS WITH MARKET SHARE	19
TABLE 14 NUMBER OF CARRIERS AND REGIONAL SUPPLY SHARE BY REGION	20
TABLE 15 CARRIERS SERVE MORE THAN ONE REGIONS	20
TABLE 16: TOP 200 MARKETS, REGIONAL POPULATION, INTERCITY TRIP DEMAND AND BUS SERVICES	21
TABLE 17: TOP 200 MAKETS LOCATED IN THE NORTH EAST REGION	23
TABLE 18: TOP 200 MARKETS LOCATED IN THE SOUTH REGION	28
TABLE 19: TOP 200 MARKETS LOCATED IN THE FLORIDA REGION	30
TABLE 20: TOP 200 MARKETS LOCATED IN THE MIDWEST REGION	33
TABLE 21: TOP 200 MARKETS LOCATED IN THE TEXAS REGION	36
TABLE 22: TOP 200 MARKETS LOCATED IN THE NORTH WEST REGION	39
TABLE 23: TOP 200 MARKETS LOCATED IN THE SOUTH WEST REGION	
TABLE 24: TOP 200 MARKETS IN THE INTER-REGIONAL GROUP	45
TABLE 25: SUMMARY OF ESTIMATED 2014 WEEKLY BUS RIDERSHIP BY REGION	53
TABLE 26: 20 MARKETS WITH HIGHEST RIDERSHIP IN 2014	
TABLE 27 REGIONAL POPULATION AND EMPLOYMENT: 2014 AND 2040	64
TABLE 28: REGIONAL MEAN GROWTH FACTOR AND ESTIMATED WEEKLY RIDERSHIP IN 2040	64
TABLE 29: 20 MARKETS WITH HIGHEST RIDERSHIP IN 2040	65

1.0 INTRODUCTION

Intercity bus ridership estimates for the year 2008 were developed as part of FHWA's Traffic Analysis Framework (TAF) Multimodal Interregional Passenger Travel Origin Destination Data project. Those estimates were based on extrapolations from the 1995 American Travel Survey. However, the intercity bus market changed considerably in the interval between 1995 and 2008 (and has continued to change since 2008) and simple extrapolations apparently did not capture the full extent of those changes. A review of the estimates by the American Bus Association (ABA) and its member companies indicated that these initial TAF estimates were likely too low.

Fortunately, there are several other types of data that, together, can be used to provide current, refined estimates of bus ridership. The specific objectives of this project were to:

- Develop and apply an approach that can be used to more reliably estimate, and annually update, intercity bus ridership volumes based on empirical observations of the number of bus services actually operated.
- Apply this approach to generate bus ridership estimates for the top 200 U.S. origin-destination pairs, based on publically available schedules.
- Develop and implement an approach to building core bus service/network supply data that can be used to support other FHWA national policy information needs.

This report documents the approach to fulfilling these requirements that the project team followed, the definition of the top 200 markets (where a market is a pair of metropolitan areas defined using the U.S. Census Bureau's Core Based Statistical Areas (CBSAs)), the characteristics of those markets, the development of schedule data for those markets, and the results of bus ridership estimation for those markets. The report is accompanied by a complete spreadsheet containing the top 200 markets, their bus service, and estimated ridership, and general transit feed service (GTFS) format schedule data developed by the project team for intercity bus service in the United States.

2.0 RESEARCH APPROACH

2.1 | DATA SOURCES

This project focused on the development, validation and initial application of a method to produce more reliable and updatable estimates of intercity bus ridership. The approach uses several mutually complementary, data sources:

- Existing GTFS data for intercity bus services Detailed service data are currently available for highway (National Highway Planning Network – NHPN), air (BTS T-100) and rail (AMTRAK) modes, but not for intercity bus in a comprehensive database. Some of the intercity bus carriers have organized their service data in GTFS format and we compiled GTFS data for all of the intercity services in the northeast corridor between Washington, D.C., New York City, and Boston, where the highest densities of intercity bus service are located.
- 2. Intercity bus schedule data from Russell's Guide We entered an arrangement with Russell's Guide to provide their most current data on bus schedules, both for the major carriers who subscribe to their official Guide as well as for other low-cost carriers who are not listed in the official Guide but who have agreed to provide their data to Russell's Guide for other uses. We developed procedures to convert these data into GTFS format.
- 3. Additional Intercity Schedule Data In addition to those service provides that already produce schedule data in GTFS format, and the data provide by Russell's Guide, many service provides make their schedules available in other format: on websites and in paper schedule format. In some cases, the actual schedules are masked and only available through on line booking systems. For this project we compiled data from various other sources and also converted them to GTFS format to arrive at a relatively complete set of schedule data for the top 200 intercity markets in the United States (see Chapter 3).
- 4. Northeast Corridor traveler survey The Northeast Corridor Commission (NECC) is a body created by congress to facilitate planning and cooperation for passenger rail projects in the Northeast U.S. In 2014, the NECC commissioned us to conduct a study of intercity bus travel along the rail corridor running between Washington DC and Boston MA, known as the Northeast Corridor (NEC). Boarding count data and survey data collected as part of this NEC study were used to estimate and validate the ridership models developed and applied in this project.
- 5. Data defining the key markets for the study Definitions of the top 200 metro-pairs for this analysis were made for this project based on extensive work undertaken by the FHWA's Traffic Analysis Framework (TAF) program. The immediate applicability of these data was demonstrated in work undertaken by us for the National Academies, with the cooperation of FHWA.

2.2 | ANALYTICAL CONCEPT AND STEPS

These data sources were used to produce estimates of current bus ridership in the top 200 intercity markets to support FHWA's Traffic Analysis Framework (TAF) Multimodal Interregional Passenger Travel Origin Destination Data. The ability to produce accurate ridership estimates was constrained in part by the types and coverage of data that are currently available. The survey data that were used in this approach provide only partial descriptions of bus ridership, as the NEC survey obviously covers only one region. The only reliable

way to provide complete national-level estimates for scheduled intercity buses was to attach the information from the surveys, in the form of parameters such as estimated bus load factors, to an enumeration of the intercity bus service for the top 200 intercity travel markets of interest. So, the central concept in this approach was that the survey data could be used to develop estimates of key parameters and that a complete description of intercity bus supply could be used to expand that information to create estimates for scheduled service in the top 200 markets.

Each analysis step undertaken to produce the ridership estimates, once the data were compiled, is introduced below and then fully described in the subsequent chapters of this report.

- Use NEC survey, county, and schedule data to develop a model of bus boardings and alightings that can applied to service outside of the NEC to estimate ridership. The NEC survey and associated data development work provides as complete a picture of intercity bus supply and demand as is available currently, and facilitated the development of a simulation model that is sensitive to bus service characteristics including frequency and the number of stops on a route. The collection of the survey data is described in Chapter 3 and its use to estimate bus ridership is described in Chapter 4.
- Define the top 200 markets using aggregated county to county TAF Multimodal Interregional Passenger Travel Origin Destination Data. The products of the earlier TAF projects were used to identify the top 200 intercity travel markets. Some heuristics were applied to select markets to focus on markets likely to have at least some intercity bus ridership. The heuristics were based on total trips across all modes, the size of the metropolitan areas at each end of the market, the distance between the metropolitan areas, and the presence of some intercity bus service. The definition of the top 200 markets is described in Chapter 3
- Estimate 2014 intercity bus ridership: Once the top 200 markets were identified and their schedule data collated, the model estimated using the NEC data was applied to the top 200 markets. This model produces estimates of 2014 ridership for each market. The application approach and the results of the application are presented in Chapter 4
- Estimate 2040 intercity bus ridership: A relatively simple socioeconomic growth approach was applied to the 2014 ridership estimates in order to estimate future year intercity bus ridership. The approach was similar to that used in the TAF Multimodal Interregional Passenger Travel Origin Destination Data project to estimate 2040 rail travel. The approach and results are presented in Chapter 5.

2.3 | RESEARCH PRODUCTS

The final product of this project is improved current (2014) and 2040 estimates of scheduled intercity bus ridership for the top 200 U.S. intercity travel markets, identified as described in this report using the FHWA's Traffic Analysis Framework (TAF) Multimodal Interregional Passenger Travel Origin Destination Data. This is provided in the form of a complete spreadsheet containing the top 200 markets, their bus service, and estimated ridership for 2014 and 2040.

In addition GTFS format schedule data developed by the project team for intercity bus service in the United States to support the ridership estimates are also provided in text files following the GTFS specification.

Finally, an important product of this research is the repeatable analytical process codified in software programs, which means that the base and future year ridership estimates can be updated as new schedule data

are made available. The rapidly changing nature of the intercity bus market means that new carriers are entering the market and new services are being added very regularly. This means that any "current" estimates are a snapshot and, several months later, the service and resulting ridership in some or many markets might have changed.

3.0 COMPILE EXISTING DATA

3.1 | OVERVIEW

Data to support the development of the bus ridership estimates were developed in several stages, initially to support estimation of bus load factors and then to support the bus ridership estimates in the top 200 markets nationwide. This Chapter describes

- The development of schedule data, including the approach to convert Russell's Guide and other schedule data into GTFS format
- The NEC survey project, including the collection of survey data, boarding counts, and schedule data
- The analysis of TAF Multimodal Interregional Passenger Travel Origin Destination Data in order to define the top 200 intercity travel markets for use in this project
- The refined schedule data for the top 200 intercity travel markets

3.2 | SCHEDULE DATA OVERVIEW

This project relied on a dataset of bus schedules for intercity passenger service providers throughout the United States. We acquired a national bus schedule database from Russell's Guide. Russell's Guide publishes a compilation of intercity bus schedules for the United States and Canada. To support the publication, Russell's Guide develops and maintains a bus schedule database. Russell's Guide shared this dataset with us for this project.

The dataset was improved in two stages. In the first stage, the dataset was improved for the NEC only, as part of the NEC survey project, which is described in more detail below. Russell's Guide added additional carriers to the database identified by us in order to provide a complete description of NEC service. We then improved the dataset for the Megabus and Bolt Bus carriers. For Mega Bus, we developed a program to automatically read the schedules from the Megabus website, and to automatically combine schedule options that were actually part of the same bus vehicle trip into one trip in the dataset. A similar process was developed for the Bolt Bus data to ensure that duplicate trips did not appear in the dataset.

In the second stage, the improvements for Megabus and Bolt Bus were extended to the entire nation: the schedules for both carriers were automatically retrieved from their websites, and then the schedule options were combined to eliminate duplicate trips.

The Russell's guide dataset was converted to GTFS format, in order to simplify and clarify the bus schedule data. The GTFS version of the schedule data was then summarized according to s CBSAs. These areas are defined by the census bureau define boundaries of urban areas along county lines. Each stop in the schedule data was assigned to one of the CBSAs if it fell in one. Then the description of each bus trip was simplified to include only the sequence of CBSAs it visited. Finally, the number of bus trips that serve each CBSA pair in a week was tabulated.

3.3 | NEC SCHEDULE DATA AND SURVEY DATA

The NECC is a body created by congress to facilitate planning and cooperation for passenger rail projects in the Northeast U.S. In 2014, the NECC commissioned us to conduct a study of intercity bus travel along the rail corridor running between Washington DC and Boston MA, known as the NEC. We saw considerable

overlap in its work with NECC and FHWA, as the NECC project allowed for a detailed investigation of bus travel in the Northeast. What follows is a description of the effort for the NECC; the results from this study were used to support the analysis described in this report, as discussed in later chapters.

SAMPLING AND SURVEY ADMINISTRATION

During the NECC study, we surveyed and counted bus riders at six sites, each for two days. Those sites included Union Station in Washington DC, South Station in Boston, and the Port Authority Bus Terminal in New York City. The three remaining sites are all areas of Manhattan that are rich in curbside bus service; Midtown, Chinatown, and the area surrounding the Javits Center. Most intercity bus travel in the NEC region originates or terminates at one of these sites.

Intercity bus passengers on sampled trips were handed a paper questionnaire as they prepared to board. Respondents had the option to return the questionnaire directly to survey staff, return the questionnaire by Business Reply Mail, or take the survey online through a link and QR code provided on the questionnaire cover. Survey staff recorded the range of unique serial numbers distributed on each bus; this information was later associated with survey responses to ensure accurate information about carrier, departure time, and departure location. Passengers who refused a survey were counted, yielding an accurate count of passengers on each surveyed bus.

We sampled a total of 322 trips, allocated proportionally based on the number of trips originating at each of the six sites. All sampled trips originated in Boston, Washington DC, or New York City, but trips going to all other major corridor destinations were sampled. Table 1 provides a list of all the cities included in the study, and Table 2 shows the number of trips sampled at each location by date. We also took care to sample trips from a variety of different bus carriers, including some smaller niche operators, as shown in Table 3.

A large number of surveys were mostly complete, but had missing or nonsensical information for origin and destination. Only the surveys with good origin and destination information could be used to estimate the origin-destination matrix produced as part of the NEC project; however, for the rest of the analysis, the surveys with bad origin or destination information were included. Table 4 presents a breakout of both types of completed surveys by response method (paper or web).

TABLE 1: CITIES INCLUDED IN STUDY

INCLUDED CITIES
Arlington VA
Baltimore MD
Bethesda MD
Boston MA (including Cambridge and Newton)
Bridgeport CT
Hartford CT
New Haven CT
Newark NJ
Newark DE
New York NY
Philadelphia PA (including Cherry Hill NJ)
Providence RI
Secaucus NJ
Springfield, MA
Stamford CT
Trenton, NJ
Washington DC (including Greenbelt MD)
Wilmington, DE
Worcester, MA

TABLE 2: SAMPLED TRIPS BY SITE AND DATE

SITE/DATE	SAMPLED TRIPS
Union Station	
9/18/2014	29
9/19/2014	28
South Station	
9/25/2014	24
9/26/2014	34
Javits Center	
10/1/2014	35
10/3/2014	34
Midtown	
10/1/2014	19
10/3/2014	19
PABT	
10/2/2014	30
10/3/2014	30
Chinatown	
10/2/2014	20
10/4/2014	20
Grand Total	322

TABLE 3: SAMPLED TRIPS BY CARRIER

CARRIER	SAMPLED TRIPS
BestBus	1
BoltBus	58
Bonanza	4
Eastern Shuttle	11
Focus Travel	10
Go Bus	9
Greyhound	60
Hola	8
Lucky Star	19
Megabus	70
Peter Pan	36
Rockledge Bus	7
Tripper	3
Vamoose Bus	6
Washington Deluxe	7
Yo! Bus	15

TABLE 4: COMPLETED SURVEYS BY METHOD

RESPONSE METHOD	BAD O/D INCLUDED	BAD O/D EXCLUDED
Paper	1416	897
Web	47	42
Total	1463	939

QUESTIONNAIRE DESIGN

The questionnaire asked a range of questions relating to the trip geography (origin, boarding location, alighting location, and destination), trip purpose, the respondent's travel habits, their use of electronic devices during the trip, the fare paid, and demographics.

ESTIMATING RIDERSHIP

We designed the study to estimate *one week* of ridership. The research team deliberately chose to survey during September and October when travel patterns are stable and "typical."

Since it was not feasible to do a full week of surveying at each location, estimating a week's ridership required a few simplifying assumptions: First, we assumed that all days except for Friday and Sunday are similar, and Friday and Sunday are similar to one another. For most sites, we have data from both a Friday and another day (the only exception is Chinatown, where we have data only from Thursday and Saturday).

We modeled ridership separately for each city pair. A direct bus would have only one such pair, but a bus making multiple stops would have a separate estimate for each of the possible boarding-alighting pairs on that trip. The selected model form was a Poisson model, which is a useful for modeling count data. Predictor variables included carrier, the populations of the boarding and alighting metropolitan areas, minutes from peak ("peak" was determined to be 6pm), the number of stops on the full route, and dummy variables for select cities with unique characteristics.

The boarding counts which had been collected by survey staff were assigned proportionally based on the surveys themselves. As an example, imagine a bus that travels from New York to Washington DC with a stop in Baltimore. Twenty riders board the bus in New York. We receive 10 returned surveys, five of which give Baltimore as their alighting location and five of which give Washington DC as their alighting location; in other words, half of the riders who returned surveys go to each location. In this case, a count of ten riders will be assigned to the New York – Baltimore trip, and ten will be assigned to the New York – Washington DC trip.

Once the model was estimated, it was applied to the schedule data, yielding a prediction for each possible boarding-alighting trip in the schedule. Those individual predictions were then summed to generate a boarding and alighting matrix.

PROCESSING SURVEY DATA

An experienced data entry firm manually entered data from the paper questionnaires through a unique process. To ensure accuracy, two people enter each survey, and the data are compared. In the event that there is a difference in how some response was recorded, a third person looks at the disparity and determines the final outcome. These data were then merged with the online data to create a single dataset.

Data Cleaning

Several steps were taken to ensure high quality data, including:

- Geocoding origin, destination, boarding, and alighting locations. This was done largely through an automatic process using Google Maps. In instances where the software could not find a match, locations were assigned manually. Locations are accurate to the town/city level.
- Comparing origins and destination locations to boarding and alighting locations, ensuring that each trip makes logical sense. In some cases, origins and destinations were adjusted.
- Verifying that the correct bus data (departure time, carrier, and route) was assigned to each survey response. These data were collected in the survey as well as by the survey staff, which allowed for verification that the right trip had been assigned to each record. Where there was a mismatch, records were assigned to a different trip or excluded.

Weighting

To ensure that the survey data were representative of the full riding population, surveys were weighted to the ridership estimates by submarket pair. Since surveys could only be weighted to travel volumes for submarket pairs that included the New York, Boston, and Washington areas, the weighted total for the survey data is slightly less than the total form the boarding and alighting matrix. The trips represented in the survey data account for 92% of the total in the boarding and alighting matrix (136,045 vs. 148,056).

GEOGRAPHY

Origins, boarding locations, alighting locations, and destinations were all assigned to one of 14 submarkets. Figure 1 is a map of the submarkets.

FIGURE 1: SUBMARKET MAP



ESTIMATED WEEKLY TRIP VOLUMES

Boarding and Alighting

By applying the ridership model to the schedule data, we generated a boarding and alighting matrix for the NEC region. Table 5 provides estimates of one week's worth of trip volumes between eight submarkets; these submarkets offer intercity bus service along the NEC.

TABLE 5: BOARDING AND ALIGHTING MATRIX

	Destination Submarket								
Origin Submarket	Fairfield County	Greater Baltimore /DC	Greater Boston/ Providence	Hartford Area	New Haven	New York City	Newark area	Philadel- phia area	Grand Total
Fairfield County	110		92	110	152	362			826
Greater Baltimore/DC		3,926	176			22,922	1,180	3,206	31,410
Greater									
Boston/Providence	115	163	3,267	1,298	264	15,741	353	296	21,497
Hartford Area	160		1,182		488	2,950			4,780
New Haven	214		223	514		567			1,519
New York City	261	23,346	17,231	2,911	617			17,852	62,217
Newark area		1,259	506					555	2,320
Philadelphia area		3,180	304			19,041	963		23,488
Grand Total	859	31,873	22,982	4,834	1,521	61,584	2,496	21,908	148,056

Origin and Destination

By weighting the survey data to the boarding and alighting matrix, we generated a matrix of trip volumes by ultimate origin and destination. While that entire matrix is large, the majority of trips are concentrated in a few submarket pairs. Table 6 shows the top 10 origin-destination by estimated weekly ridership. These estimates are based on the subset of complete surveys with accurate boarding and alighting information, and on the boarding and alighting estimates for those submarket pairs that include New York City, Boston, and Washington DC. The top three submarket pairs represent 62% of all intercity bus travel along the NEC.

TABLE 6: TOP 10 ORIGIN-DESTINATION PAIRS

SUBMARKET PAIR	BUS TRIPS/WEEK
Baltimore/DC/Northern VA <-> New York City	41,704
Greater Philadelphia area <-> New York City	25,569
New York City <-> Eastern Mass./Rhode Island	24,998
Baltimore/DC/Northern VA <-> Greater Philadelphia area	5,588
Baltimore/DC/Northern VA <-> Northeastern NJ	3,132
Northeastern NJ <-> Eastern Mass./Rhode Island	2,088
Mercer/Burlington Counties, NJ <-> New York City	1,461
Baltimore/DC/Northern VA <-> Eastern Mass./Rhode Island	1,213
New York City <-> New Haven/Middlesex Counties, CT	1,185
Hartford County CT <-> Eastern Mass./Rhode Island	1,163
Total	108,101

3.4 | DEFINING THE TOP 200 MARKETS

The top 200 intercity travel markets were defined based on the national long distance trip tables produced by the TAF Multimodal Interregional Passenger Travel Origin Destination Data project. That nationwide, long

distance origin destination (OD) trip table is a combination of four different modes: auto, bus, air and rail. Long distance trips for that project were defined as any trip longer than 100 miles.

The auto and bus OD tables were developed primarily using the 1995 American Travel Survey ATS) as the baseline, as well as some additional sources to support elements of data development and validation. The ATS was selected as the primary source of auto and bus travel data as it is the only comprehensive source of national origin-destination data on long-distance passenger travel. The air and rail trip tables for 2008 were developed directly using data from the Bureau of Transportation Statistics (BTS), Federal Aviation Authority (FAA), and Amtrak.

The auto trip generation was done for two purposes: business and non-business. Trip attractions were generated by linear regression specifications based on (origin county) employment and population. The balanced productions were then estimated by distributing the attractions using a destination choice model in multinomial logit formulation for each trip purpose.

Bus trips were estimated using home-based trip generation rates and a trip distribution model. A single, nation-wide set of trip rates was developed that related the number of annual long distance bus trips per household with age, income and auto ownership. This set of rates produced an estimation of residential-based bus trips for 2008. These trips were then distributed via a state-level destination choice model.

For bus and auto, special generators also considered for trips 1) crossed border entry points between the United States and Canada or Mexico and 2) destined for popular recreation locations such as national parks, Las Vegas and Orlando. Data from the national parks service (NPS), BTS for cross border inbound trips, visitors bureau from Las Vegas and Orlando were allocated to their production/attraction zones.

The bus trips were later updated using data provided by the ABA based on their motorcoach census and a special survey of its membership. The table below is based on these updated results, published in January 2015, but the remainder of the analysis in the report was completed prior to the updating of the bus trip table.

The 2008 air OD table was developed by blending three primary sources of data. Two national datasets, Airline Origin and Destination Survey Data (DB1B) and T-100 data, describe air passenger trips between airports; and a collection of airport ground-access surveys describes air passenger trips from trip origins (e.g. homes, offices, hotels) to airports and from airports to trip destinations. The combination of trip origin to airport, airport to airport, and airport to trip destination portrays a complete air passenger trip from origin to destination.

The 2008 rail OD table was created using a similar approach to the 2008 air OD table, by blending data on station-to-station trips provided by Amtrak with data and models for trips accessing stations.

As shown in Table 7, the developed national long distance trip table covers a 1.6 billion passenger trips, with an average annual growth rate of 3.9% from 1995 to 2008.

TABLE 7: NATIONAL LONG DISTANCE TRIPS BY MODE

PARAMETER	AUTO	BUS	AIR	RAIL	TOTAL
1995 ATS	813,858,000	20,445,000	161,165,000	4,994,000	1,000,462,000
2008 Estimate	1,225,711,728	190,665,970	221,161,444	11,980,162	1,649,519,304
Share 1995	81.3%	2.0%	16.1%	0.5%	100%
Share 2008	74.3%	11.6%	13.4%	0.7%	100%
Total Growth (1995 to 2008)	50.6%	832.6%	37.2%	139.9%	64.9%
Annual Total Growth (1995 to 2008)	3.2%	18.7%	2.5%	7.0%	3.9%

For this project, the top 200 markets were defined based on the total number of the long distance trips (all four modes combined), with some additional restrictions. A top market, in this study, is the total two-way passenger travel demand between a pair of zones. The project team selected the zoning system of CBSAs. A CBSA is typically an aggregation of multiple counties. Using a County-CBSA look up table, the original county-to-county OD trip tables produced during an interim step of the TAF Multimodal Interregional Passenger Travel Origin Destination Data project were aggregated into a total of 443,211 pairs of CBSA that have distances between them of greater than 100 miles. For each pair of CBSAs (A and B), a total market demand is the summation of the trips from A to B and trips from B to A. 439,588 pairs of CBSAs have at least 1 trip between the zones.

As the distance between an OD pair increases, the advantage of air travel against road based transportation and rail grows. Table 8 shows the mode share for difference distance brackets. Those pairs with greater distance, e.g. Los Angeles to New York, might have high total demand, but are unlikely to generate a significant number of bus trips. The distance threshold for a top market for this study was therefore set to be between 100 and 1000 miles.

DISTANCE BAND (IN MILES)	NUMBER OF PAIRS	AUTO	BUS	AIR	RAIL
100 - 499	81,074	90.6%	2.6%	5.2%	1.6%
500 - 999	145,633	60.5%	2.1%	36.8%	0.5%
1000 - 1999	153,027	19.7%	1.4%	78.7%	0.2%
2000 - 2999	56,425	1.7%	0.8%	97.4%	0.1%
3000 +	7,052	2.6%	0.7%	96.7%	0.0%

TABLE 8 LONG DISTANCE TRIP MODE SHARE BY DISTANCE BAND

Some of the very large metropolitan areas contain multiple CBSAs that may extend to more than 100 miles apart. Travel demand between the CBD of such a metropolitan area and its surrounding suburban towns do not fit this study's interest in intercity long-distance bus travel. To screen out these cases, the project team applied a population threshold to exclude trips from and to any CBSA with populations of less than 100,000.

The presence of existing bus service was used to identify markets with at least some intercity bus ridership. As explained earlier in this chapter, the project team built a dataset of nationwide intercity bus service for this project. These data were used to screen out any pair of CBSAs with on bus service.

The top 200 markets are those pairs with the highest total number of intercity trips that meet all three selection criteria. Table 9 shows some summary statistics of the 200 markets against the national total.

	N	MEAN DISTANCE	MEAN TRIPS	MINIMUM TRIPS	FIRST QUARTILE TRIPS	THIRD QUARTILE TRIPS	MAXIMUM TRIPS
Top 200 market	200	236	2,223,884	748,160	1,049,860	2,051,696	22,479,798
National total	443,211	1,106	4,727	0	4	586	22,479,798

TABLE 9: COMPARING TOP 200 MARKET SIZE AGAINST THE NATIONAL TOTAL

Table 10 shows a profile of the top 200 markets relative to all of the long distance travel in the United States: the account for 21% of the total long distance trips in the continental US, and 69% of the population is included in at least one end of a market.

TABLE 10 SUMMARY OF TOP 200 MARKETS

PARAMETER	TOP MARKETS	US TOTAL	SHARE
Total number of trips	444,776,724	2,094,968,702	21%
Average number of trips (per market)	2,223,884	4,727	
2008 population	196,256,406	283,806,818	69%

Figure 2 shows that, among the top 200 markets, three outlier markets exist with high demand: Los Angeles to Riverside, Los Angeles to San Diego and New York to Philadelphia. Table 11 shows the count share as well as the market share by category. Over half of the markets have total trips in the one to two million range, which contributes 32.7% of the top markets' demand; the top 3 markets by themselves have a 14% share of the top markets' demand.

FIGURE 2: HISTOGRAM OF TOP 200 MARKETS' DEMAND

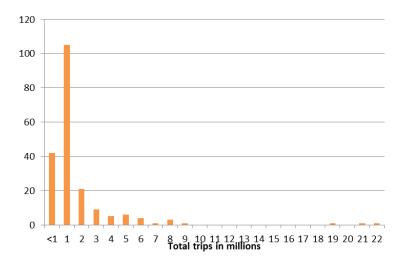
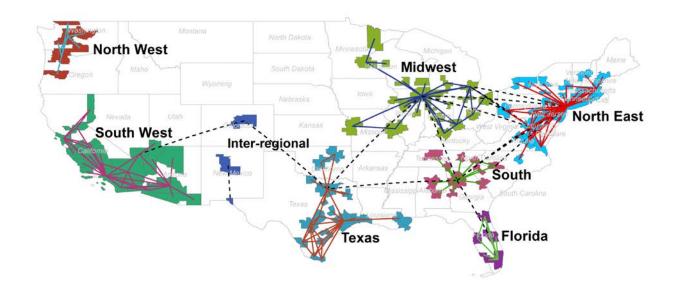


TABLE 11: DISTRIBUTION OF	TOP 200 MARKET DEMAND
----------------------------------	------------------------------

TOP TRIP NUMBER	COUNT	COUNT SHARE	MARKET SHARE
<1,000,000	42	21.0%	8.0%
1,000,000 - 1,999,999	105	52.5%	32.7%
2,000,000 - 2,999,999	21	10.5%	11.0%
3,000,000 - 6,999,999	24	12.0%	18.8%
7,000,000 - 9,999,999	5	2.5%	9.6%
19,000,000 - 23,000,000	3	1.5%	14.1%

The top 200 markets include 133 CBSAs. Figure 3 shows those CBSAs, together with a line representation of each of the markets. Excluding a small number of relatively long distance pairs, there are clear clusters of markets, most obviously in the Northeast and Southwest parts of the US. Other important regional markets include Texas, Florida, the Chicago area, the South and the Northwest. The seven regional market grouping are referred as a region in the remainder of this report. The remaining markets, most of which connect large metropolitan areas across regions, are grouped into Region 8 "inter-regional". A more detailed description of each region is presented in Section 3.6.

FIGURE 3: TOP 200 MARKETS BY REGION



3.5 | SCHEDULE DATA FOR THE TOP 200 MARKETS

Using the processes described in Section 3.2, the project team derived the number of weekly scheduled bus services for the top 200 markets. Figure 4 displays the services. On average, there are 136 buses per week among the top markets, for a total of 27,259 services between the top market pairs.

Table 12 shows that four ODs pairs have high number of services, two in the 1,000 to 1,200 thousand range and two in the 1,600 to 1,800 range. Unlike the outlier markets in terms of the total number of intercity trips, all four pairs are within the North East region. Over half of the top 200 markets have less than one hundred buses per week and in total sum to a 21.2% share of all of the service in the top 200 markets, and is almost identical to that of the top four by themselves.

FIGURE 4: TOP 200 MARKETS' SCHEDULE DATA

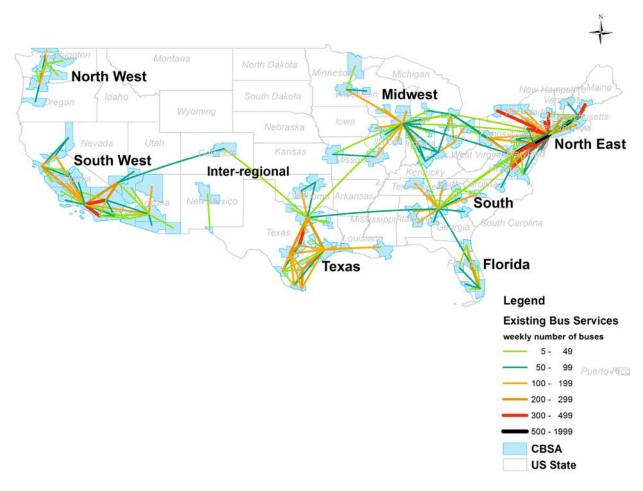


TABLE 12 DISTRIBUTION OF TOP 200 MARKETS' SUPPLY

NUMBER OF BUS SERVICES PER WEEK	COUNT	COUNT SHARE	MARKET SHARE
<100	117	58.5%	21.2%
100-199	48	24.0%	24.0%
200 - 299	18	9.0%	16.0%
300 - 499	13	6.5%	17.6%
1,000 - 1,199	2	1.0%	8.3%
1,600 - 1,899	2	1.0%	12.8%

The schedule data include fifty-four carriers nationwide, thirty-seven of which serve the top 200 markets. Sixteen carriers have more than 1% shares nationwide; ten of which have more than 1% shares in the top 200 market as well. Two other carriers dedicated to the top 200 markets (Hola bus and Eastern Shuttle) also have

more than 1% shares of the top 200 only markets. Table 13 listed those carriers with their market shares. Greyhound is the largest carrier that provides more than half of all the services, followed by Megabus that has a 16% share of services in the top 200 markets.

TABLE 13 TOP BUS	CARRIERS WITH	MARKET SHARE
------------------	----------------------	--------------

CARRIER	TOTAL SERVICES	NATIONAL SHARE	TOP MARKET SERVICES	TOP MARKET SHARE
Greyhound Lines Inc.	101,620	60%	14,771	54%
Megabus	10,174	6%	4,407	16%
Jefferson Lines	9,346	6%	66	0%
Burlington Trailways	6,524	4%	105	0%
Coach USA Erie	5,457	3%	256	1%
Miller Transportation	3,598	2%	112	0%
Susquehanna Trailways	2,982	2%	49	0%
Indian Trails	2,660	2%	98	0%
Adirondack Trailways	2,584	2%	519	2%
Peter Pan Bus Lines	2,300	1%	1,035	4%
Southeastern Stages Inc.	2,065	1%	84	0%
Bolt Bus	1,910	1%	1,426	5%
Bonanza Bus Lines	1,760	1%	371	1%
Martz Trailways	1,739	1%	256	1%
Black Hills Stage Lines, Inc.	1,682	1%	0	0%
Bieber Transportation Group	1,616	1%	364	1%
Hola Bus	741	0%	741	3%
Eastern Shuttle	385	0%	385	1%
36 others	10,363	6%	2,214	8%

Most carriers focus their services within a region; and most regions have a handful of carriers. The only exception is the North East region, which have twenty-five carriers registered in our schedule database. Not surprisingly, Greyhound ranks top in every single market. Table 14 looks at the regions by number of carriers and their share; Table 15 looks at the regional shares for those five carriers that serve more than one region.

TABLE 14 NUMBER OF CARRIERS AND REGIONAL SUPPLY SHARE BY REGION

REGION	TOTAL SERVICES	TOTAL CARRIERS	MIN SHARE	MAX SHARE
North East	13,755	25	0.3%	32.8%
South	1026	4	5.5%	63.5%
Florida	512	3	3.1%	71.7%
Midwest	2,968	8	1.8%	54.8%
Texas	3,851	4	0.4%	90.8%
North West	529	3	2.6%	53.3%
South West	3,850	4	0.7%	85.7%
inter-regional	768	2	29.7%	70.3%

TABLE 15 CARRIERS SERVE MORE THAN ONE REGIONS

CARRIER	NORTH EAST	SOUTH	FLORIDA	MID WEST	TEXAS	NORTH WEST	SOUTH WEST	INTER- REGIONAL
Greyhound	31%	4%	2%	11%	24%	2%	22%	4%
Megabus	54%	5%	3%	18%	7%		8%	5%
Bolt Bus	70%					17%	13%	
Jefferson lines				79%	21%			
Lake Front Lines		50%		50%				

3.6 | MARKETS BY REGION

This section presents a portrait of each of the eight regions along the dimensions of population, total long distance trips, and total scheduled bus services. To allow for cross-region comparison, the same color scale is used for all maps. A complete list of all the CBSAs with market size and service frequency can be found in the spreadsheet accompanying this report.

Table 16 shows some comparison among the eight regions. It highlights the North East region and the South West region as the two largest segments of the market in terms of total intercity travel, each contributing to one third of total long distance travel amongst the top 200 markets. The North East region has a 30% share of population, which is consistent with its share of the long distance travel demand, but a significantly higher (50%) share of the bus service supply. In contrast, the South West region has only a 21% share of the population and a lower (14%) bus services supply share.

REGION	TOTAL POP (IN 1000)	TOTAL TRIPS (IN 1000)	SERVICE PER WEEK	TRIP RATES (PER CAPITA)	SERVICE RATES (PER 1000 CAPITA)	SHARE OF POP	SHARE OF DEMAND	SHARE OF SUPPLY
North East	58,126	137,117	13,755	2.36	0.24	30%	31%	50%
South	13,177	14,696	1,026	1.12	0.08	7%	3%	4%
Florida	12,961	12,644	512	0.98	0.04	7%	3%	2%
Midwest	38,220	46,315	2,968	1.21	0.08	19%	10%	11%
Texas	22,417	45,638	3,851	2.04	0.17	11%	10%	14%
North West	6,785	21,087	529	3.11	0.08	3%	5%	2%
South West	40,476	146,208	3,850	3.61	0.10	21%	33%	14%
inter-regional	66,897*	21,072	768	0.31	0.01	2%	5%	3%
Total	196,256	444,777	27,259	2.27	0.14	100%	100%	100%

TABLE 16: TOP 200 MARKETS, REGIONAL POPULATION, INTERCITY TRIP DEMAND AND BUS SERVICES

*: Duplicated population for inter-regional trips' ends were excluded from the total.

REGION 1: NORTH EAST

The North East (Figure 5) region covers one third of the top 20 market's population and demand. This area, which includes the NEC survey study area described earlier in this report, contributes half of the top market's bus service supply. It includes four of the top 10 markets by long distance travel demand and all of the top 10 markets by bus service supply. With its 19 million population, New York metropolitan area is the core of the region, together with three other large metropolitan areas: Boston, Philadelphia, and Washington D.C. Its 53 markets construct a complicated regional network. Inside the region, a consistent pattern of long distance travel demand (Figure 6) and bus service supply (Figure 7) is observed. Table 17 lists all the market pairs with the sum of population (i.e., total population including both ends of the market), the two-way total number of long distance trips in 2008, and the number of bus services per week in 2014 as well as their rankings in the region. The trip rates shown in the table are calculated as the ratio of the two-way total number of long distance trips in 2008 and the total population, which is then divided by two to represent the annual per capita trip rate in one direction within the pair of CBSAs.

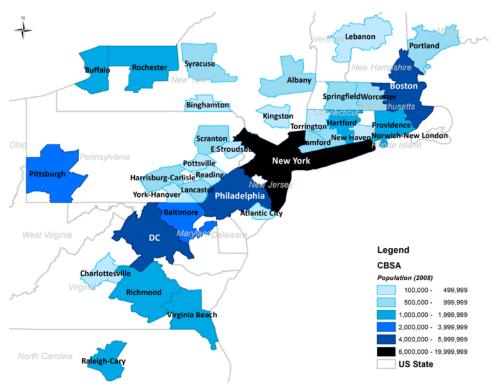
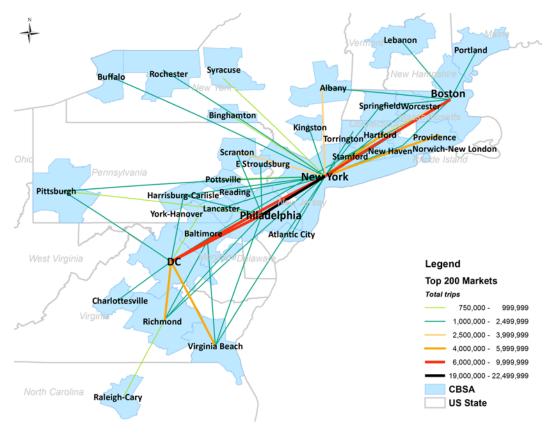


FIGURE 5: POPULATION IN THE NORTH EAST REGION

FIGURE 6: LONG DISTANCE TRIPS IN THE NORTH EAST REGION



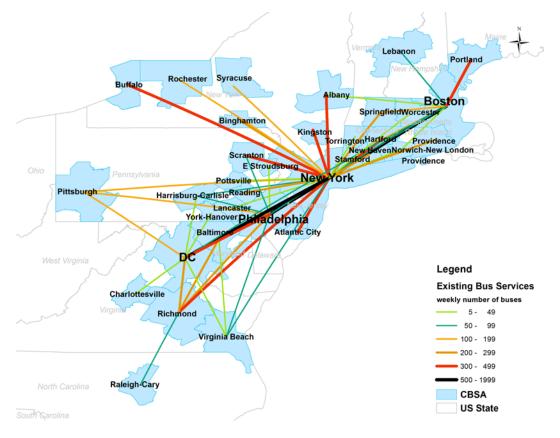


FIGURE 7: SCHEDULED BUS SERVICES IN THE NORTH EAST REGION

TABLE 17: TOP 200 MAKETS LOCATED IN THE NORTH EAST REGION

ORIGIN	DESTINATION	2008 TOTAL TRIPS (IN 1000)	SUM OF POP (IN 1000)	2008 TRIP RATES	2014 SERVICE PER WEEK	DEMAND RANKING	SUPPLY RANKING	POP RANKING
New York	Philadelphia	19,030	24,845	0.38	1,677	3	2	2
New York	DC	9,656	24,365	0.20	1,811	4	1	4
Philadelphia	DC	8,528	11,197	0.38	383	6	9	59
Boston	New York	8,493	23,530	0.18	1,088	7	4	5
Richmond	DC	5,647	6,584	0.43	253	13	25	100
Hartford	New York	5,417	20,197	0.13	363	15	11	13
Baltimore	New York	5,170	21,674	0.12	1,187	17	3	7
Virginia Beach	DC	4,156	7,016	0.30	28	22	165	96
New York	Providence	4,037	20,603	0.10	201	23	35	11
Baltimore	Philadelphia	3,939	8,506	0.23	386	24	8	78

ORIGIN	DESTINATION	2008 TOTAL TRIPS (IN 1000)	SUM OF POP (IN 1000)	2008 TRIP RATES	2014 SERVICE PER WEEK	DEMAND RANKING	SUPPLY RANKING	POP RANKING
Albany	New York	3,602	19,861	0.09	328	27	15	16
New York	Scranton	3,140	19,556	0.08	345	30	13	21
Boston	Hartford	3,042	5,713	0.27	158	31	45	134
New York	Springfield	2,324	19,694	0.06	204	41	33	19
New Haven	New York	1,848	19,853	0.05	294	56	18	17
Boston	Portland	1,768	5,037	0.18	468	60	5	146
Boston	Springfield	1,745	5,210	0.17	130	62	64	142
Boston	DC	1,738	9,881	0.09	92	63	88	71
Harrisburg	New York	1,734	19,538	0.04	70	66	108	22
New York	Richmond	1,714	20,232	0.04	318	68	16	12
New York	Virginia Beach	1,705	20,665	0.04	51	70	136	10
Harrisburg	DC	1,665	5,889	0.14	28	72	166	128
E Stroudsburg	Philadelphia	1,633	6,004	0.14	28	75	167	119
New York	Reading	1,583	19,410	0.04	274	82	23	24
Baltimore	Richmond	1,574	3,893	0.20	204	83	34	167
Buffalo	New York	1,560	20,131	0.04	367	86	10	14
Boston	Philadelphia	1,541	10,361	0.07	56	89	122	63
New York	Torrington	1,508	19,195	0.04	42	92	144	28
New York	New London	1,494	19,271	0.04	49	93	139	26
Harrisburg	Philadelphia	1,474	6,370	0.12	90	94	92	111
Philadelphia	Scranton	1,447	6,388	0.11	81	97	102	109
Charlottesville	DC	1,445	5,553	0.13	28	99	168	136
Boston	Lebanon	1,325	4,694	0.14	84	110	98	151
Lancaster	New York	1,315	19,509	0.03	38	112	154	23
Pittsburgh	DC	1,295	7,709	0.08	132	113	61	85
Atlantic City	New York	1,260	19,277	0.03	424	116	6	25
New York	Worcester	1,260	19,791	0.03	14	117	191	18

ORIGIN	DESTINATION	2008 TOTAL TRIPS (IN 1000)	SUM OF POP (IN 1000)	2008 TRIP RATES	2014 SERVICE PER WEEK	DEMAND RANKING	SUPPLY RANKING	POP RANKING
New York	Pittsburgh	1,213	21,358	0.03	162	121	44	8
Albany	Boston	1,142	5,377	0.11	48	134	141	139
Baltimore	Virginia Beach	1,076	4,325	0.12	28	144	169	163
New York	Rochester	1,069	20,041	0.03	146	146	50	15
Philadelphia	Virginia Beach	1,066	7,497	0.07	56	147	123	87
Boston	New Haven	1,058	5,369	0.10	60	149	119	140
Kingston	New York	1,031	19,188	0.03	400	152	7	29
Philadelphia	Richmond	1,028	7,064	0.07	277	155	21	95
Boston	New London	1,006	4,787	0.11	43	157	143	149
New York	Syracuse	908	19,651	0.02	177	167	38	20
New York	Pottsville	826	19,154	0.02	14	180	192	30
Boston	Stamford	791	5,418	0.07	42	186	145	138
Binghamton	New York	781	19,252	0.02	269	189	24	27
Raleigh-Cary	Richmond	777	2,314	0.17	98	190	84	190
Philadelphia	Pittsburgh	769	8,190	0.05	133	193	59	81
DC	York-Hanover	759	5,783	0.07	28	197	170	131

REGION 2: SOUTH

The South region (Figure 8) covers 7% of the population out of the top 200 markets, with a 3% share of the long distance travel demand and a 4% share of the bus service supply. Atlanta at the center of the region in terms of being at the hub of the defined markets within the region as shown in Figure 9 and Figure 10. Table 18 lists all the 11 markets in this region: population, total trips and bus services per week as well as the rankings.

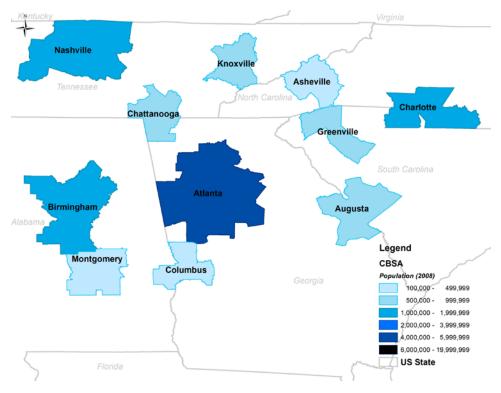
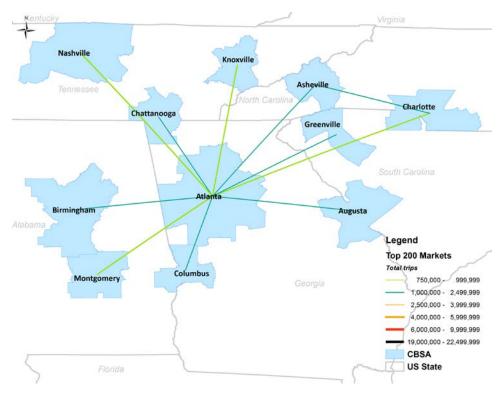


FIGURE 8: POPULATION IN THE SOUTH REGION





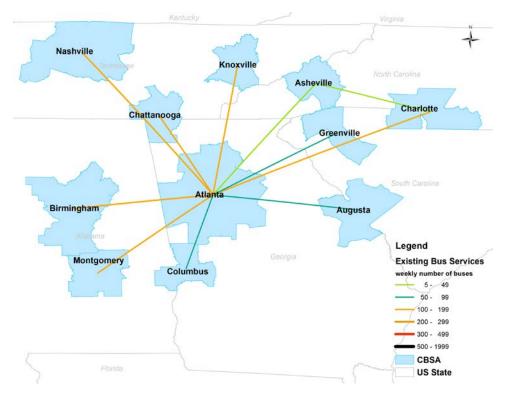


FIGURE 10: SCHEDULED BUS SERVICES IN THE SOUTH REGION

ORIGIN	DESTINATION	2008 TOTAL TRIPS (IN 1000)	SUM OF POP (IN 1000)	2008 TRIP RATES	2014 SERVICE PER WEEK	DEMAND RANKING	SUPPLY RANKING	POP RANKING
Atlanta	Chattanooga	2,162	5,895	0.18	199	46	36	127
Atlanta	Birmingham	2,068	6,494	0.16	116	49	72	106
Asheville	Atlanta	1,812	5,785	0.16	7	57	199	130
Atlanta	Columbus	1,471	5,664	0.13	63	95	115	135
Atlanta	Augusta	1,399	5,911	0.12	70	103	109	126
Atlanta	Greenville	1,211	6,001	0.10	56	123	124	120
Asheville	Charlotte	1,072	2,110	0.25	14	145	193	193
Atlanta	Nashville	991	6,927	0.07	136	159	57	97
Atlanta	Charlotte	956	7,078	0.07	135	162	58	94
Atlanta	Montgomery	785	5,742	0.07	104	187	80	132
Atlanta	Knoxville	769	6,067	0.06	126	194	65	117

TABLE 18: TOP 200 MARKETS LOCATED IN THE SOUTH REGION

REGION 3: FLORIDA

Florida covers 7% of the population and 3% of the total long distance trips in the top 200 markets. This demand is accommodated with a 2% share of the scheduled bus services. Miami, located at the southern end of the region, is largest population center (Figure 11). This region has a relatively simple pattern of market connections representing long distance travel demand (Figure 12) and bus service supply (Figure 13). Table 19 lists all of the eight markets in this region and tabulates population, total trips and bus services per week as well as the rankings.

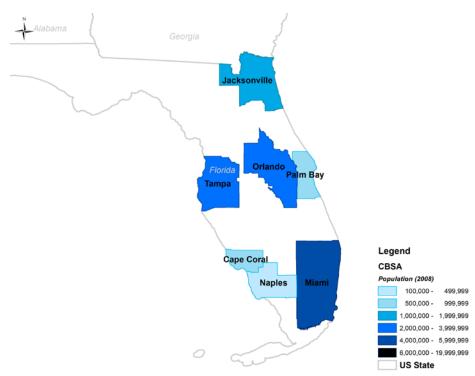
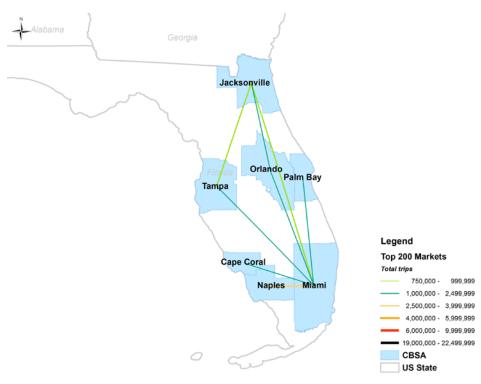


FIGURE 11: POPULATION IN THE FLORIDA REGION

FIGURE 12: LONG DISTANCE TRIPS IN THE FLORIDA REGION



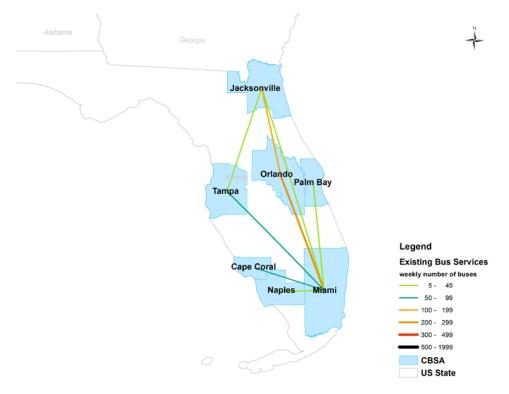


FIGURE 13: SCHEDULED BUS SERVICES IN THE FLORIDA REGION

TABLE 19: TOP 200 MARKETS LOCATED IN THE FLORIDA REGION

ORIGIN	DESTINATION	2008 TOTAL TRIPS (IN 1000)	SUM OF POP (IN 1000)	2008 TRIP RATES	2014 SERVICE PER WEEK	DEMAND RANKING		POP RANKING
Miami	Naples	2,647	5,730	0.23	31	35	164	133
Miami	Orlando	2,319	7,469	0.16	205	42	32	88
Miami	Tampa	1,767	8,149	0.11	62	61	118	82
Jacksonville	Orlando	1,594	3,368	0.24	101	81	82	178
Cape Coral	Miami	1,563	6,008	0.13	52	85	134	118
Miami	Palm Bay	1,029	5,951	0.09	28	154	171	123
Jacksonville	Tampa	883	4,047	0.11	21	170	189	165
Jacksonville	Miami	842	6,728	0.06	12	178	198	99

REGION 4: MIDWEST

The Midwest region covers 19% of the population and about 10% of the long distance travel demand and 11% of the bus service supply. Out of the 26 CBSAs that form the markets within the region, the Chicago

metropolitan area dominants in terms of population (Figure 14) and serves as the center of the long distance demand (Figure 15) and bus service supply (Figure 16) network. In addition, there are multiple connections to Minneapolis in Minnesota; and high connectivity within Ohio as well. Detail information on the 33 markets is shown in Table 20.

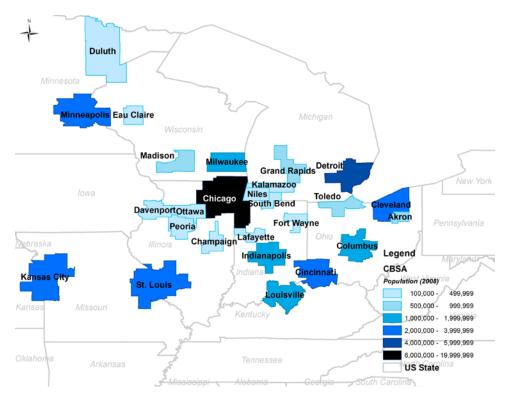


FIGURE 14: POPULATION IN THE MIDWEST REGION

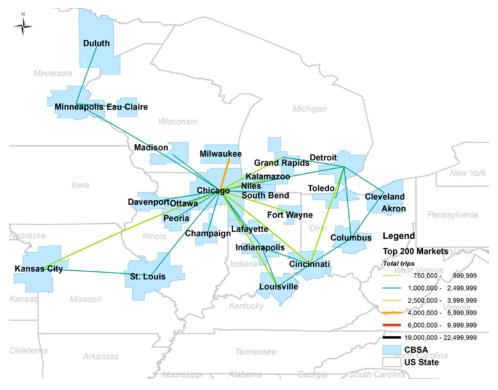
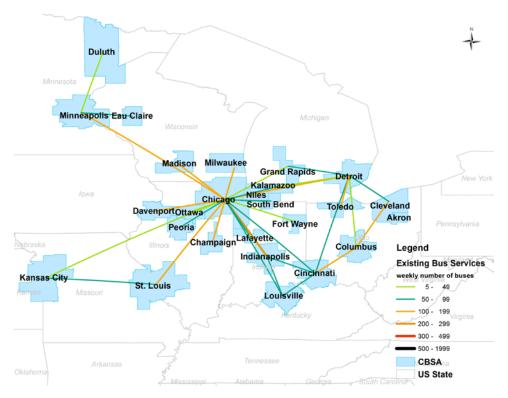


FIGURE 15: LONG DISTANCE TRIPS IN THE MIDWEST REGION





ChicagoDeChicagoIndChicagoStDetroitGrChicagoPeCincinnatiIndCincinnatiLoCincinnatiCo	ilwaukee			RATES	PER WEEK			RANKING
Chicago Ind Chicago St Detroit Gr Chicago Pe Cincinnati Ind Cincinnati Lo Cincinnati Co		5,198	11,119	0.23	158	16	46	60
Chicago St Detroit Gr Chicago Pe Cincinnati Inc Cincinnati Lo Cincinnati Co	etroit	2,220	13,995	0.08	139	45	53	41
Detroit Gr Chicago Pe Cincinnati Ind Cincinnati Lo Cincinnati Co	dianapolis	2,118	11,285	0.09	206	47	31	57
Chicago Pe Cincinnati Inc Cincinnati Lo Cincinnati Co	t. Louis	1,960	12,386	0.08	131	54	63	52
Cincinnati Inc Cincinnati Lo Cincinnati Co	rand Rapids	1,789	5,202	0.17	64	58	114	143
Cincinnati Lo Cincinnati Co	eoria	1,737	9,942	0.09	70	64	110	68
Cincinnati Co	dianapolis	1,699	3,871	0.22	76	71	106	168
	ouisville	1,652	3,400	0.24	77	73	104	176
Chicago Mi	olumbus	1,649	3,928	0.21	146	74	51	166
	inneapolis	1,613	12,800	0.06	137	77	56	51
Chicago Ma	adison	1,542	10,131	0.08	177	88	39	65
Cleveland De	etroit	1,437	6,513	0.11	84	101	99	103
Columbus De	etroit	1,394	6,198	0.11	28	104	172	112
Chicago Sc	outh Bend	1,289	9,886	0.07	63	114	116	70
Indianapolis Lo	ouisville	1,239	2,960	0.21	98	118	85	182
Detroit To	oledo	1,193	5,074	0.12	112	125	73	145
Chicago Da	avenport	1,172	9,947	0.06	105	129	75	67
Cleveland Co	olumbus	1,146	3,861	0.15	132	133	62	169
Chicago La	afayette	1,124	9,762	0.06	87	136	94	74
Champaign Ch	hicago	1,108	9,794	0.06	166	137	43	72
Kansas City St	t. Louis	1,089	4,819	0.11	70	143	111	148
Eau Claire Mi	inneapolis	1,065	3,389	0.16	52	148	135	177
Duluth Mi	inneapolis	1,034	3,504	0.15	42	151	146	173
Chicago Ot	ttawa	1,031	9,723	0.05	42	153	147	76
Chicago Gr	rand Rapids	980	10,346	0.05	36	160	156	64
Cincinnati De			6,580	0.07	56	161	125	101

ORIGIN	DESTINATION		SUM OF POP (IN 1000)	2008 TRIP RATES	2014 SERVICE PER WEEK	DEMAND RANKING	SUPPLY RANKING	POP RANKING
Chicago	Kalamazoo	952	9,893	0.05	78	163	103	69
Chicago	Cincinnati	951	11,725	0.04	76	164	107	54
Chicago	Kansas City	846	11,572	0.04	28	177	173	56
Chicago	Fort Wayne	815	9,981	0.04	28	181	174	66
Chicago	Louisville	782	10,814	0.04	98	188	86	61
Chicago	Niles	758	9,729	0.04	56	198	126	75
Detroit	Kalamazoo	752	4,749	0.08	50	199	137	150

REGION 5: TEXAS

The Texas region (Figure 17) covers 11% of the population and 10% of the top 200 market's demand for long distance travel. Its 31 markets form a multi-center network of demand around Dallas, Houston, Austin and San Antonio, within which Houston is the largest attraction for demand (Figure 18). A relatively similar pattern shows on the supply side as well as in Figure 19, with the Dallas metro area served by the highest number of bus services this time. Besides the North East region, this is the only other region that supply shares (14%) exceeds the share of demand, and notably includes two bus services ranked in the top 20 among all the 200 markets: Austin–Dallas (ranking 12), and Austin–San Antonio (ranking 20). Table 21 lists all of the market details and rankings.

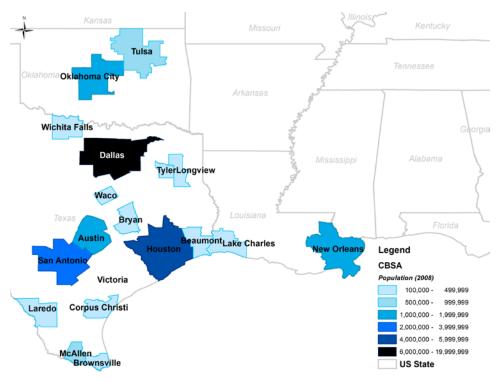
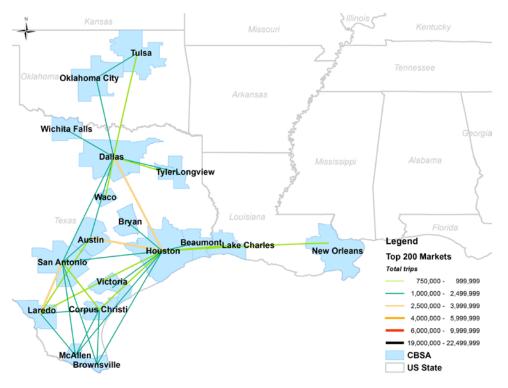


FIGURE 17: POPULATION IN THE TEXAS REGION

FIGURE 18: LONG DISTANCE TRIPS IN THE TEXAS REGION



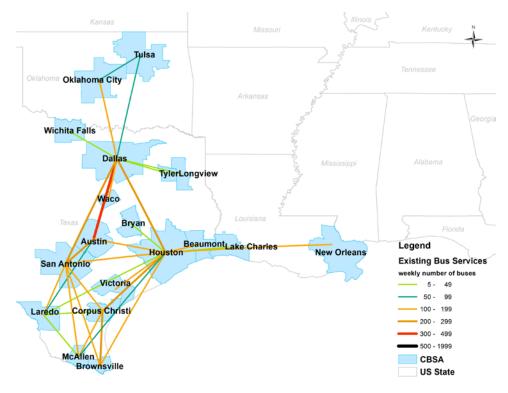


FIGURE 19: SCHEDULED BUS SERVICES IN THE TEXAS REGION

TABLE 21: TOP 200 MARKETS LOCATED IN THE TEXAS REGION

ORIGIN	DESTINATION	2008 TOTAL TRIPS (IN 1000)	SUM OF POP (IN 1000)	2008 TRIP RATES	2014 SERVICE PER WEEK	DEMAND RANKING	SUPPLY RANKING	POP RANKING
Dallas	Houston	3,635	12,028	0.15	224	26	28	53
Laredo	San Antonio	2,810	2,268	0.62	157	33	47	191
Austin	Houston	2,730	7,381	0.18	123	34	68	90
Houston	San Antonio	2,448	7,760	0.16	122	39	70	84
Bryan	Houston	2,389	5,936	0.20	14	40	194	124
Austin	Dallas	2,046	7,953	0.13	358	52	12	83
Beaumont	Houston	1,717	6,106	0.14	106	67	74	116
Dallas	Oklahoma City	1,712	7,506	0.11	105	69	76	86
Corpus Christi	McAllen	1,628	1,142	0.71	104	76	81	197
Brownsville	Corpus Christi	1,604	808	0.99	237	79	27	199
Dallas	San Antonio	1,595	8,331	0.10	277	80	22	80
McAllen	San Antonio	1,573	2,758	0.29	126	84	66	183

ORIGIN	DESTINATION	2008 TOTAL TRIPS (IN 1000)	SUM OF POP (IN 1000)	2008 TRIP RATES	2014 SERVICE PER WEEK	DEMAND RANKING	SUPPLY RANKING	POP RANKING
Houston	Victoria	1,446	5,842	0.12	139	98	54	129
Corpus Christi	Laredo	1,350	652	1.04	28	109	175	200
Dallas	Wichita Falls	1,271	6,447	0.10	28	115	176	108
Austin	San Antonio	1,236	3,684	0.17	280	119	20	170
Oklahoma City	Tulsa	1,213	2,122	0.29	91	122	89	192
Dallas	Longview	1,192	6,505	0.09	42	126	148	104
Brownsville	San Antonio	1,161	2,424	0.24	147	130	49	188
Brownsville	Houston	1,154	6,121	0.09	139	131	55	115
Laredo	McAllen	1,104	964	0.57	28	139	177	198
Houston	McAllen	1,055	6,455	0.08	83	150	101	107
Dallas	Tulsa	940	7,216	0.07	77	165	105	92
Houston	New Orleans	902	6,862	0.07	123	169	69	98
Corpus Christi	Houston	865	6,144	0.07	209	172	30	114
Corpus Christi	San Antonio	859	2,447	0.18	126	176	67	187
Dallas	Tyler	829	6,501	0.06	49	179	140	105
Houston	Laredo	814	5,965	0.07	35	182	157	122
Austin	Laredo	813	1,890	0.22	88	183	93	195
Houston	Lake Charles	775	5,921	0.07	42	191	149	125
Dallas	Waco	771	6,530	0.06	144	192	52	102

REGION 6: NORTH WEST

The North West region (Figure 20) has only 3% of the population but relative high per capita demand to make up to 5% of the total of the top 200 markets' demand for long distance travel. The highest demand between Seattle and Bellingham is most likely a result of cross board demand to Canada. The other high demand is between the two major metropolitan areas, Seattle and Portland as shown in Figure 21. The bus service supply is consistent with the long distance travel demand (Figure 22); albeit the region only has a 2% share out of the top 200 markets' scheduled bus service. Table 22 has the details for the eight markets in the region.

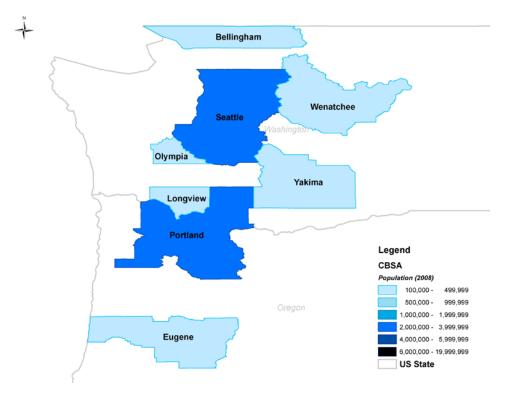
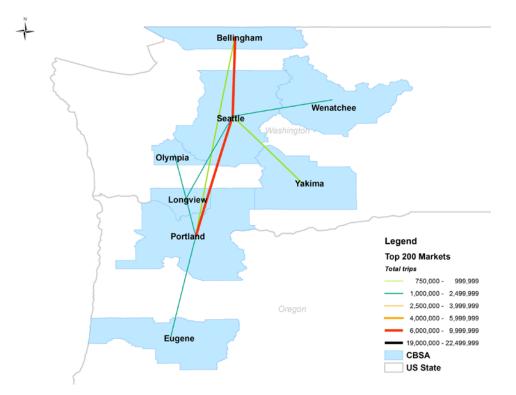


FIGURE 20: POPULATION IN THE NORTH WEST REGION

FIGURE 21: LONG DISTANCE TRIPS IN THE NORTH WEST REGION



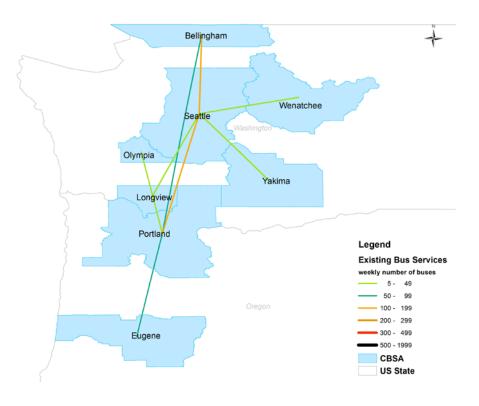


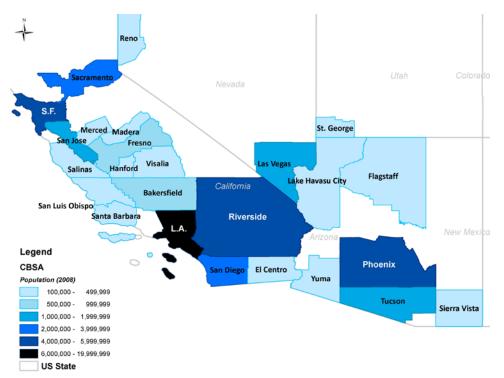
FIGURE 22: SCHEDULED BUS SERVICES IN THE NORTH WEST REGION

ORIGIN	DESTINATION	2008 TOTAL TRIPS (IN 1000)	SUM OF POP (IN 1000)	2008 TRIP RATES	2014 SERVICE PER WEEK	DEMAND RANKING	SUPPLY RANKING	POP RANKING
Bellingham	Seattle	6,708	3,541	0.95	105	9	77	172
Portland	Seattle	6,351	5,552	0.57	156	11	48	137
Eugene	Portland	2,012	2,554	0.39	87	53	95	184
Olympia	Portland	1,605	2,453	0.33	44	78	142	186
Seattle	Wenatchee	1,403	3,453	0.20	14	102	195	174
Longview	Seattle	1,231	3,446	0.18	37	120	155	175
Bellingham	Portland	915	2,404	0.19	58	166	120	189
Seattle	Yakima	862	3,579	0.12	28	175	178	171

REGION 7: SOUTH WEST

The South West region contains several of the largest metropolitan areas. Labeled as L.A. in Figure 23, Los Angeles metropolitan area has the second largest population among all CBSAs (12 million); followed by three

others in the four to six million range (Phoenix, S.F. for San Francesco, and Riverside) from north to east across the region. The South West region has the highest demand for long distance travel in various dimensions: it has largest market share (33%) among all the regions; its per capital trip rate (3.61) is nearly two times as high as the other regions (3.61 vs 1.92); and the average market size in terms of total long distance trips is also twice that in the other regions (3.45 million vs 1.85 million). It contains 10 of the top 20 markets in terms of long distance travel demand, and notably the top 2: Los Angeles—Riverside and Los Angeles— San Diego (Figure 24). This region has a relatively low bus service supply that totals up to a 14% share of all supply in the top 200 markets (Figure 25). Table 23 lists all the markets in the region with its population, demand and supply as well as the rankings.





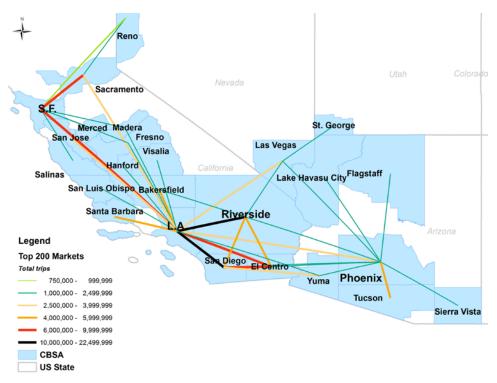
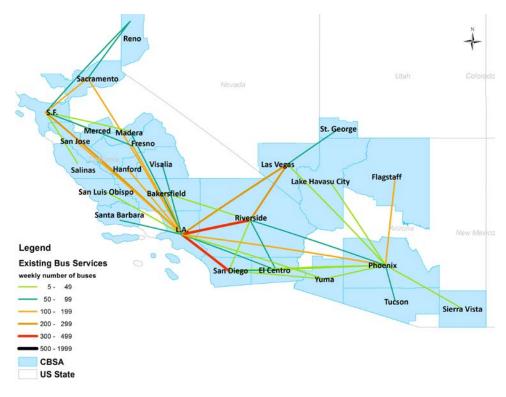


FIGURE 24: LONG DISTANCE TRIPS IN THE SOUTH WEST REGION

FIGURE 25: SCHEDULED BUS SERVICES IN THE SOUTH WEST REGION



ORIGIN	DESTINATION	2008 TOTAL TRIPS (IN 1000)	SUM OF POP (IN 1000)	2008 TRIP RATES	2014 SERVICE PER WEEK	DEMAND RANKING	SUPPLY RANKING	POP RANKING
Los Angeles	Riverside	22,480	16,989	0.66	338	1	14	33
Los Angeles	San Diego	21,107	15,874	0.66	312	2	17	34
El Centro	San Diego	8,933	3,165	1.41	56	5	127	180
Los Angeles	San Francisco	7,273	17,147	0.21	289	8	19	32
El Centro	Los Angeles	6,565	13,037	0.25	63	10	117	48
Sacramento	San Francisco	6,262	6,384	0.49	188	12	37	110
Bakersfield	Los Angeles	5,510	13,673	0.20	174	14	41	43
El Centro	Riverside	5,027	4,280	0.59	56	18	128	164
Los Angeles	Santa Barbara	4,513	13,278	0.17	56	19	129	45
Phoenix	Tucson	4,339	5,294	0.41	91	20	90	141
Riverside	San Diego	4,309	7,117	0.30	28	21	179	93
Las Vegas	Los Angeles	3,659	14,739	0.12	240	25	26	39
Los Angeles	Phoenix	3,237	17,155	0.09	105	29	78	31
Los Angeles	San Jose	3,006	14,692	0.10	176	32	40	40
Los Angeles	Sacramento	2,592	14,983	0.09	117	36	71	36
San Diego	Yuma	2,527	3,195	0.40	42	37	150	179
Los Angeles	Visalia	2,299	13,299	0.09	84	43	100	44
Phoenix	Yuma	2,292	4,476	0.26	35	44	158	156
Phoenix	Sierra Vista	2,106	4,411	0.24	28	48	180	160
Flagstaff	Phoenix	2,061	4,410	0.23	105	50	79	161
Salinas	San Francisco	2,049	4,683	0.22	42	51	151	153
El Centro	Phoenix	1,850	4,446	0.21	35	55	159	157
Phoenix	San Diego	1,768	7,283	0.12	35	59	160	91
Merced	San Francisco	1,737	4,521	0.19	35	65	161	154
Las Vegas	Phoenix	1,544	6,148	0.13	28	87	181	113
Los Angeles	Yuma	1,524	13,067	0.06	14	91	196	47

TABLE 23: TOP 200 MARKETS LOCATED IN THE SOUTH WEST REGION

ORIGIN	DESTINATION	2008 TOTAL TRIPS (IN 1000)	SUM OF POP (IN 1000)	2008 TRIP RATES	2014 SERVICE PER WEEK	DEMAND RANKING	SUPPLY RANKING	POP RANKING
Lake Havasu City	Phoenix	1,459	4,478	0.16	28	96	182	155
Hanford	Los Angeles	1,444	13,022	0.06	133	100	60	49
Los Angeles	San Luis Obispo	1,393	13,138	0.05	42	105	152	46
Fresno	Los Angeles	1,388	13,782	0.05	174	107	42	42
Fresno	San Francisco	1,321	5,184	0.13	58	111	121	144
Los Angeles	Madera	1,188	13,021	0.05	56	127	130	50
Reno	Sacramento	1,177	2,525	0.23	86	128	96	185
Las Vegas	Riverside	1,138	5,982	0.10	219	135	29	121
Las Vegas	St. George	1,106	2,003	0.28	56	138	131	194
Madera	San Francisco	1,096	4,423	0.12	35	141	162	159
Bakersfield	Riverside	1,025	4,916	0.10	14	156	197	147
Phoenix	Riverside	1,000	8,398	0.06	91	158	91	79
Reno	San Francisco	905	4,689	0.10	86	168	97	152

REGION 8: INTER-REGIONAL

The Inter-regional markets forms a small part of the total demand for long distance travel (5%) and of the bus service supply side (3%). The markets includes connections form the North East (particularly New York) to locations in the adjacent regions of the South and the Midwest (Figure 26 and Figure 27) while the highest level of bus service supply is associated with Atlanta. Detailed data describing the inter-regional markets is presented in Table 24.

FIGURE 26: LONG DISTANCE TRIPS AMONG REGIONS

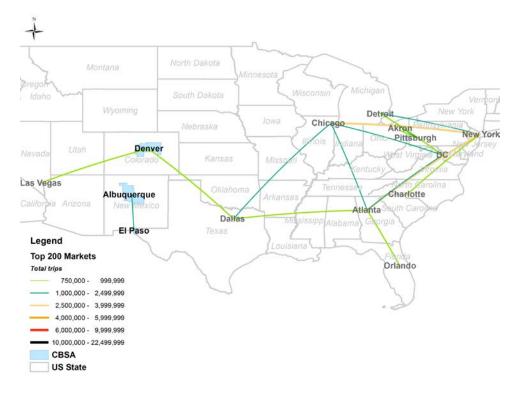
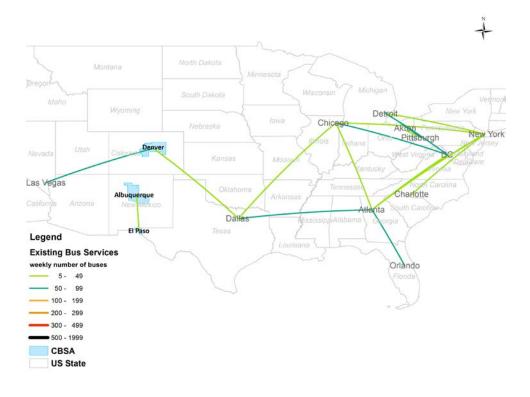


FIGURE 27: SCHEDULED BUS SERVICES AMONG REGIONS



ORIGIN	DESTINATION		SUM OF POP (IN 1000)	2008 TRIP RATES	2014 SERVICE PER WEEK	DEMAND RANKING		POP RANKING
Chicago	New York	3,242	28,576	0.06	28	28	183	1
Atlanta	New York	2,512	24,383	0.05	23	38	188	3
Chicago	DC	1,537	14,928	0.05	56	90	132	38
Atlanta	DC	1,389	10,734	0.06	28	106	184	62
Cleveland	Pittsburgh	1,379	4,439	0.16	70	108	112	158
Atlanta	Chicago	1,201	14,946	0.04	28	124	185	37
Detroit	New York	1,149	23,432	0.02	21	132	190	6
Chicago	Dallas	1,099	15,870	0.03	50	140	138	35
Albuquerque	El Paso	1,095	1,588	0.34	28	142	186	196
Atlanta	Dallas	877	11,676	0.04	98	171	87	55
Atlanta	Orlando	863	7,431	0.06	100	173	83	89
Denver	Las Vegas	863	4,372	0.10	56	174	133	162
Dallas	Denver	797	8,807	0.05	42	184	153	77
Detroit	DC	791	9,783	0.04	70	185	113	73
Charlotte	New York	766	20,709	0.02	28	195	187	9
Akron	Pittsburgh	764	3,050	0.13	35	196	163	181
Atlanta	Philadelphia	748	11,215	0.03	7	200	200	58

4.0 ESTIMATED BUS RIDERSHIP FOR THE TOP 200 MARKETS

The previous chapters have identified the top 200 CBSA pairs (or "markets") with the most intercity travel demand and, potentially, bus passenger trips. The previous chapters have also described the comprehensive intercity schedule database, and the intercity passenger origin-destination study conducted in the NEC. This chapter describes the estimation method for this project, which brings together the schedule database and the NEC study results to estimate the intercity bus passenger volume for each of the top 200 markets. This chapter begins with a description of the estimation data sources, continues to describe the estimation approach, followed by details on how the approach was applied, and concludes with the results of the application.

4.1 | ESTIMATION DATA SOURCES

The estimation process used three main data sources: the schedule database, the NEC study survey data, and a spatial dataset containing information about the CBSAs. This section describes each data source and how they relates to the estimation process.

SCHEDULE DATABASE

The schedule database is a complete description of intercity passenger bus service within the 48 states. It is in GTFS format, and contains the locations of all intercity bus stops and the times for intercity bus services. The core elements of this database are the *trips*, which are sequences of stop locations where the bus stops to pick up or drop off passengers. Each trip begins where the scheduled service starts, and ends where the scheduled service terminates. The database also contains the number of times each week that a trip is repeated (from 1 to 7), and the location of each stop, given as latitude/longitude coordinates.

The schedule database models variation in bus service over the course of a week (rather than a month, season, or year), because that is the primary way in which intercity bus service varies in time in the real world. The estimation process also works at a week level, and estimates the total passenger volumes on a per-week basis.

The database contains many trips, each with its own sequence of stops and weekly frequency. Together, these trips define the intercity bus services that are available to passengers for traveling within the 48 states. The estimation process uses this dataset to determine how each CBSA pair is served by intercity buses.

NORTHEAST ORIGIN-DESTINATION STUDY

The NEC study was conducted to estimate origin-destination intercity bus passenger volumes for cities in the NEC. As part of the study, surveys were distributed to passengers as they boarded busses in major northeast cities. The surveys asked the passengers to give their origin and destination locations for the bus trip. Because the study collected information on passenger origins and destinations directly, it is a good data source for calibrating and validating the estimation method for this project.

Using data from the surveys, the study estimated an origin-destination matrix. The passenger volumes from that matrix are used as a point of comparison for the results of this project's estimation process.

CORE-BASED STATISTCAL AREAS

CBSAs are a geographic unit defined by the US Census Bureau. Each CBSA is centered on an urban area defined as an area of dense urban development. CBSAs are the unit used to defined the top 200 markets: each market is a pair of CBSAs. CBSAs are also used in the estimation process to define a spatial framework.

4.2 | ESTIMATION APPROACH

The three estimation data sources are used to support the estimation approach. The approach is centered around the schedule database, and uses the northeast origin-destination matrix for calibration and validation, and the CBSAs for a spatial framework. This section describes the estimation approach. First, the approach is described for estimating ridership for a single bus trip. Then the approach is extended to the entire schedule database.

The principal idea behind the estimation approach is that the bus schedule data is a very important predictor of passenger trip volumes. The schedule data places an upper limit on the number of passenger trips between two locations: there cannot be more passengers than can fit on the busses running between the two locations. The schedule data also places a lower limit on the number of passengers between two locations: the busses cannot be too empty or else the service would not be profitable and would not exist. If we can choose an average per-bus passenger volume somewhere between the upper and lower limits, then the estimation becomes a simple multiplication of number of busses (from the schedule data) by the average passenger load.

ESTIMATION STEPS FOR ONE BUS TRIP

The estimation process for a single bus trip follows these four steps:

Simplify the Stop Sequence

The first step is to simplify the stop sequence of the trip. Each trip consists of a series of stops, each with a particular latitude/longitude location. To simplify the trip, the stop sequence is converted to a sequence of CBSAs by determining which CBSA each stop falls in. Stops that do not fall in any CBSA are dropped from the sequence. If a trip stops within one CBSA more than once in a row, the stops within the CBSA are collapsed to one representative stop. The result of the simplification is a sequence of CBSAs served by the trip.

Stops that fall outside of a CBSA are dropped because passenger trips are very unlikely to start or end outside of a CBSA – CBSAs cover much of the 48 states in terms of their population.

Multiple stops within one CBSA are collapsed into one stop because we do not expect many trips to start and end within the same CBSA, and because we are not interested in distinguishing stop locations for stops that are in the same CBSA.

Enumerate Possible OD Pairs

The second step is to enumerate all the OD pairs that the bus trip serves. For example, a bus trip that starts in city A, makes an intermediate stop in city B, and then terminates in city C has three possible OD pairs: A-B, A-C, and B-C. Each of these OD pairs is a trip that a passenger could take. The estimation process will need to estimate how many passengers travel on this bus for each of the three OD pairs.

Assign Scores to Pairs

The third step is to assign a score to each of the enumerated OD pairs. The score relates to the relative number of passengers each OD pair will have, compared to the other OD pairs. An OD pair with a high score will have more passenger trips than an OD pair with a low score.

For this project the scores are assigned based on the populations of the two CBSAs and the distance between them. The scores are then scaled to sum to one.

Add Passengers One At A Time

The fourth step is an iterative loop:

- 1. Randomly choose one of the OD pairs based on the relative scores.
- 2. Check if adding one passenger to the chosen OD pair would violate the bus capacity constraint. If so, return to step 1 to choose another OD pair. If not, go ahead and add the passenger to the OD pair.
- 3. Calculate the average passenger-miles per vehicle-miles. If the target average load has been achieved, then the process is done. If not, return to step 1 to choose another OD pair and add another passenger.

The bus capacity constraint is a number giving the maximum number of people that can ride on a bus at one time. The iterative process ensures that the bus is never over capacity.

The target load factor is an important parameter for the estimation method. It specifies the average personmiles per vehicle-mile for bus vehicle trips. The estimation results are directly related to this parameter; doubling the parameter value will double the total number of estimated passenger trips.

At the end of the iterative process, each OD pair will have an estimated passenger volume associated with it. The average passenger load derived from these OD estimates will be near the target load, and the derived maximum load will never be over the maximum capacity.

ESTIMATION FOR THE ENTIRE SCHEDULE DATABASE

The previous section detailed the estimation approach for a single bus trip. The approach is easily extended to the entire schedule database. First, all the trips in the database are simplified to form CBSA sequences (as described above). Then, trips with identical CBSA sequences are grouped together. Then the per-week trip frequencies are used to determine how many times each unique CBSA sequence occurs in a week, across the entire schedule database.

Next, the estimation process is applied to each unique CBSA sequence in exactly the same way, except the target load factor and the maximum capacity parameters are multiplied by the number of times the CBSA sequence occurs in a week.

At the end of the process, each unique CBSA sequence has its own set of OD passenger volume estimates. All the OD estimates are then aggregated by OD pair, to form a complete set of estimated of passenger volumes. The results are then tabulated, with one row for each estimate, and columns for origin, destination, and number of passengers.

DISCUSSION OF THE ESTIMATION APPROACH

The estimation approach has several benefits. First, it maintains explicit consistency with the schedule database, because the bus trips from the database form the basis for the estimation.

The estimation approach allows setting a target load factor, which is the average number of people on board the bus over the distance of its service. The target load factor allows calibrating the estimation process to match real-world observed loads. The estimation approach also allows setting a maximum bus capacity, which is the maximum number of people that can be on a bus at any time.

The approach also allows controlling the relative demand for different OD pairs through assigning different scores. This allows calibrating the model to match, for example, passenger trip distance distributions.

4.3 | APPLYING THE ESTIMATION APPROACH

The estimation approach was applied to the schedule database. As part of the implementation, the project team developed a scoring approach and estimated the target load factor, and then implemented the approach in a computer program. The final result was a comma-delimited text file giving the weekly passenger volume for each of the top 200 markets.

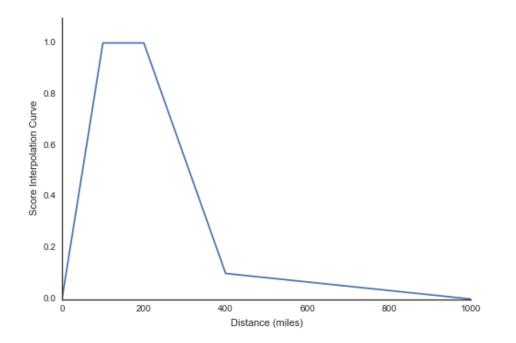
SCORING APPROACH

The scoring function takes the origin and destination locations (which are CBSA's), and returns a number between zero and one. The overall score is the product of two parts – one for the distance between the origin and destination, and one for the product of the population at the origin and population at the destination.

The distance part score is calculated by using a piecewise linear interpolation function that takes a distance in miles and returns a number between zero and one (Figure 28). The distance scoring function is highest between 100 and 200 miles, and is lower towards lower and higher distances. This reflects the fact that shorter trips are often served by other modes of ground transportation, and longer trips are often served by airplane.

The population part score is calculated by interpolating on the line defined by the points $(1.3 \times 10^8, 0)$ and $(2.4 \times 10^{14}, 1)$ using the product of the populations of the origin and destination CBSAs. The interpolation line is designed to give zero at the smallest possible CBSA combination, and one at the largest possible CBSA combination.

The overall score is the product of the two part scores. OD pairs that have a higher population will have a higher number of passenger trips assigned. In addition, OD pairs with a distance closer to the 100 - 200 miles range will have more passenger trips assigned.



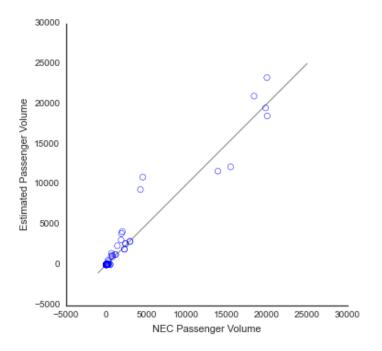
ESTIMATING THE TARGET LOAD FACTOR

The value for the target load factor parameter was estimated using results from the NEC study. Passenger volumes from the NEC study origin-destination matrix were compared to the results from this study's estimation method. The target load factor was then adjusted to minimize the differences between the two sets of results. The final estimate for the load factor was 23 passenger-miles per vehicle mile. This means that there is an average of 23 people on a bus over the distance of the bus route.

4.4 | VALIDATION

Using the estimated target load factor of 23 passenger-miles per vehicle mile, the estimation method was used to generate a final estimate of passenger OD volumes. To validate the estimation method, the final results were again compared to the results from the NEC study. The results are shown below in Figure 29. The diagonal line shows where the points would fall if they estimates matched perfectly.

The validation results show that the estimated passenger volumes generally follow the pattern and scale of the NEC passenger volumes, which gives us confidence in our scoring approach and target load factor estimate, and in our ability to apply the model to estimate volumes for the entire schedule database.



4.5 | RESULTS: BUS RIDERSHIP IN THE TOP 200 MARKETS

This section presents the estimated weekly bus ridership in 2014 for the top 200 markets, based on application of the approach described above. The ridership results are presented in units of person trips.

Overall, the top 200 markets were estimated to carry 382,437 passengers every week. The New York to Philadelphia market has the largest ridership, of 44,269 passengers while over 61.5% of the top 200 markets have a weekly ridership of less than 1,000. Figure 30 is a histogram of the estimated weekly ridership.

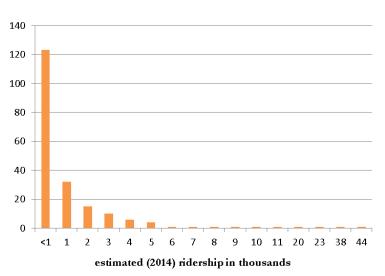
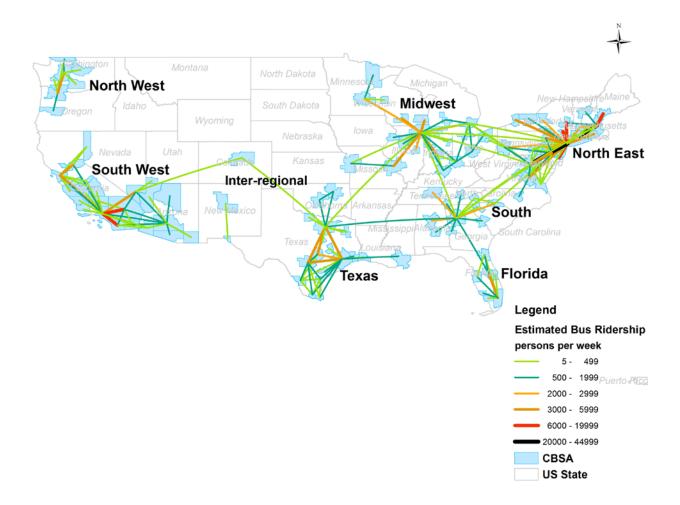


FIGURE 30: HISTOGRAM OF ESTIMATED WEEKLY BUS RIDERSHIP

Figure 31 show the top 200 markets' ridership using a similar color scheme to that used earlier in the report to present intercity travel demand. Not surprisingly, the spatial pattern of the ridership is highly consistently with that of the bus service supply. Table 25 provides some summary statistics at the regional level. The North East region, with the highest number of people per bus, has a 59% share of the total ridership. The Southwest, Midwest, and Texas are the second to fourth regions in terms of ridership, each with around 10% of the top 200 markets' bus trips.





REGION	N	MIN	MAX	MEAN	AVERAGE PERSONS PER BUS	SHARE
North East	53	8	44,269	4,262.79	16.50	59.1%
South	11	35	2856	1,266.09	13.57	3.6%
Florida	8	80	4131	996.88	15.58	2.1%
Midwest	33	58	3637	1,161.73	12.92	10.0%
Texas	31	10	5115	1,155.45	10.21	9.4%
North West	8	39	3679	779.13	12.10	1.6%
South West	39	20	10,719	1,281.26	12.98	13.1%
inter-regional	17	3	1074	249.94	5.53	1.1%
Total	200	3	44,269	1,912.19	14.25	100%

TABLE 25: SUMMARY OF ESTIMATED 2014 WEEKLY BUS RIDERSHIP BY REGION

Table 26 lists the top 20 markets. The Northeast region again has thirteen markets ranked in top 20. Other markets are located in the Southwest, Texas and Florida. The three markets with the highest number of persons per bus were in Los Angeles—Riverside, Sacramento—San Francisco, and Houston—San Antonio (which has 28.4 persons per bus in its 3,470 total ridership, ranking 26).

RIDERSHIP RANKING	ORIGIN	DESTINATION	REGION	ESTIMATED WEEKLY RIDERSHIP	AVERAGE PERSONS PER BUS
1	New York	Philadelphia	North East	44,269	26.4
2	New York	DC	North East	38,016	21.0
3	Boston	New York	North East	23,804	21.9
4	Baltimore	New York	North East	20,237	17.0
5	Boston	Portland	North East	11,236	24.0
6	Los Angeles	Riverside	South West	10,719	31.7
7	Atlantic City	New York	North East	9,618	22.7
8	Albany	New York	North East	8,213	25.0
9	Los Angeles	San Diego	South West	7,492	24.0
10	Kingston	New York	North East	6,595	16.5
11	Hartford	New York	North East	5,739	15.8
12	Sacramento	San Francisco	South West	5,353	28.5
13	Philadelphia	DC	North East	5,222	13.6
14	Austin	Dallas	Texas	5,115	14.3
15	New York	Scranton	North East	4,948	14.3
16	Buffalo	New York	North East	4,893	16.1
17	Dallas	San Antonio	Texas	4,546	16.4
18	Dallas	Houston	Texas	4,432	19.8
19	Los Angeles	San Francisco	North East	4,216	14.6
20	Miami	Orlando	Florida	4,131	20.2

TABLE 26: 20 MARKETS WITH HIGHEST RIDERSHIP IN 2014

The rest of the section presents the results for each region individually, including a figure showing the ridership distribution using the same color scale as maps shown earlier in this report..

The Northeast region is the largest and busiest bus market region. Figure 32 shows the region's 53 markets and demonstrate the ridership concentration around New York City metropolitan area. Outside connections to New York, Boston—Portland and Philadelphia—DC are estimated to be high demand ODs.

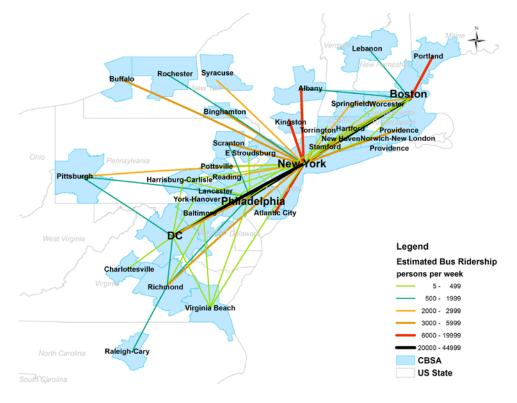


FIGURE 32 ESTIMATED 2014 WEEKLY BUS RIDERSHIP IN THE NORTH EAST REGION

The South region, as shown in Figure 33, with its main center in Atlanta, has a simple and relatively evenly distributed ridership pattern. The three markets with over 2000 passengers per week are from Atlanta to Charlotte, Birmingham and Nashville.

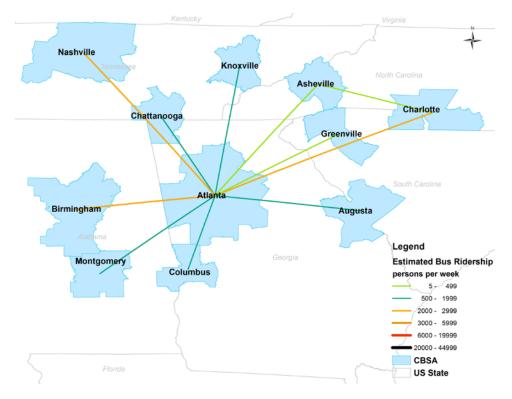


FIGURE 33: ESTIMATED 2014 WEEKLY BUS RIDERSHIP IN THE SOUTH REGION

The Florida region includes the Miami—Orlando pair which is ranked in the top 20 in terms of total ridership. The second largest in this region is the Jacksonville—Orlando pair. Figure 34 shows the results for all of the top 200 markets in the region.

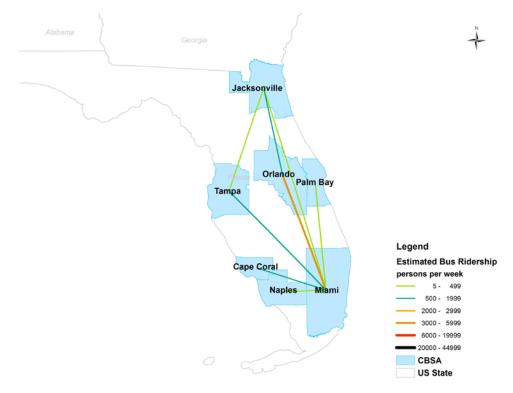


FIGURE 34: ESTIMATED 2014 WEEKLY BUS RIDERSHIP IN THE FLORIDA REGION

The Midwest region markets are clustered around the Chicago area. The top five markets are from Chicago to Indianapolis, St. Louis, Madison, Milwaukee and Detroit. The two pairs connecting St. Louis, Chicago—St. Louis and Kansas City—St. Louis, both have high persons per bus rates (around 26.5). Figure 35 show the top 5 markets in dark orange.

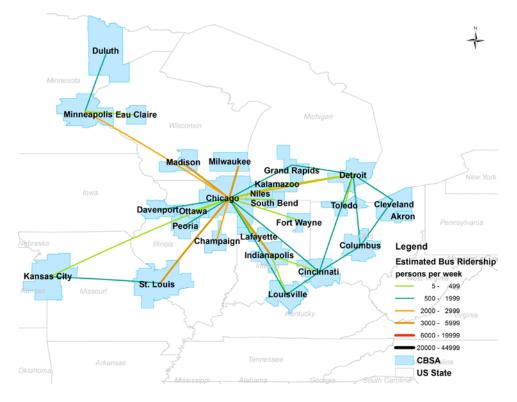


FIGURE 35: ESTIMATED 2014 WEEKLY BUS RIDERSHIP IN THE MIDWEST REGION

The Texas region includes several CBSAs with high populations. There are several markets – Austin—Dallas, Dallas—San Antonio, and Dallas—Houston – that are all included in the top 20 markets national wide. As shown in Figure 36, other significant pairs in the region include Houston—San Antonio and Austin—Houston.

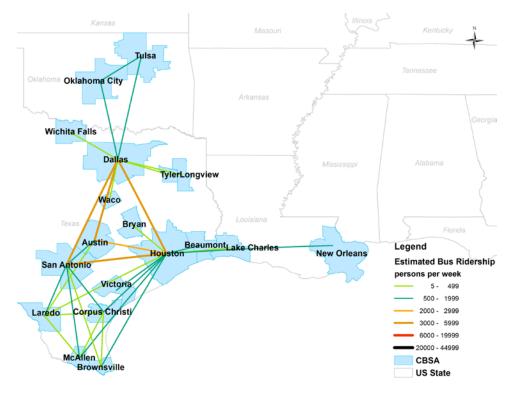


FIGURE 36: ESTIMATED 2014 WEEKLY BUS RIDERSHIP IN THE TEXAS REGION

The Northwest region is a relatively simple long distance bus market, with its highest ridership between Seattle and Portland. Other regional markets with relatively high passenger volumes are Eugene—Portland and Bellingham—Seattle. Figure 37 shows the ridership distribution of the region.

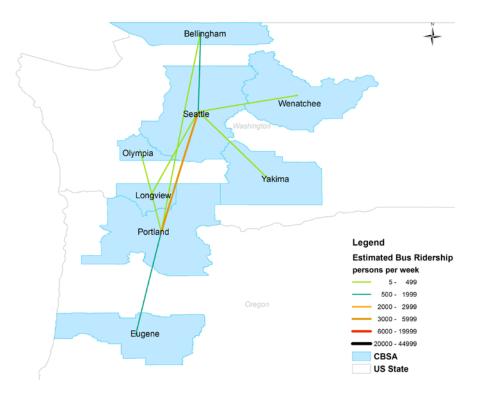


FIGURE 37: ESTIMATED 2014 WEEKLY BUS RIDERSHIP IN THE NORTH WEST REGION

The Southwest region's biggest markets are Los Angeles—Riverside and Los Angeles—San Diego. As shown in Figure 38, other large markets include Sacramento—San Francisco, Los Angeles—San Francisco and

Las Vegas-Los Angeles.

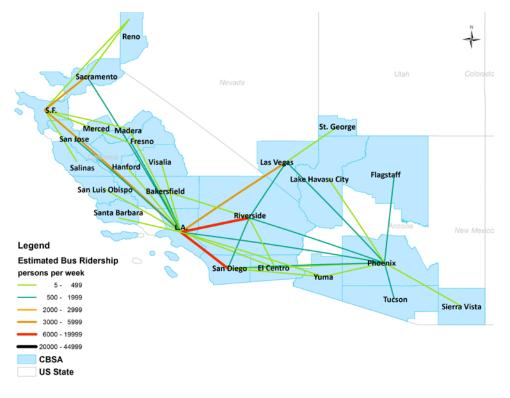
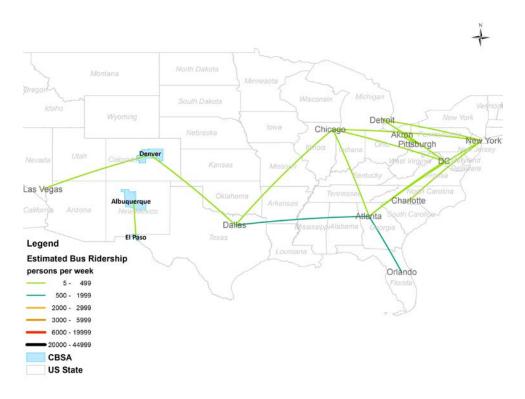


FIGURE 38: ESTIMATED 2014 WEEKLY BUS RIDERSHIP IN THE SOUTH WEST REGION

The Inter-regional markets have relatively low ridership and significantly lower rates of persons per trip. The largest market in this group is 1,074 persons per week between Atlanta and Orlando. As shown in Figure 39, although the northern part of the inter-regional grouping appears to have more connections, the ridership is relatively higher in the south part.

FIGURE 39: ESTIMATED 2014 WEEKLY BUS RIDERSHIP AMONG REGIONS



5.0 FUTURE YEAR (2040) BUS RIDERSHIP ESTIMATES

This chapter described the approach used in this project to estimate future ridership in 2040. The method, using population/employment growth factors, is taken from the earlier TAF Multimodal Interregional Passenger Travel Origin Destination Data project, where it was used to grow rail data from 2008 to 2040.

The approach used to estimate future ridership in 2040 is a relatively simple growth factor approach based on forecasts of population and employment. The approach embodies the assumption that bus ridership would increase to meet the demand for travel but would not go above that expected purely based on population and employment growth. Since the estimated ridership derived in Chapter 4 is for 2014, growth factor are based on forecasts of population and employment group between 2014 and 2040.

A growth factor for each OD pair was calculated using population and employment growths at both ends of the market:

$$GF_{ij} = \frac{P_{i,2040} + P_{j,2040} + E_{i,2040} + E_{j,2040}}{P_{i,2014} + P_{j,2014} + E_{i,2014} + E_{j,2014}}$$

GFij - Growth factor for CBSA OD ij

Pi,yyyy - population for CBSA i in year yyyy

Ei,yyyy - employment for CBSA i in year yyyy

The 2014 estimates and 2040 forecasts of population and employment by county were obtained from Woods and Poole CEDDS data.

Table 27 sums the population and employment at the regional level. From 2014 to 2040, the population grew from 205 million to 265 million in the 133 CBSAs that form the top 200 markets (28.8% increase) and the employment increases to 170 million in 2040 from 118 million in 2014 (43.6% increase).

REGION	NUMBER OF CBSAS	TOTAL POPULATION 2014 (IN 1000)	TOTAL POPULATION 2040 (IN 1000)	TOTAL EMPLOYMENT 2014 (IN 1000)	TOTAL EMPLOYMENT 2040 (IN 1000)
North East	34	58,154.7	68,370.0	34,660.6	46,574.5
South	11	14,235.5	20,523.2	8,155.9	12,581.3
Florida	7	14,223.5	20,108.8	7,713.7	12,089.8
Midwest	26	39,168.3	45,591.7	23,528.4	32,050.8
Texas	19	25,047.0	37,061.7	14,350.1	22,637.5
North West	8	7,382.1	10,306.2	4,368.6	6,553.0
South West	25	43,002.0	56,574.9	22,930.7	33,487.8
inter-regional	3	4,545.9	6,567.0	2,597.4	3,965.9
Total	133	205,759.1	265,103.6	118,305.5	169,940.6

TABLE 27 REGIONAL POPULATION AND EMPLOYMENT: 2014 AND 2040

The estimated 2040 weekly ridership sums to 491,152 with a 28% increase from the 2014 total. Table 28 summarizes the 2040 weekly ridership at the regional level and Table 29 lists the top 20 markets with their 2040 ridership, ranking and growth factors. Most pairs remained in the top 20 ranking from 2014 to 2040 with some position switches, with the only exception being Houston—San Antonio replacing Los Angeles—San Francisco.

REGION	MEAN GROWTH FACTOR	TOTAL ESTIMATED WEEKLYRIDERSHIP IN 2040	MEAN ESTIMATED WEEKLY RIDERSHIP IN 2040
North East	1.24	274,152	5,172.68
South	1.50	20,941	1,903.73
Florida	1.45	11,760	1,470.00
Midwest	1.22	46,715	1,415.61
Texas	1.55	55,979	1,805.77
North West	1.44	8,987	1,123.38
South West	1.37	66,537	1,706.08
inter-regional	1.34	6,081	357.71
Total	1.35	491,152	2,455.76

RIDERSHIP RANK IN 2040	ORIGIN	DESTINATION	REGION	GROWTH FACTOR	WEEKLY RIDERSHIP IN 2040	RIDERSHIP RANK IN 2014
1	New York	Philadelphia	North East	1.19	52,742	1
2	New York	DC	North East	1.24	47,055	2
3	Boston	New York	North East	1.20	28,498	3
4	Baltimore	New York	North East	1.21	24,434	4
5	Los Angeles	Riverside	South West	1.34	14,403	6
6	Boston	Portland	North East	1.24	13,948	5
7	Atlantic City	New York	North East	1.19	11,465	7
8	Albany	New York	North East	1.19	9,801	8
9	Los Angeles	San Diego	South West	1.28	9,570	9
10	Austin	Dallas	Texas	1.61	8,225	14
11	Kingston	New York	North East	1.19	7,846	10
12	Dallas	San Antonio	Texas	1.58	7,204	17
13	Dallas	Houston	Texas	1.56	6,924	18
14	Hartford	New York	North East	1.19	6,848	11
15	Philadelphia	DC	North East	1.30	6,793	13
16	Sacramento	San Francisco	South West	1.26	6,728	12
17	Miami	Orlando	Florida	1.47	6,072	20
18	New York	Scranton	North East	1.19	5,878	15
19	Buffalo	New York	North East	1.18	5,795	16
20	Houston	San Antonio	Texas	1.55	5,368	26

TABLE 29: 20 MARKETS WITH HIGHEST RIDERSHIP IN 2040

6.0 RECOMMEND APPROACH FOR FURTHER REFINEMENT AND UPDATING

One emphasis of the work described in this report was to develop a process that could be refined and updated. The final chapter of the report discusses aspects of this, such as the work undertaken to date to encode the process into software, and the recommendation of the project team on how to further refine the work beyond the scope of this project.

6.1 | REPEATABLE ANALYTICAL PROCESS

An important product of the research is a repeatable analytical process codified in software, which means that the base and future year ridership estimates can be updated as new schedule data are made available. The rapidly changing nature of the intercity bus market means that new carriers are entering the market and new services are being added very regularly. This means that any "current" estimates are a snapshot and, several months later, the service and resulting ridership in some or many markets might have changed.

The work that is described in this report has been produced using a set of Python code developed by the project team. Python (<u>https://www.python.org/</u>) is a programming language that is released under an open source license, making it freely usable and distributable, even for commercial use, which means that the software required to execute Python code is freely available.

There are two key components to the software developed as part of this project:

- 1. Schedule data development, comprised of several tools
 - a. Conversion of bus schedule data from Russell's Guide format to GTFS format
 - b. Automated collected of schedule data from select carrier websites and conversion GTFS data
 - c. Conversion of other tabular schedule data to GTFS format
- 2. Simulation of bus ridership, comprised of several tools
 - a. Processing of GTFS schedule data into simplified formats (i.e., CBSA sequences)
 - b. Simulation of ridership subject to constraints

Given revised inputs, such as a new round of schedule data from Russell's Guide or updated service from other carriers, the various elements of the software can be re-run to produce updated ridership estimates.

6.2 | RECOMMENDATIONS FOR REFINEMENT

While this project has focused primarily on using data from the NEC to understand intercity bus travel behavior and on collecting schedule data and estimating bus ridership in the top 200 intercity travel markets in the United States, the approach could be extended and refined in several ways.

- Extending the application to additional, smaller markets: while the top 200 intercity travel markets covers a significant part of the overall intercity travel market, a more complete enumeration of intercity bus travel would require the collection of schedule data for additional markets and the application of the ridership estimation method in those markets
- Collection of additional bus ridership calibration data: the estimates of bus ridership relied on the use of survey data in just the NEC. While this is the single largest intercity bus corridor in the United

States, it is unclear whether bus load factors in the remainder of the country are the same as those in the NEC. Additional targeted bus ridership counts and surveys could improve our understanding of bus travel behavior in different regions of the country and improve the reliability of bus ridership estimates.