FHWA-NJ-2016-001

Integration of Bus Stop Counts Data with Census Data for Improving Bus Service

FINAL REPORT
April 2016

Submitted by

Catherine T. Lawson, Ph.D.
Chair, Geography and Planning Department
Director, Albany Visualization and Informatics Lab
Director, Lewis Mumford Center
Director, Masters in Urban and Regional Planning (MRP) Program
AS 218 1400 Washington Avenue
Albany, New York 12222
(518) 442-4775

NJDOT Research Project Manager
Giri Venkiteela, PhD

In cooperation with

New Jersey Department of Transportation
Bureau of Research
And
U. S. Department of Transportation
Federal Highway Administration

DISCLAIMER STATEMENT

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No.	2.Government Accession No.	3. Recipient's Catalog No.
FHWA-NJ-2016-001		

4. Title and Subtitle	5. Report Date
Integration of Bus Stop Counts Data with Census Data for Improving Bus Service	April 2016
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
Dr. Catherine Lawson	
9. Performing Organization Name and Address	10. Work Unit No.
University at Albany 1400 Washington Ave. Albany, NY, 12222	11. Contract or Grant No.
12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
New Jersey Department of Transportation PO 600 Federal Highway Administration U.S. Department of Transportation	14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

This research project produced an open source transit market data visualization and analysis tool suite, The Bus Transit Market Analyst (BTMA), which contains user-friendly GIS mapping and data analytics tools, and state-of-the-art transit demand modeling microsimulation capabilities. BTMA combines both archived transit operations data (e.g., automatic fare-box data), and new open data resources, particularly GTFS and US Census.

17. Key Words	18. Distribution Statement
NJ TRANSIT, Bus Market Analyst, Microsimulation, Open Source, Open Data, Census, Demographics, ACS, CTPP, Open Trip Planner, GTFS, GTFS Editor, Regression Model, Fare box, Survey Data,	

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

Form DOT F 1700.7 (8-69)

ACKNOWLEDGEMENTS

The authors wish to thank Paul Thomas of New Jersey Department of Transportation, Janice Pepper, Tom Marchwinski, Fei Yang, Lyndsey Scofield, Matt Safer, Matt Ledger, and Aaron Reisner of NJ TRANSIT, the University Transportation Research Center, University at Albany Graduate Programmer David Vadney, and the dedicated Albany Visualization and Informatics Lab staff members without whom this project would not have been possible.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

The Integration of Bus Stop Counts Data with Census Data for Improving Bus Service research project produced an open source transit market data visualization and analysis tool suite, *Bus Transit Market Analyst* (heretofore referenced as *Transit Market Analyst)*.

The *Transit Market Analyst* combines the rich source of archived transit operations data (e.g., automatic farebox data), with new open data resources, particularly GTFS and US Census. It is web-based and open source, allowing for easy deployment and consistent non-proprietary upgrades.

The *Transit Market Analyst* was used to analyze three market areas and one case study scenario. It combines GIS mapping and data visualizations by using GTFS routes as backbones to define market areas. This allows for easy implementation of new market areas based on choosing routes from GTFS. The American Community Survey (ACS) Application Programming Interface (API) is used to display demographic information in dynamic maps and graphs. A set of interactive tools: a choropleth map, a bar graph and a table, allow users to interrogate census information from the market area. Additionally, the tool offers Census Transportation Planning Products (CTPP) data in a similar set of map and table tools.

The *Transit Market Analyst* also has a set of open source tools for analyzing and visualizing machine-readable automatic farebox data and ridership survey data. The survey tools include a set of interactive graphs and maps that allow users to filter by route, demographics, and customer data.

Lastly, *Transit Market Analysts* contains a state of the art microsimulation transit demand modeling tool that uses Open Trip Planner as a microsimulation routing "engine." With GTFS as a geo-spatial backbone, demographic information is used to generate origin/destination trip tables and are algorithmically plotted as "bus riders" throughout the market area census tracts. The bus riders are then microsimulated through Open Trip Planner.

All of the market area models underestimate full day ridership, despite often over-estimating peak-time ridership. This is expected due to the microsimulation model algorithm only accounting for work trips. It also points to how peak ridership behaves differently from market area to market area. Farebox data indicates fairly steady ridership throughout the day with a more concentrated AM Peak than PM Peak.

The microsimulation modeling algorithm requires additional research to better predict bus-route level ridership. There are a number of approaches that could be explored for improving origin/destination location generation, such as using parcel data, point-based establishment and employment data, and expanded survey data.

BACKGROUND

Problem Statement

Providing a safe, efficient, and cost-effective bus network is a daunting challenge, requiring reliable data and high quality planning tools. Recent products made available by the United States Census Bureau, including the American Community 5-year Series and the 2010 decennial Census, create a golden opportunity for planners and researchers to refine the empirical basis for their population-based decisions. By delving into the demographics and characteristics of each of its markets, the New Jersey Department of Transportation (NJDOT) and New Jersey Transit (NJ TRANSIT) will better understand these characteristics, improving the lives of its stakeholders.

NJDOT sought guidance from a qualified team of fellow researchers and planners to assist them in identifying the changing transit ridership trends present in the 2010 Census. Dr. Catherine Lawson's research team at the Albany Visualization And Informatics Lab (AVAIL), from the State University of New York at Albany, used a data science approach to conduct demographic, economic and transportation analytics for small-medium sized market areas in the state of New Jersey. AVAIL developed a software tool suite that uses technology to aid in transit planning by mapping and visualizing various datasets and by developing demand models that put these data to work for the benefit of the State's people and economy. This demonstrates the commitment of NJDOT to achieve the tech-forward ideas set forth in its Long-Range Transportation plan and the Strategic Plan of the United States Department of Transportation (USDOT).

Research Objectives

The AVAIL team specializes in visualization and informatics, creating web-based visualization tools that allow for cooperative analysis. Starting in 2013, AVAIL was tasked with developing and using web-based visualization software to analyze NJTransit ridership trends. The final data visualization and informatics tool suite, capable of offering an enhanced perspective on NJTransit's own transportation assets, is designed to:

- Identify the key demographic factors that influence transit ridership;
- Assess the social, economic and systemic determinants that exist within the 2010 census;
- Develop solutions that provide persistent competitive advantage for NJTransit; and
- Provide the tools for continued success subsequent to the completion of this research project.

Introduction

In order to yield a more robust model for transit ridership, AVAIL employed a data-centric approach to the research project, utilizing a wide variety of data, and working closely with NJTransit to utilize their knowledge and experience when applying data analysis hypotheses to a broader New Jersey population. NJTransit already makes its schedules, fares and routes available publicly through its developer support program in the General Transit Feed Specification (GTFS) format. This format is database ready and allows for cross-reference with a number of publically available and agency produced machine readable datasets. AVAIL incorporated the following data sources to provide the most complete picture of transit stakeholder characteristics possible.

- American Community Survey (ACS);
- 2010 Census Data;
- Census Transportation Planning Products (CTPP);
- Longitudinal Employment Household Dynamics (LEHD) Longitudinal Origin Destination Employment Statistics (LODES);
- NJTransit farebox data;
- NJTransit survey data; and
- Available Land-Use Data.

AVAIL utilizes data science in developing transit market planning tools. Data science combines elements of computer science with domain specified techniques for analyzing data. Additionally, AVAIL uses open source software, and whenever possible, open data. The problem of proprietary software plagues transportation planning. Notes Daniel Sun et al. (2011) in *Development of Web-Based Transit Trip-Planning System Based on Service-Oriented Architecture* "The majority of transit trip planners exist as proprietary systems based on particular vendor products. With the incorporation of more functional components, system maintenance and regular transit information updates become burdensome tasks for transit agencies. In addition, the proprietary nature of the systems makes it difficult to take advantage of the rapid advancement of geospatial information and web technologies." Open source software, in contrast, has source code that is available for modification or enhancement by anyone. This openness provides opportunities for additional progress by teams with new ideas, while providing feedback on these features and improvements to the original software creators. Open sourcing research products allow planning agencies to make updates to the software either in house or through a third party and to receive the benefits of all future updates as they are made by other agencies.

Similar to open source software, open data can be used and distributed freely by anyone. Using open source to build public good data systems has broad reaching benefits for everyone.

Summary of the Literature Review

Uses of Open Data and Open Source Software in Transit Planning

One of the first uses of data science to assist transit planning began in 2005, when the Tri-County Metropolitan Transportation District of Oregon (Tri-Met), partnered with Google, to develop an open data scheduling strategy. Their efforts resulted in the creation of the General Transit Feed Specifications (GTFS). This open data approach made a unique contribution with the generation of static schedule information (e.g., stop location, route geometrics, and stop times) in a standard format (see https://developers.google.com/transit/gtfs/). By providing universal access to GTFS (including the instructions for creating it via the internet), a community of GTFS producers and consumers demonstrated the value of this open data route planning specification by drafting a variety of planning applications both for users and for agencies as noted by Antrim and Barbeau (no date),

- Trip planning and maps– applications that assist a transit customer in planning a trip from one location to another using public transportation;
- Ridesharing – applications that assist people in connecting with potential ridesharing matches;
- Timetable creation – create a printed list of the agency's schedule in a timetable format;
- Mobile applications –applications for mobile devices that provide transit information;
- Data visualization – applications that provide graphic visualizations of transit routes, stops, and schedule data;
- Accessibility – applications that assist transit riders with disabilities in using public transportation
- Planning analysis – applications that assist transit professionals in assessing the current or planned transit network;
- Interactive Voice Response (IVR) – applications that provide transit information over the phone via an automated speech recognition system;
- Real-time transit information – applications that use GTFS data along with a real-time information source to provide estimated arrival information to transit riders It should be noted that only a subset of all applications that use GTFS data are presented in this;

AVAIL uses GTFS as a spatial backbone in developing its web-based tool-suite. This allows transit planners to utilize for transit planning and analyses, the same data being developed to guide bus riders. AVAIL also uses GTFS to connect to Census Population, Transportation and Economic Demographics, agency datasets (farebox and survey), and Open Trip Planner for microsimulation.

Summary of the Work Performed

The Research Product: Transit Market Analyst

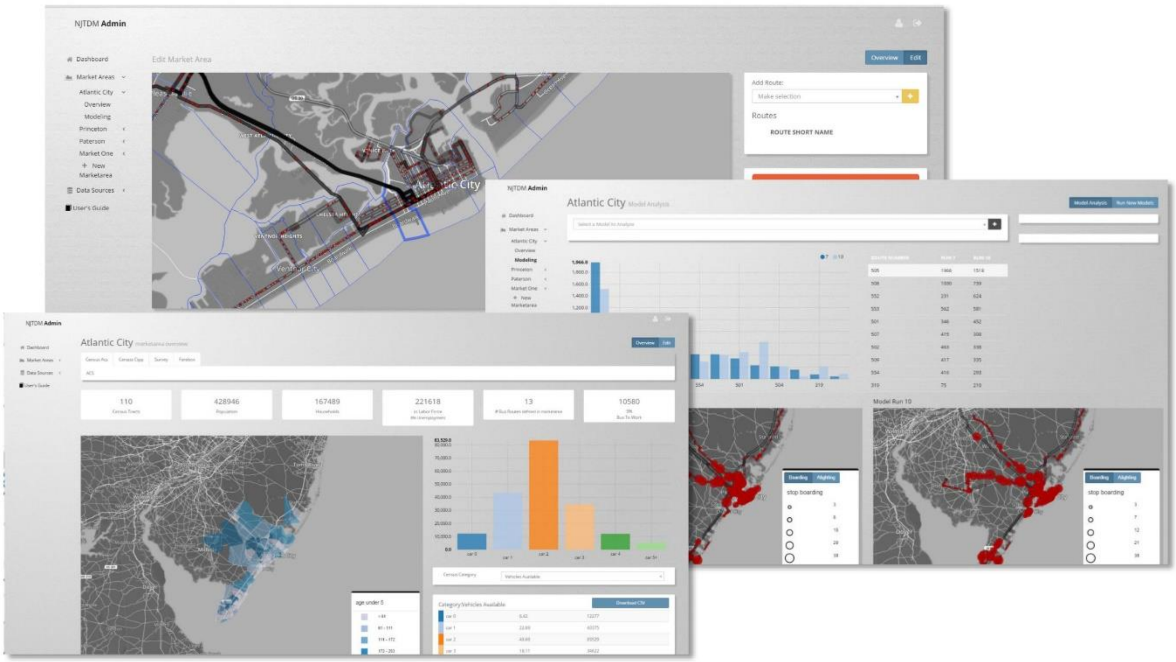


Figure 1. Example Screenshots of Transit Market Analyst

The Integration of Bus Stop Counts Data with Census Data for Improving Bus Service research project has produced as a final deliverable, an open source transit market data visualization and analysis tool suite, *Transit Market Analyst*.

The Transit Market Analyst combines the rich source of archived transit operations data (e.g., automatic farebox data), with new open data resources, particularly GTFS and US Census. It is web-based and open source, allowing for easy deployment and consistent non-proprietary upgrades. The Transit Market Analyst employs leaflet (http://leafletjs.com/) and D3.js (http://d3js.org/), both open source software, to create interactive maps organized by census tract geographies. This collection of tools and methodologies are intended to allow planners to assess changing transit demand in customizable market areas defined simply by GTFS routes and census geographies. The web-tool aggregates a number of data sets which are universally available in the US, such as the American Community Survey (ACS), Census Transportation Planning Products (CTPP) and The Longitudinal Employment and Household Dynamics (LEHD) survey, with data generated by transit agencies like GTFS and ridership surveys. These data sets are then run through an algorithm to approximate public transportation ridership. Custom developed software combined with Open Trip Planner is then used to microsimulate ridership in a given market area. The collection of tools and methodologies together, illuminate dynamics of

Figure 2. Open Trip Planner Logo

public transportation ridership in a given area and allow planners to investigate various market scenarios as well as write, edit and save GTFS.

GTFS Tools

The Transit Market Analyst combines GIS mapping and data visualizations by using GTFS routes as backbones to define market areas. This allows for easy implementation of new market areas based on choosing routes from GTFS. As GTFS routes are added to a market area the web-tool automatically chooses census tracts that contain bus stops on the GTFS routes. Additional census tracts can be easily added to the market area via pointing and clicking on the map.

Figure 3. Transit Market Analyst, Market Area Editing Tools

The GTFS routes that define the market area are also included on the maps for reference or as filters for some of the various data visualizations.

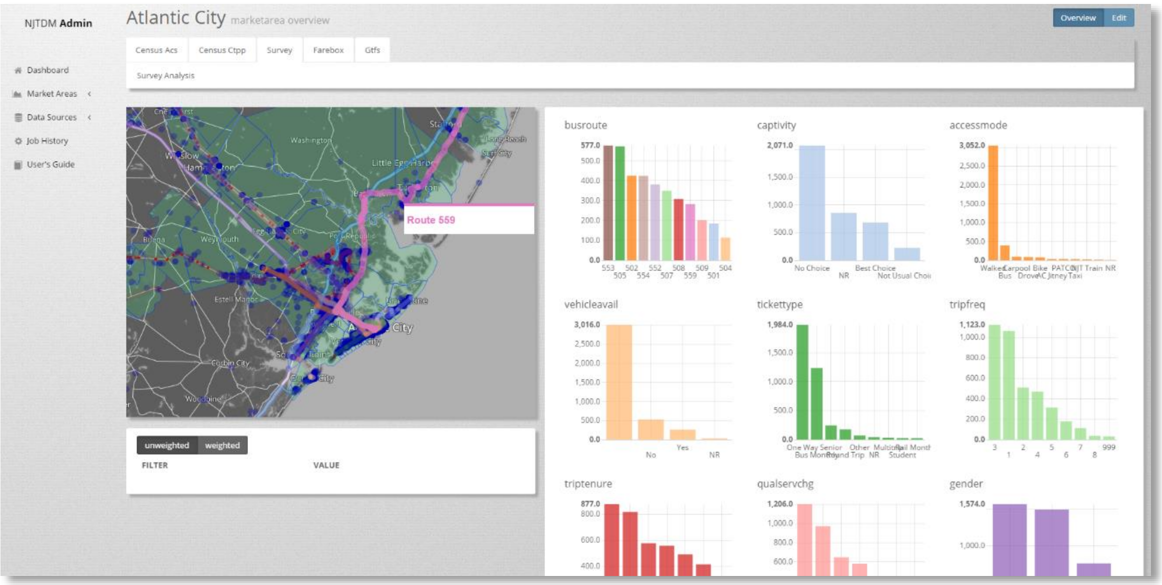


Figure 4. Transit Market Analyst, Survey Tools

With GTFS as the spatial tool for defining market areas, NJTransit and AVAIL found it critical for *Transit Market Analyst* users to be able to manipulate GTFS within the tool suite. The resulting tools are capable of adding and removing bus stops from a GTFS route via the mapping interface. A routing engine is employed to re-route the GTFS based on the user modifications. A scheduling interface is connected to the map for scheduling routes and trips. When a schedule and a route are ready, the GTFS can be saved and exported or employed as the GTFS for a market area.

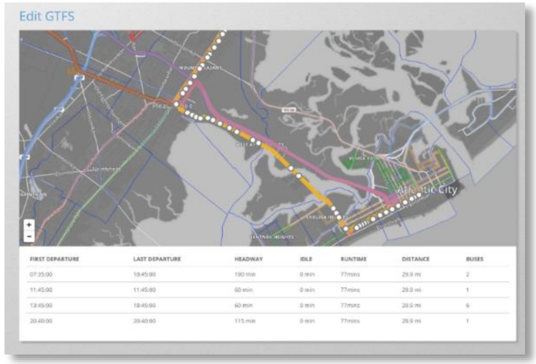


Figure 5. Transit Market Analyst, GTFS Editing Tools

Demographics Tools

The American Community Survey (ACS) Application Programming Interface (API) is used to display demographic information in dynamic maps and graphs. A set of interactive tools: a choropleth map, a bar graph and a table, that allow users to interrogate census information from the market area. Additionally, the tool offers Census Transportation Planning Products (CTPP) data in a similar set of map and table tools. The CTPP tools provides information from the census tracts on where bus riders live and work, identifying origin/destination flows between census tracts. The ACS and

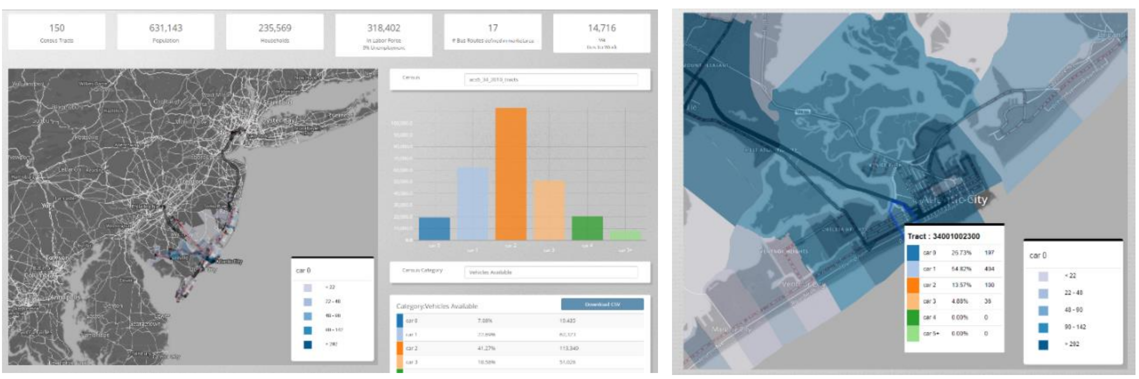


Figure 6. Transit Market Analyst, Demographic Analysis Tools

CTPP visualizations (maps, graphs and tables) are exportable from Transit Market Analyst by svg (image), csv, geojson and shapefile for use in other applications.

Farebox and Survey Visualization Tools

The Transit Market Analyst has a set of open source tools for analyzing and visualizing machine-readable automatic farebox data and ridership survey data. The survey tools include a set of interactive graphs and maps that allow users to filter by route, demographics, and customer data. The farebox tools also include interactive graphs and map that are filterable by time of day and farezone. The farebox tool also allows users to create average days based on user-selected date ranges. Left unfiltered the farebox tools show an average of all uploaded farebox data. A user can choose one day or a collection of days and can export the filtered farebox data (along with farezone filters) for use elsewhere in the site or as a csv file. Farebox and Survey data are used to “validate” the outcome of the microsimulation demand models.

Microsimulation Transit Market Demand Modeling using Open Trip Planner.

Transit Market Analysts contains a state of the art microsimulation transit demand modeling tool that uses Open Trip Planner (OTP) as a microsimulation routing "engine." With GTFS as a geo-spatial backbone, demographic information is used to generate origin/destination (O/D) trip tables.

Each row in an O/D table is treated as a "bus rider." Each bus rider is algorithmically plotted throughout the market area census tracts, placed in close proximity to bus-stops. The bus riders are then microsimulated through OTP.

Each bus rider takes a trip based on the GTFS schedule, as if they were real bus riders using their smartphones to navigate their way to work on the bus. OTP returns the three fastest-by-travel-time routes from the origin point to the destination point by departure time. One of those routes is chosen at random.

The bus ride is then microsimulated and the details are saved as legs and trips. This is repeated for each trip from the trip table and the output is collected together into data visualizations of bus stop O/Ds and route ridership data by time. These model outputs are then analyzed in comparison to farebox and survey data through a series of interactive maps and graphs to validate the microsimulated model.

OTP generates data for each leg of a micro-simulated rider's journey, including time (duration), distance, and route ids of the GTFS routes and stop locations utilized for each leg of the trip.

Table 1. Leg Data Output from Open Trip Planner Microsimulation

Block _id	Trip_id	Mode	Duration (sec)	Distance (sec)	Route	Route _id	GTFS_ trip_id	On_ stop_id	Off_stop_id	Dest_ id
121	705941	BUS	1050000	8069	502	133	18512	41985	38583	2438341
121	705941	BUS	786000	4464	555	134	18530	38583	205	2438342
121	705942	BUS	198000	1992	507	136	19206	205	42737	2438343

The Transit Market Analyst then aggregates each leg of a rider's trip into data for one trip.

Table 2. Aggregated Trip Data Output from Open Trip Planner Microsimulation

Run_ id	Start_ time	End_ time	Duration (sec)	Transit_ Time (sec)	Wait_ Time (sec)	Walk_ dist	Walk_ time	From_lat	From_lon	To_lat	To_long
981	10:26	10:36	1428000	1044	329	76.120	68	40.936	-74.184	40.918	-74.185
982	10:26	11:06	1428000	1399	286	76.120	68	40.936	-74.184	40.912	-74.176
983	10:26	11:16	1428000	1599	409	76.120	68	40.936	-74.184	41.345	-74.189

Trip Table Generation

In order to plot riders throughout the market area for microsimulation, Transit Market Analyst builds Origin/Destination trip tables based on CTPP census tract flows modified by a regression equation specified with ACS variables that are highly correlated with bus to work (public transportation ridership to work). What follows is a step by step explanation of trip table development

CTPP Models

Trip tables based solely on CTPP data are not high performing models but they are simple models and they are the first step in understanding the process of building trip tables for Open Trip Planner microsimulation demand modeling.

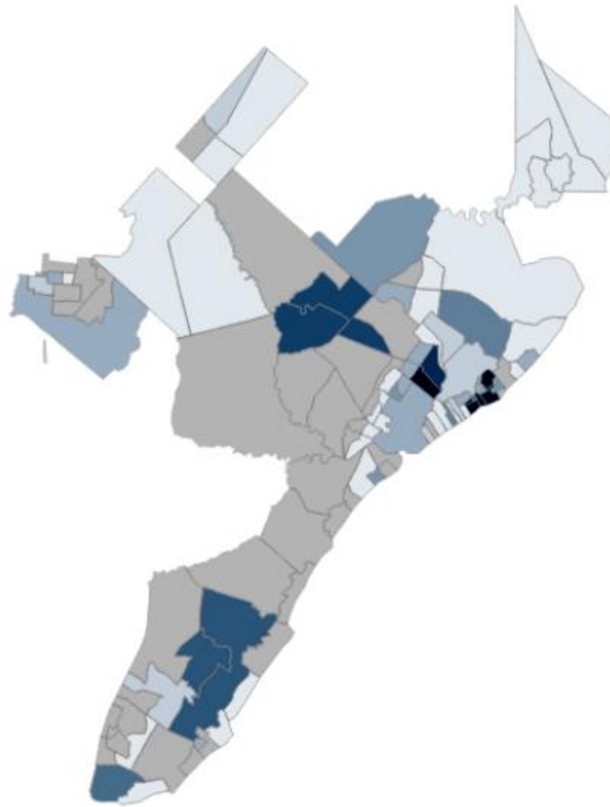


Figure 7. Transit Market Analyst, Choropleth Map of Atlantic City Trip Table

The CTPP model pulls origin (home) and destination (work) information for bus riders directly from CTPP for AM peak ridership. To generate PM Peak travel times the AM travelers have their Origins and Destinations reversed and their travel times (travel times from ACS) are pushed forward 8 hours from their AM Peak travel times. Those indicating PM travel to work are also introduced.

Market Area Specific Regression Models: Atlantic City Example

AVAIL specified a regression model for Atlantic City that has an R-Squared of 0.61. The model uses the dependent variable of bus-to-work in the initial regression model and keeps only the independent variables which are statistically significant and contribute to increases in the R-Squared and Adjusted R-Squared. NJTransit staff recommended using only ACS variables that could be easily estimated for future forecasting even when there were higher performing variables in the regression equation. The higher performing variables are not included in the final models. For example, persons earning a bachelor's degree, while statistically significant, would be very difficult to ascertain.

Microsimulation Flow Chart

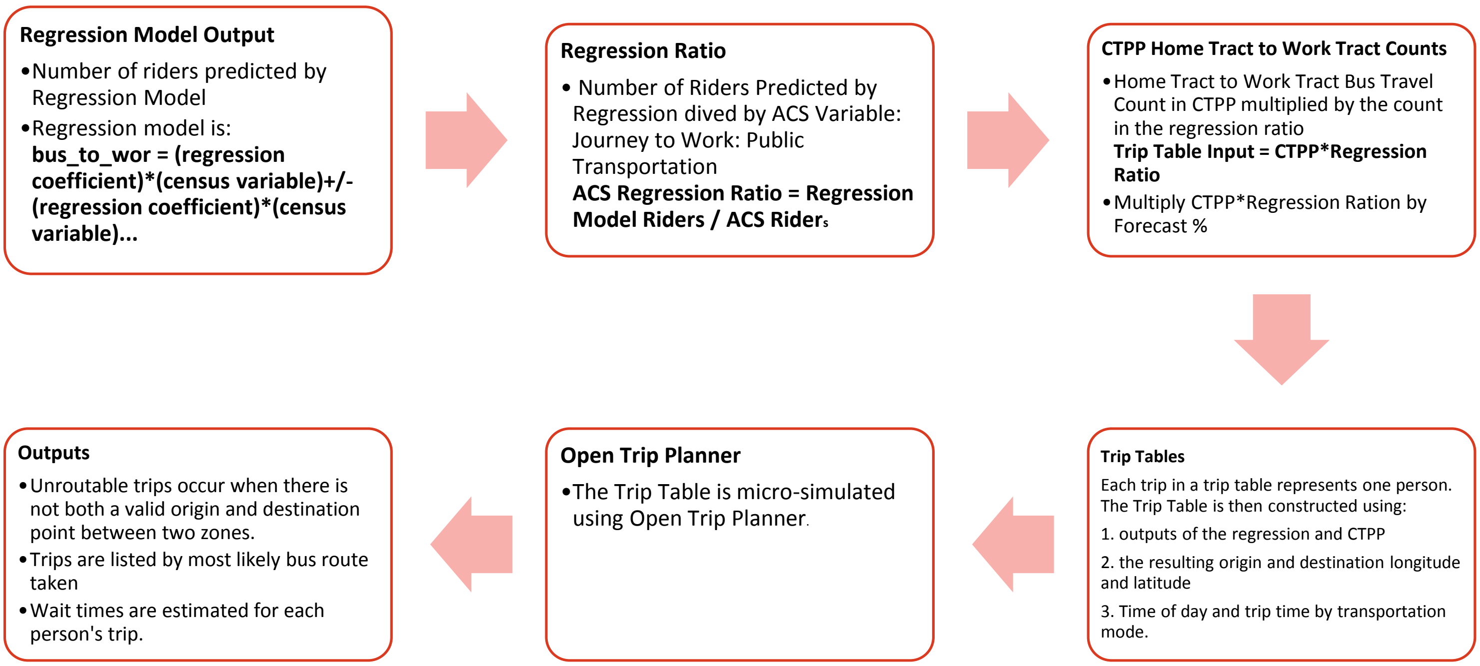


Figure 8. Flow Chart: Implementing Regression Models for Microsimulation

Regression Equation: Case Study – Atlantic City

Table 3. Atlantic City, Regression Model Development

	Atlantic City1	Atlantic City2	Atlantic City4	Atlantic City5
Dependent Variable	bus_to_wor	bus_to_wor	bus_to_wor	bus_to_wor
Constant	14.8622	-36.62	-36.82*	-41.505**
car_0_hous	0.42**	.36**	.28**	.230**
arts		.015**	.14**	.163**
emp_den			.02**	.019**
bachelor's			.01*	
R Sq.	0.32	0.50	0.54	0.61
N	110	110	110	110

** T-value >2.5 and P-value <.05. * T-value >2.5 or P-value <.05.

This equation uses the following three variables:

- bus_to_wor = Journey to Work by Public Transportation by Bus or Trolley Bus
- car_0_hous = Households, Zero Vehicles Available
- arts = Employment in Arts Sector
- emp_den = Employment/Area

The recommended final specification is:

bus_to_wor = – 41.505 + (0.230 x (car_0_hous)) + (0.163 x (arts)) + (0.019 x (emp_den))

<u>Implementing Regression Models For Microsimulation: Case Study – Atlantic City Census Tract 34001012200</u>

To predict ridership in our microsimulation, we use the following process for each tract in the market area. To show this process on a smaller scale, we will show a Regression Model example using a single census tract, 34001012200 and the Atlantic City Regression Equation. This census tract has 379 individuals indicating bus-to-work as their mode of transportation.

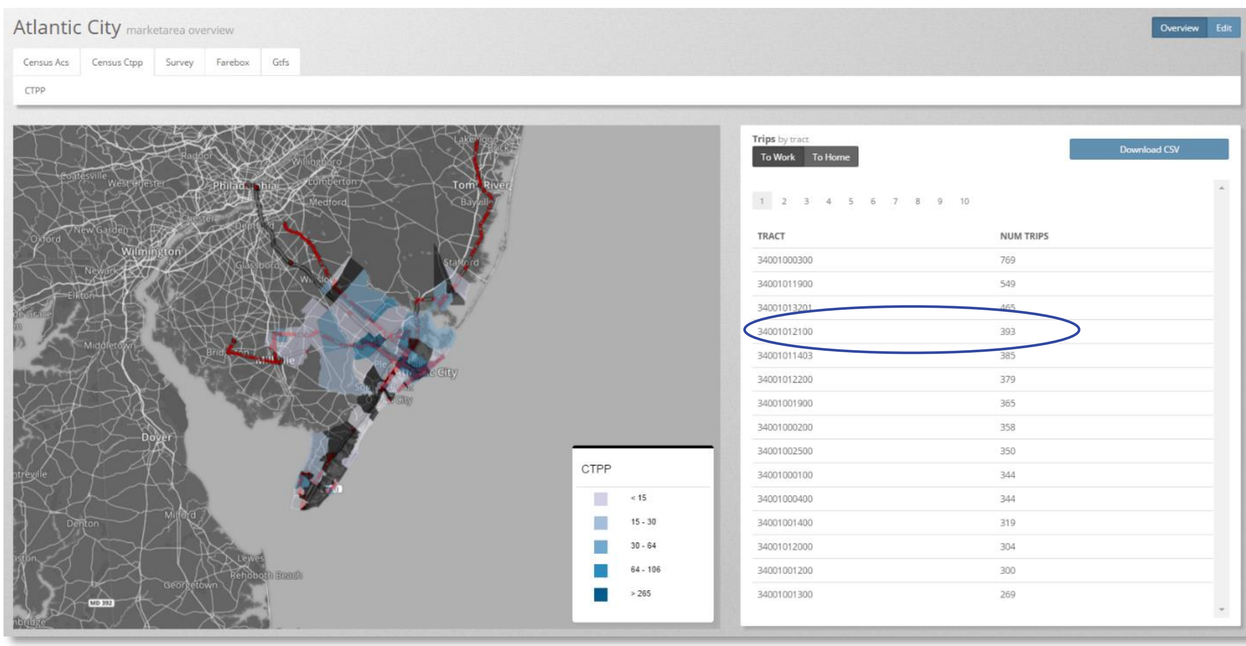


Figure 9. Transit Market Analyst, CTPP Flows Map

Find the number of riders predicted by the regression model for the tract.
To find the number of riders AVAIL has developed an algorithm that pulls specific census data for each census tract into the regression model. In this example, we will collect the variables for the Atlantic City Regression Equation and the corresponding data for census tract 34001012200. One can access this information using the Admin Tool by accessing the overview map and scrolling over the desired census tract.

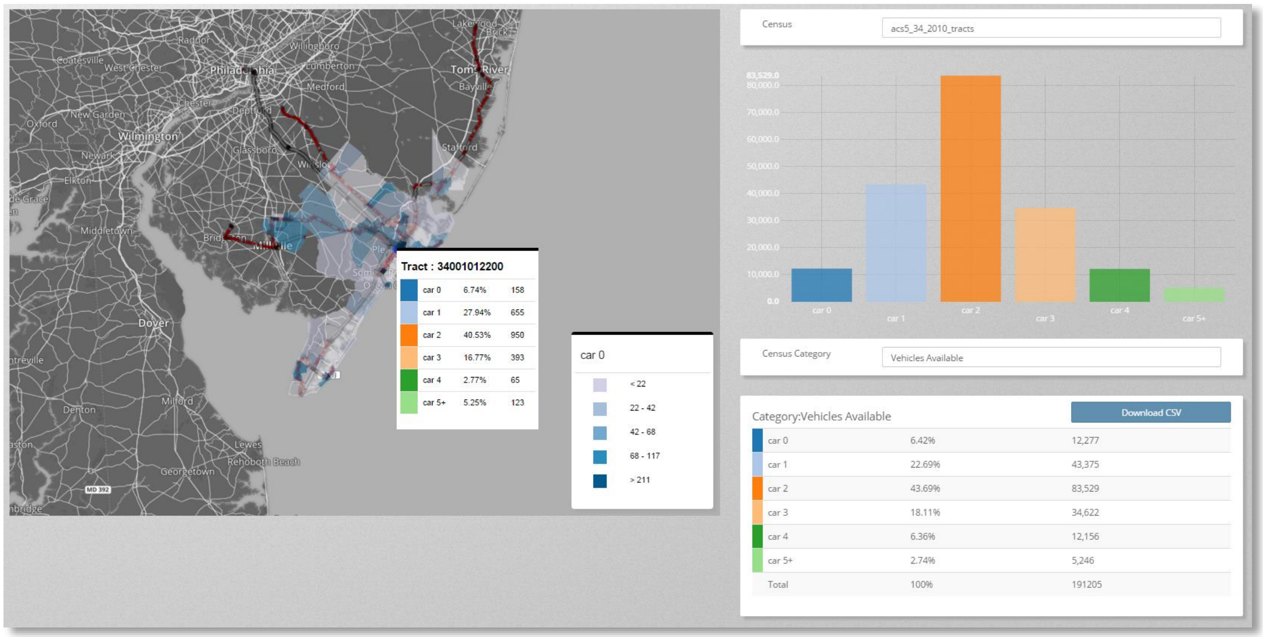


Figure 10. Census Tract ACS Variables

The trip table generating algorithm gathers the census tract data for each regression model variable from the ACS 5-Year dataset.

Table 4. Trip Table Algorithm, Census Tract Variables

Equation Variable	Description	Census Category	Amount in Census Tract 34001012200
bus_to_wor	Journey to Work by Public Transportation by Bus or Trolley Bus	Journey To Work	388
car_0	Households, Zero Vehicles Available	Household	196
arts	Employment in the Arts Sector	Labor Force	991
emp_den	Employment/Area	Total Employment/ Total Area	2251

The Regression Model Equation
In Atlantic City census tract 34001012200, the Atlantic City Regression Model is run as follows

bus_to_wor = – 41.505 + (0.230 x (car_0_hous)) + (0.163 x (arts)) + (0.019 x (emp_den))

When filled with census data for the tract:

bus_to_wor = – 41.505 + (0.230 x (196)) + (0.163 x (991)) + (0.019 x (2251))

The number of riders in census tract 34001012200 predicted by the Atlantic City Regression is 208.

ACS Regression Ratio
We take the number of riders predicted by the regression, and divide that by ACS variable Journey to Work: Public Transportation/Bus for the tract, giving us the ratio of predicted riders to census counted riders.

Regression Model Riders / ACS Riders = ACS Regression Ratio

208/388=0.54

CTPP Home Tract to Work Tract Counts
Then, for each home tract to work tract bus travel count in the CTPP, we multiply the count by the ACS Regression Ratio to find the ridership numbers for input into the trip table.

Trip Table Input = CTPP*Regression Ratio

All resulting trips are added to our trip table to be simulated by the modeling software.
The resulting trip table shows the number of bus trips from the origin point (census tract 3400101220) to each corresponding work census tract.

Table 5. Trip Table Output

Riders from Home Tract 34001012200		
Work Tract	Riders	Trip Table Output = CTPP*Regression Ratio (0.54)
34001002400	160	86
34001002300	60	32
34001001400	60	32
34001011900	25	14
34001000400	25	14
34001001100	20	11
34001013201	15	8
34001013302	10	5
34001011702	4	2
Total	379	205

This process is repeated for every census tract in the Market Area.

Generating AM and PM Bus-Riders

To generate AM Peak travel, the trip algorithm removes the percentage of bus-riders whose travel times are indicated as PM bus to work from the ACS travel time to work variable. PM Peak bus-riders in the trip table include those indicating PM travel to work. Additionally, the PM ridership includes the AM bus-to-work riders, their Origins and Destinations reversed (from work to home) and pushed forward 8 hours from their AM Peak travel times (to model the 8 hour work day).

Scenario Modeling (Forecasting)

The microsimulation demand modeling tools include customizable inputs for forecasting future conditions at the regional level (e.g., increases or decreases in population and employment), at the route level (e.g., increase or decreases at the census tract level), and at the stop level (e.g., using specific land use changes within walking distance of the stop)

When forecasts are applied to a census tract during model building, those percentages are applied to the trip table as the very last step before microsimulation. If a -10% population growth was applied to tract 34001012200, for instance, the ridership numbers for each of the work tracts would be reduced by 10%. The results for a -10% growth in tract 34001012200 is illustrated in Table 4.

Table 6. Forecasted Ridership Trip Table Output

Riders from Home Tract 34001012200			
Work Tract	Riders	Trip Table Output = CTPP*Regression Ratio (0.54)	Forecasted Ridership (- 10%)
34001002400	160	86	77.8
34001002300	60	32	29.2
34001001400	60	32	29.2
34001011900	25	14	12.2
34001000400	25	14	12.2
34001001100	20	11	9.7
34001013201	15	8	7.3
34001013302	10	5	4.9
34001011702	4	2	1.9
Total	379	205	184.2

Model Validation Tools

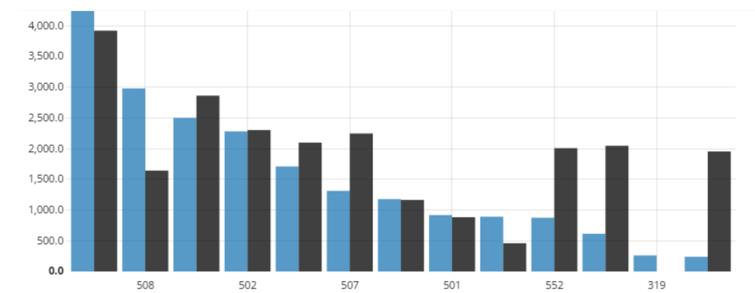
The visualization suite includes tools for comparing microsimulation models to farebox and survey results.

Demand Model Analysis

The Transit Market Analyst has a set of demand model analysis features. The model analysis tools compare models to farebox data as well as to other models. Below are a set of example graphs and tables comparing Atlantic City Regression Model outputs to farebox data from July 2013. The Transit Market Analyst tools for demand modeling and model analysis currently have hard-wired defaults for AM and PM Peak times.

In figures 6 and 8 below, the blue bars represent modeled bus ridership, the black bars represent actual farebox data from July 2013. The farebox data is an average of all the data in the tool.

In Figure 4, 5, 7 and 8 the stacked bar graphs show ridership by time of day for each route in the market area. Figures 4, 5 and 6 display the full day model compared with full day farebox. Figures 7, 8 and 9 display the route level data associated with PM Peak (3:00PM-7:00PM) of both the model and farebox.



AM	PM	
6 - 10	15 - 19	
8045	7357	20304
124.93%	113.45%	113.45%

ROUTE NUMBER	COLOR KEY	RUN 108	RUN FBOX
505		4541	3921
508		2980	1643
553		2501	2863
502		2281	2303
554		1710	2099
507		1313	2247
509		1178	1165
501		918	884
504		892	460
552		875	2008
559		614	2047
319		261	0
551		240	1954
		20304	23594

Figure 11. Regression Model Output and farebox Route Comparison graph and table filtered by PM Peak

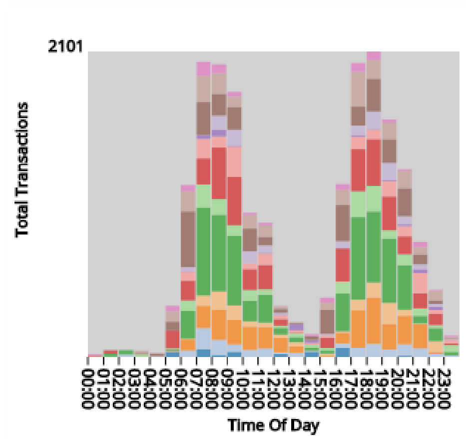


Figure 12. Regression Model Output by Route and Time of Day

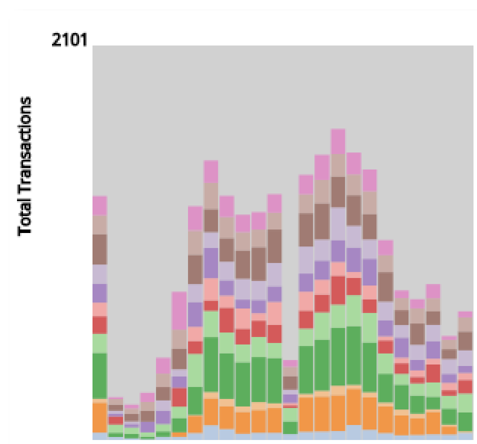
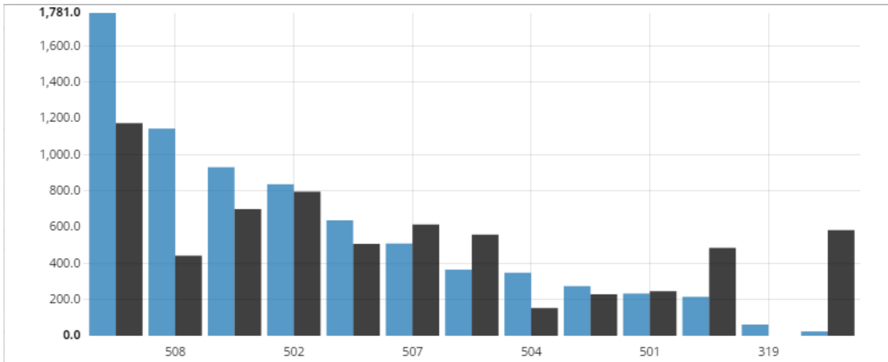


Figure 13. Farebox Graph by Route and Time of Day



AM	PM	Full
6 - 10	15 - 19	0 - 24
8045	7357	20304
124.93%	113.45%	

ROUTE NUMBER	COLOR KEY	RUN 108	RUN FBOX
505		1781	1173
508		1143	442
553		930	699
502		836	795
554		637	507
507		509	614
552		365	558
504		348	153
509		274	229
501		233	246
559		215	485
319		62	0
551		24	583
		7357	6484

Figure 14. Regression Model Output and farebox Route Comparison graph and table filtered by PM Peak

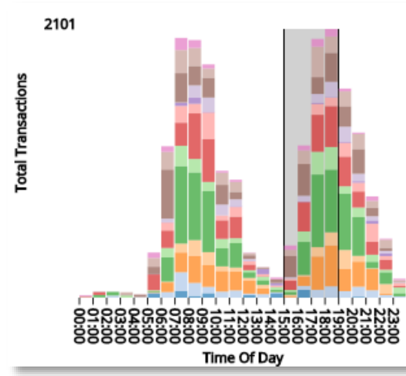


Figure 16. Regression Model Output by Route and Time of Day filtered by PM Peak

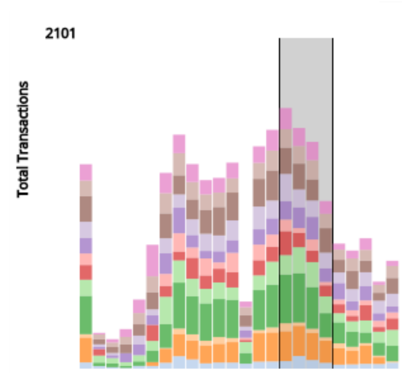


Figure 15. Farebox Graph by Route and Time of Day Filtered by PM Peak

Limitations of Transit Market Analyst Tool Suite

There are a number of limitations with the demand modeling tools as they are currently designed.

For Small and Mid-size Bus-Transit Markets

The Transit Market Analyst tool-suite is designed to analyze bus-to-work ridership in small and mid-sized market areas. The tools are not calibrated for more complex transit environments.

Non-Work Related Trips

The Transit Market Analyst demand modeling tool currently only accounts for bus-to-work ridership. Any analysis comparing Models to farebox for validation must take into account the absence of non-work related trips. Future advancements to the demand modeling tools should include an analysis of agency survey data to create a ratio for each market area of the work to non-work related trips. To account for the full range of bus riders, an additional long term goal should be to include trip purposes other than work (e.g., shopping, medical). Those riders should be included in the trip table using point-based trip generation based on land-marks and employment and bus-stops related data.

Generating AM and PM Peak Times

The demand modeling trip tables are currently built using an assumed 8 hour work day. The AM and PM Peak settings in Transit Market Analyst are hard-wired into the demand modeling and analysis tools. Future advancements of the tool should include means for setting AM and PM Peak times for model running so that the Peak times of the models more closely align with the farebox data.

Predicting Route Level Ridership Using Census Tract Geographies

The microsimulation modeling algorithm requires additional research to better predict bus-route level ridership. The algorithm currently distributes riders randomly throughout their home and work census tracts within a mile of a bus-stop to increase the likelihood that they will find a bus in the Open Trip Planner microsimulation. This decreases the accuracy of model generated route level ridership. They are declared bus riders according to the census, but they are being distributed to routes inaccurately.

Latent demand

The current version of the Transit Market Analyst lacks sufficient underlying behavioral data to develop latent demand models based solely on socio-demographic data. Future research is needed to examine whether different probabilities should be applied for individuals in households previously unserved by bus services to account for the likelihood of bus-ridership when bus service is introduced to previously underserved census geographies.

CTPP Perturbation

In order to be granted permission from the Census Bureau to use the raw ACS Data to develop the CTPP, a method referred to as perturbation is used to address disclosure concerns. The method uses a technique of adding random data during the processing. For example, some origins and destinations are randomized from the original raw data. As a result, there is some error purposely embedded in the CTPP data.

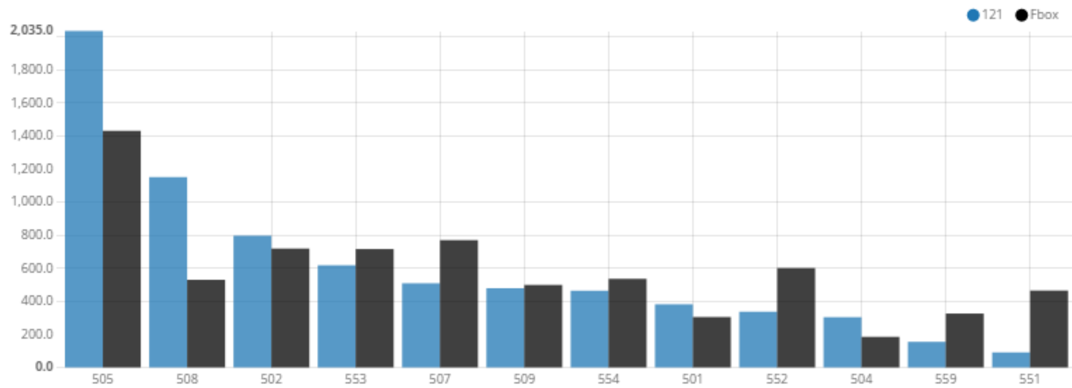
Software bugs

As a new technological approach to bus service planning, this project experienced a number of software bugs and design issues. These software problems caused significant delays and in some cases, drove the development trajectory.

Conclusions and Recommendations

Atlantic City

Atlantic City AM Findings



AM	PM	Full
Hours: 6 - 10	Hours: 15 - 19	Hours: 0 - 24
Model Total: 7322	Model Total: 6744	Model Total: 18638
FB%: 103.38%	FB%: 93.76%	FB%: 72.21%

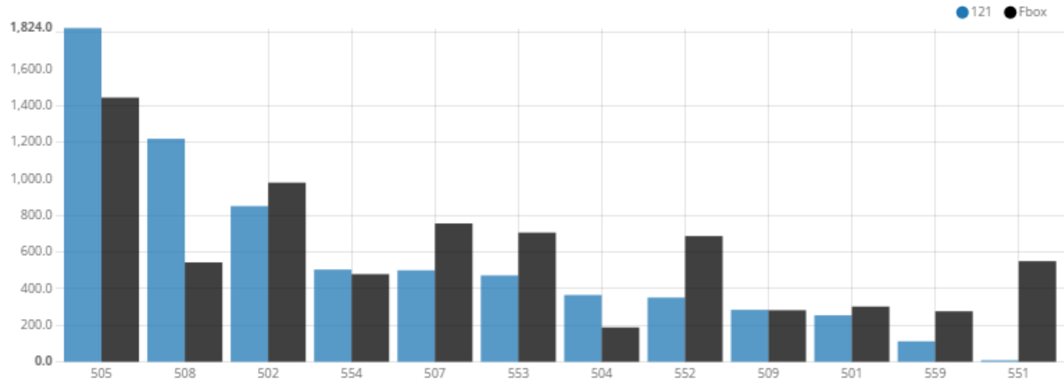
ROUTE NUMBER	COLOR KEY	RUN 121	RUN FBOX
505		2035	1431
508		1151	530
502		797	719
553		618	716
507		509	770
509		479	499
554		464	536
501		382	305
552		337	601
504		304	185
559		155	326
551		91	465
		7322	7083

Figure 17. Transit Market Analyst, Atlantic City AM Model Analysis Visualizations

- The Atlantic City Regression Model (See Appendix B) predicts a total AM ridership of 7322 for the market area. The farebox (July 2013, weekdays) averages a total AM Peak ridership of 7083.
- Using the Atlantic City Regression Model to test AM bus ridership, the microsimulation accuracy is within 3.38% of farebox.
- Microsimulation overestimates routes 508 and 505. These routes are both downtown on-island routes that compete directly with Jitney service and as a result farebox numbers are below microsimulation estimates. One possible explanation is that this represents bus riders that are not utilizing NJTransit.

- The microsimulation underestimates AM ridership on route 551, 552 and 553 which are long distance low-service bus routes that caters to non-work related patrons. This is likely related to entertainment activities downtown.
- The Microsimulation also underestimates AM ridership on route 507. The 507 runs parallel to the 509 and both routes overlap in leaving the Island.

Atlantic City PM Findings



AM	PM	Full
Hours: 6 - 10	Hours: 15 - 19	Hours: 0 - 24
Model Total: 7322	Model Total: 6744	Model Total: 18638
FB%: 103.38%	FB%: 93.76%	FB%: 72.21%

ROUTE NUMBER	COLOR KEY	RUN 121	RUN FBOX
505		1824	1444
508		1219	543
502		851	979
554		504	479
507		500	756
553		472	706
504		365	188
552		351	687
509		284	282
501		254	302
559		112	276
551		8	550
		6744	7192

Figure 18. Transit Market Analyst, Atlantic City PM Model Analysis Visualizations

- The Atlantic City Regression Model predicts a total PM ridership of 6744 for the market area. The farebox (July 2013, weekdays) averages a total PM Peak ridership of 7192.
- Using the Atlantic City Regression Model to test PM bus ridership, the microsimulation accuracy is within 6.24% of farebox.

- One possible explanation for the model under-performing during the PM Peak relates to the default peak time settings of the microsimulation modeling algorithm and the default PM Peak settings of the analysis tools. It appears that the PM Peak in Atlantic City starts slightly earlier than the default filter (see figure 19). Simultaneously, in place of actual commute times, the PM Peak settings in the microsimulation modeling algorithm – the return trips - are set 8 hours after AM commute times. As seen in figures 19 microsimulated ridership falls later than the default parameters for PM Peak and later than the farebox ridership. If the analysis tools allowed a user to adjust the PM Peak of the model, later in the day by one hour, and simultaneously adjust the PM Peak of the farebox an hour earlier, the microsimulation ridership increases to 7,611 and the farebox increases to 7,966. This would bring the model to within 4.46% of farebox

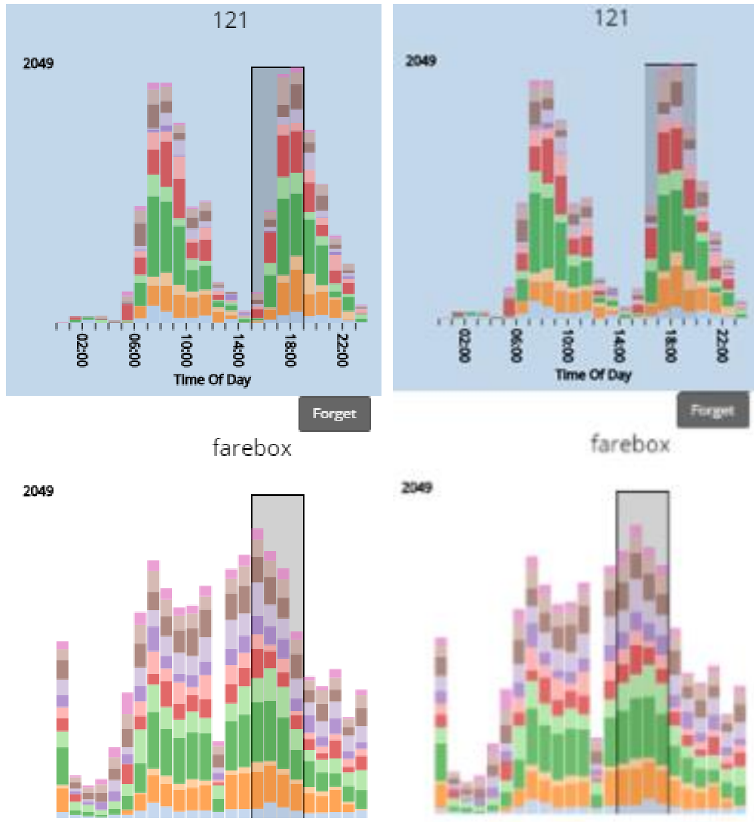
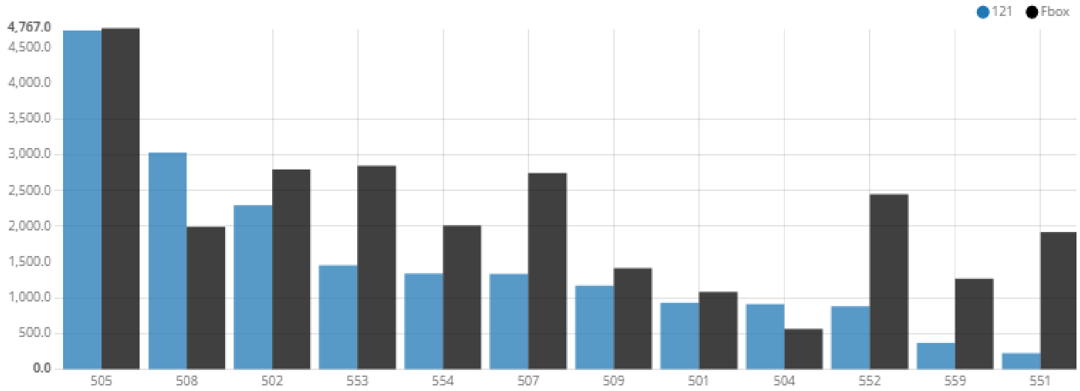


Figure 19. Transit Market Analyst, Atlantic City PM Peak Visualizations

- Microsimulation overestimates routes 508 and 505. These routes are both downtown on-island routes that compete directly with Jitney service and as a result farebox numbers are below microsimulation estimates. One possible explanation is that this represents bus riders that are not utilizing NJTransit.
- The microsimulation underestimates PM ridership on route 551, 552 and 553 which are both long distance low-service bus routes that caters to non-work related patrons. This is likely related to entertainment activities downtown.
- The Microsimulation also underestimates PM ridership on route 507. The 507 runs parallel to the 509 and both routes overlap in leaving the Island.

Atlantic City Full Day Findings



AM	PM	Full
Hours: 6 - 10	Hours: 15 - 19	Hours: 0 - 24
Model Total: 7322	Model Total: 6744	Model Total: 18638
FB%: 103.38%	FB%: 93.76%	FB%: 72.21%

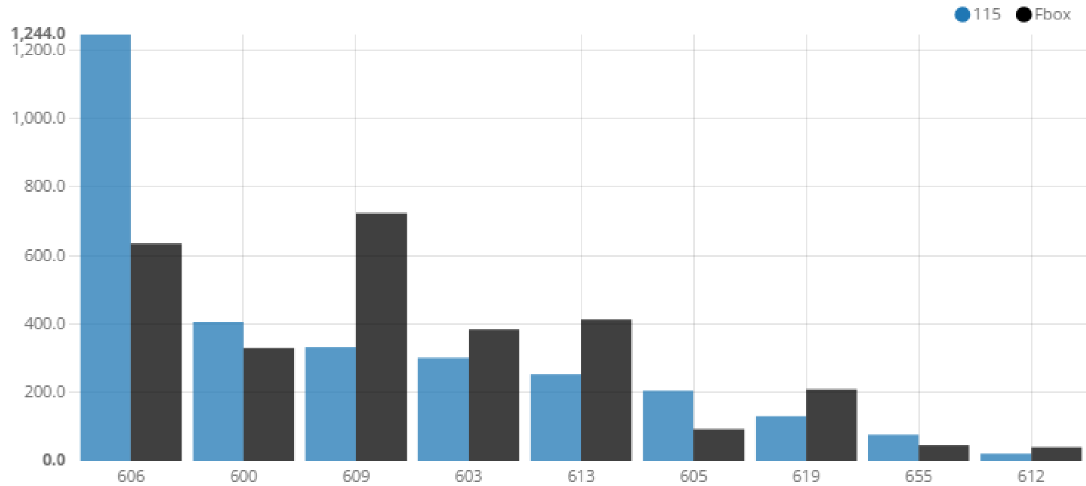
ROUTE NUMBER	COLOR KEY	RUN 121	RUN FBOX
505		4737	4767
508		3027	1990
502		2291	2792
553		1449	2840
554		1334	2009
507		1329	2740
509		1167	1411
501		927	1077
504		909	560
552		880	2443
559		369	1266
551		219	1915
		18638	25810

Figure 20. Transit Market Analyst, Atlantic City Full Day Model Analysis Visualizations

- The Atlantic City Regression Model predicts full day ridership of 18,638 for the market area. The farebox (July 2013, weekdays) averages a total full day ridership of 25,810.
- Using the Atlantic City Regression Model to test AM bus ridership, the microsimulation accuracy is within 27.79% of farebox.
- The overall underestimation of full day ridership is likely due to how the microsimulation algorithm does not account for non-work related trips.
- Microsimulation continues to overestimate Routes 508 and additionally, route 504. These routes are both downtown on-island routes that likely compete with Jitney service.

Princeton/Trenton

<u>Princeton/Trenton AM Peak Findings</u>



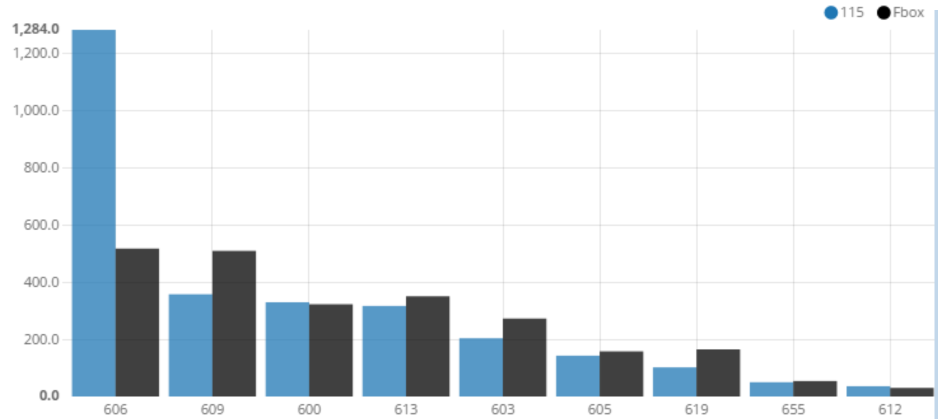
AM	PM	Full
Hours: 6 - 10	Hours: 15 - 19	Hours: 0 - 24
Model Total: 2970	Model Total: 2824	Model Total: 6806
FB%: 103.45%	FB%: 118.49%	FB%: 83.43%

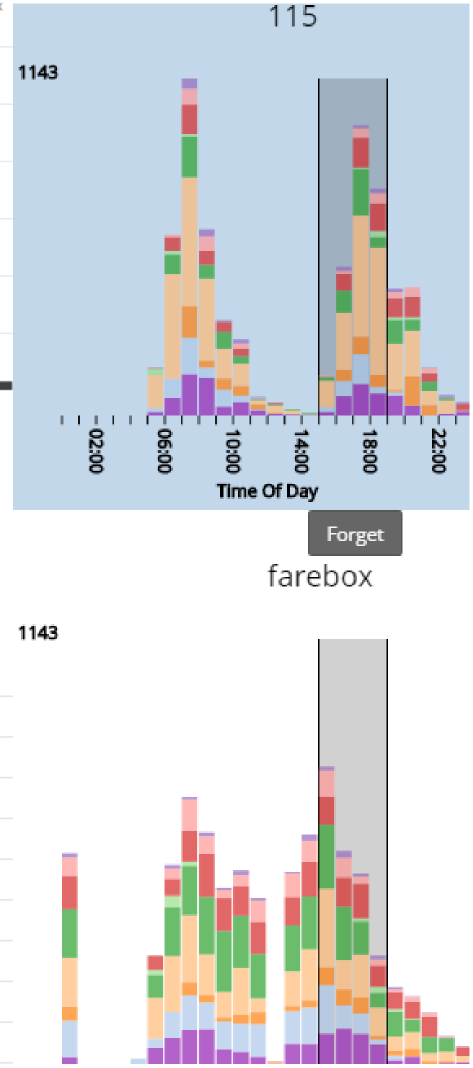
ROUTE NUMBER	COLOR KEY	RUN 115	RUN FBOX
606		1244	634
600		406	329
609		332	723
603		301	384
613		254	413
605		205	93
619		130	209
655		77	47
612		21	41
		2970	2873

Figure 21. Transit Market Analyst Princeton/Trenton AM Model Analysis Visualizations

- The Princeton/Trenton Regression Model (See Appendix D) predicts a total AM Peak ridership of 2970 for the market area. The farebox (July 2013, weekdays) averages a total AM Peak ridership of 2873.
- Using the Princeton/Trenton Regression Model to test AM Peak bus ridership, the microsimulation accuracy is within 3.45% of farebox.
- Microsimulation overestimates route 606 and underestimes 609 and 613.

Princeton/Trenton PM Peak Findings





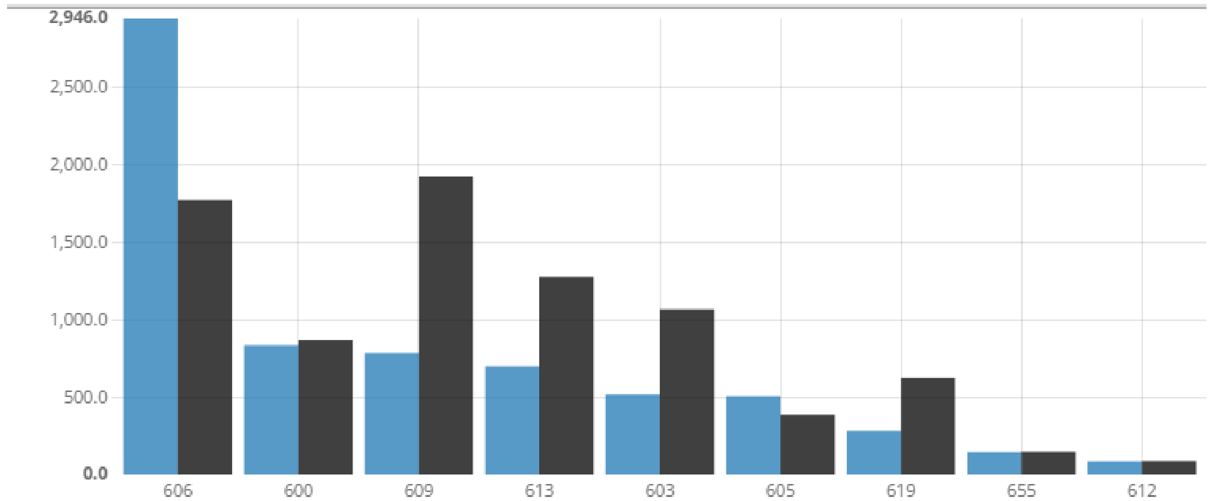
AM	PM	Full
Hours: 6 - 10	Hours: 15 - 19	Hours: 0 - 24
Model Total: 2970	Model Total: 2824	Model Total: 6806
FB%: 103.45%	FB%: 118.49%	FB%: 83.43%

ROUTE NUMBER	COLOR KEY	RUN 115	RUN FBOX
606		1284	518
609		358	510
600		330	323
613		317	351
603		204	273
605		143	158
619		102	165
655		50	54
612		36	30
		2824	2382

Figure 22. Transit Market Analyst Princeton/Trenton PM Model Analysis Visualizations

- The Princeton/Trenton Regression Model (See Appendix D) predicts a total PM Peak ridership of 2824 for the market area. The farebox (July 2013, weekdays) averages a total PM Peak ridership of 2382.
- Using the Princeton/Trenton Regression Model to test AM Peak bus ridership, the microsimulation accuracy is within 18.49% of farebox.
- One possible explanation for the model under-performing during the PM Peak relates to the default PM Peak settings of the analysis tools. It appears that the PM Peak in Trenton/Princeton starts slightly earlier than the default filter (see figure X). If we adjust the PM Peak of farebox an hour earlier than the default PM Peak, the farebox ridership increases to 2,794. This would bring the model to within 1.07% of farebox
- Microsimulation overestimates PM ridership on route 606 and underestimates PM ridership on Route 609.

Princeton/Trenton Full Day Findings

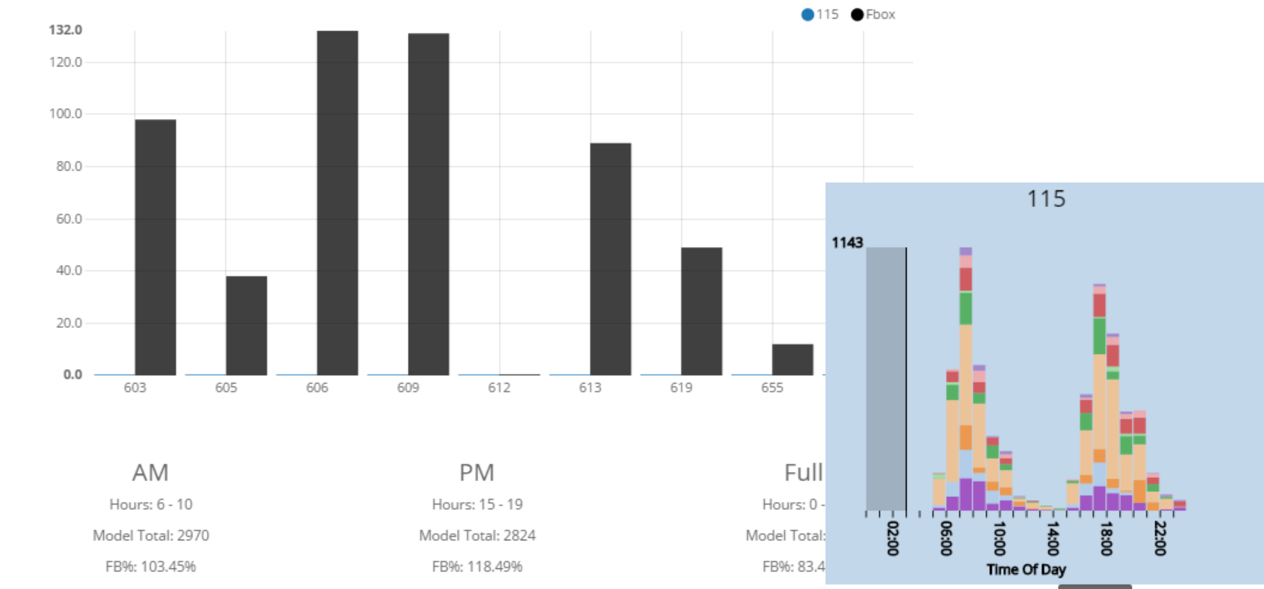


AM	PM	Full
Hours: 6 - 10	Hours: 15 - 19	Hours: 0 - 24
Model Total: 2970	Model Total: 2824	Model Total: 6806
FB%: 103.45%	FB%: 118.49%	FB%: 83.43%

ROUTE NUMBER	COLOR KEY	RUN 115	RUN FBOX
606		2946	1773
600		835	869
609		784	1923
613		700	1276
603		519	1067
605		506	387
619		284	625
655		146	148
612		86	88
		6806	8156

Figure 23. Transit Market Analyst Princeton/Trenton Full Day Model Analysis Visualizations

- The overall underestimation of full day ridership is due to how the microsimulation algorithm does not account for non-work related trips.
- Farebox data from Trenton/Princeton shows a high level of ridership during the midnight hour (see figure 23). Given that the microsimulation modeling algorithm does not account for trips at that time of the day, we can easily remove these trips from the full day farebox data.



ROUTE NUMBER	COLOR KEY	RUN 115	RUN FBOX
603		0	98
605		0	38
606		0	132
609		0	131
612		0	0
613		0	89
619		0	49
655		0	12
600		0	18
		0	567

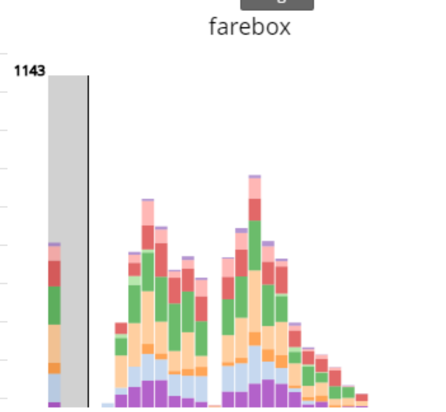


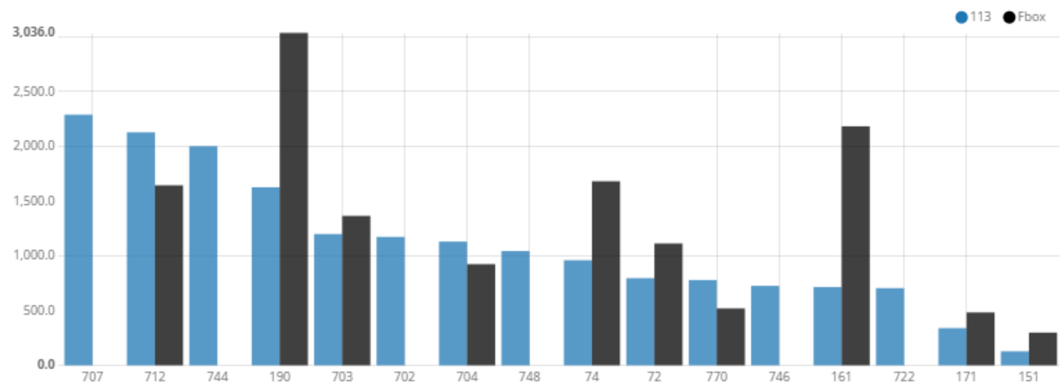
Figure 24. Transit Market Analyst Princeton/Trenton Full Day Analysis Visualizations

- We can use the model validation analysis filtering tool to see how many riders on average, take the bus between 12AM and 1PM. The total ridership for that period of time is 567. If we adjust the full day farebox ridership accordingly we get a total ridership of 7,589. With midnight ridership removed, the full day model performs to within 10.39% of farebox.
- The full day model overestimates ridership on the 606 but significantly underestimates ridership on the 609, 613, and 619.

Paterson

Of all of the regression models in this report, the Paterson model is the lowest performing model, this is in part, due to the size of the market area and the high level of non-work related bus ridership. It should also be noted that there is a limited number of routes with farebox data available for validation.

<u>Paterson AM Peak Findings</u>



AM	PM	Full
Hours: 6 - 10	Hours: 15 - 19	Hours: 0 - 24
Model Total: 1761	Model Total: 1454	Model Total: 3459
FB% 136.00	FB%: 167.38	FB%: 82.64

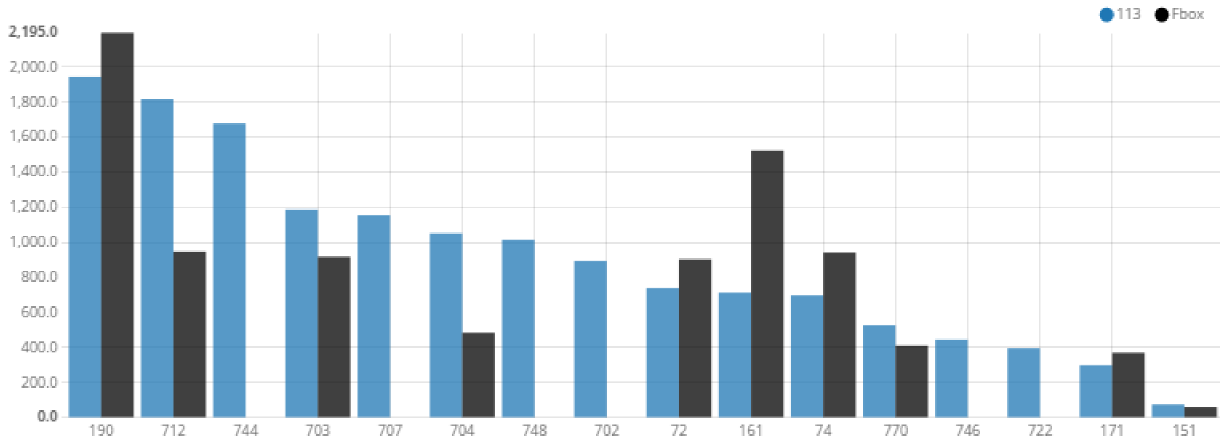
ROUTE NUMBER	COLOR KEY	RUN 113	RUN FBOX
707		2289	0
712		2128	1644
744		2001	0
190		1626	3036
703		1199	1365
702		1173	0
704		1130	924
748		1044	0
74		960	1681
72		796	1113
770		779	520
746		726	0
161		715	2183
722		705	0
		17611	12949

Figure 25. Transit Market Analyst, Paterson AM Model Analysis Visualizations

- The Paterson Regression Model (See Appendix C) predicts a total AM ridership of 17611 for the market area. The farebox averages a total AM Peak ridership of 12949.
- Using the Paterson Regression Model to test AM bus ridership, the microsimulation accuracy is within 36% of farebox.

- Part of the explanation for this drastic difference between predicted ridership and farebox ridership is that there are a number of routes that are included in the market area for which farebox contains no data. This includes route numbers 707, 744, 702, 748, 746, and 722. On those routes combined, the microsimulation model predicts a total of 7,398 riders during AM Peak. If you reduce the microsimulation ridership accordingly you have a total AM Peak ridership of 10,213. This would bring the model to within 21.13% of farebox.

Paterson PM Peak Findings



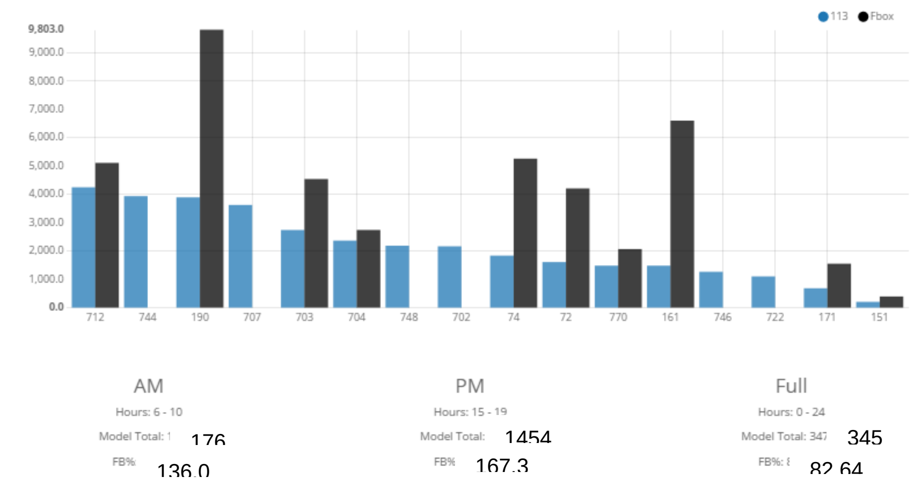
AM	PM	Full
Hours: 6 - 10	Hours: 15 - 19	Hours: 0 - 24
Model Total: 1761	Model Total: 1454	Model Total: 3459
FB%: 136.0	FB%: 167.3	FB%: 82.64

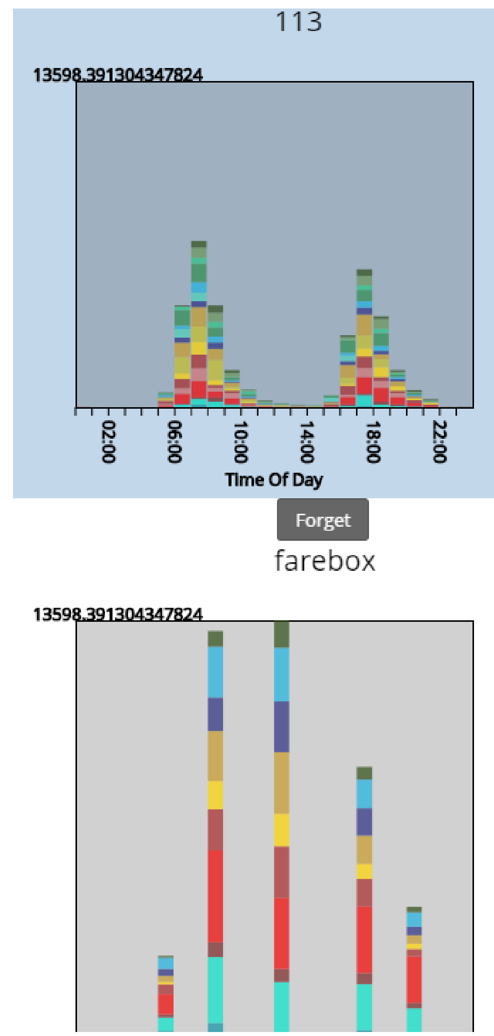
ROUTE NUMBER	COLOR KEY	RUN 113	RUN FBOX
190		1944	2195
712		1815	948
744		1677	0
703		1186	916
707		1155	0
704		1050	484
748		1013	0
702		893	0
72		737	904
161		713	1524
74		697	942
770		526	409
746		445	0
722		396	0
171		298	368
		14545	8690

Figure 26. Transit Market Analyst, Paterson PM Model Analysis Visualizations

- The Paterson Regression Model (See Appendix C) predicts a total AM ridership of 14545 for the market area. The farebox averages a total AM Peak ridership of 8690.
- Using the Paterson Regression Model to test AM bus ridership, the microsimulation accuracy is within 67.38% of farebox.
- Part of the explanation for this drastic difference between predicted ridership and farebox ridership is that there are a number of routes that are included in the market area for which farebox contains no data. This includes route numbers 707, 744, 702, 748, 746, and 722. On those routes combined, the microsimulation model predicts a total of 5,579 riders during AM Peak. If you reduce the microsimulation ridership accordingly you have a total AM Peak ridership of 8,966. This would bring the model to within 3.1% of farebox.

Paterson Full Day Findings





ROUTE NUMBER	COLOR KEY	RUN 113	RUN FBOX
712		4245	5106
744		3934	0
190		3890	9803
707		3621	0
703		2738	4536
704		2364	2739
748		2184	0
702		2163	0
74		1834	5256
72		1608	4205
770		1481	2065
161		1480	6596
746		1264	0
722		1101	0
171		683	1549
		34590	41855

Figure 27. Transit Market Analyst, Paterson Full Day Model Analysis Visualizations

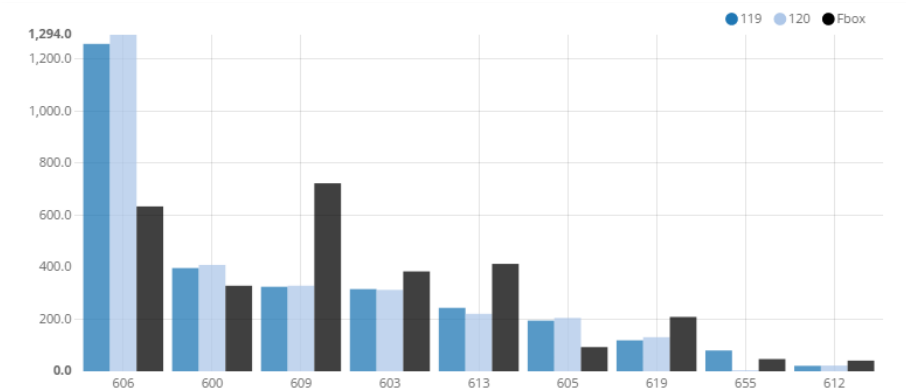
- The Paterson Regression Model full day results predicts full day ridership of 34,590 for the market area. The farebox averages a total full day ridership of 41,855.
- The overall underestimation of full day ridership is most easily explained by the size of the Paterson market area and by how the microsimulation algorithm does not account for non-work related trips.

Scenario Case Study: Ridership on Route 655

The Route 655 case study is a real-world scenario based on GTFS change. The 655 was introduced to the Princeton/Trenton Market Area by NJTransit to address a perceived need but was later removed based on low performance.

This scenario case study requires running two different models for Princeton/Trenton. The first microsimulation model includes the 655 in the Princeton/Trenton market area. For the second microsimulation model, the 655 is removed. Then, to compare the two models against farebox data that contains route 655 (July, 2013), the route is reintroduced to the Princeton/Trenton market area.

AM Peak Findings



AM	PM	Full
Hours: 6 - 10	Hours: 15 - 19	Hours: 0 - 24
Model Total: 2956	Model Total: 2848	Model Total: 6787
FB%: 102.97%	FB%: 119.50%	FB%: 83.20%

ROUTE NUMBER	COLOR KEY	RUN 119	RUN 120	RUN FBOX
606		1259	1294	634
600		397	409	329
609		325	329	723
603		316	313	384
613		244	221	413
605		195	205	93
619		119	131	209
655		80	0	47
612		21	22	41
		2956	2924	2873

Figure 28. Transit Market Analyst, Route 655 AM Model Analysis Visualizations

- For model Run 119, route 655 is part of the Princeton/Trenton market area. In Run 120 it has been removed.
- The first thing we notice is that Run 119 had 80 riders on the 655 during AM Peak. The farebox shows only 47 riders during AM Peak. Run 119 overestimates AM Peak ridership on the 655 by 33.
- When 655 is removed (Run 120) 32 of the original 80 riders were unable to be routed. However, these riders were still accounted for in the trip table and couldn't find service in the microsimulation. The remaining 48 riders found their way onto the service network. These 32 riders may account for the overestimation of ridership on the 655 in Run 119.
- There is a further shift of ridership based on the removal of the 655. Route 613 shows a significant reduction in ridership as well. In the farebox data, the 613 shows an average of 413 riders during AM Peak. With the 655 route included in the market area, Run 119 shows 244 riders on the 613. With the 655 removed, that number is reduced to 221 riders.
- Taken together, the 48 riders from the 655, and the 23 riders from the 613, account for 71 riders in the AM Peak that increased ridership on many of the routes in the market area. Most of those riders show up on the 606 (31) and to a lesser extent the 619 (11) and the 605 (10) for a total of 52 riders. The rest of the ridership shuffled around the service network in insignificant ways.

Discussion

GTFS is the market area geo-spatial backbone that allows the entire suite of software tools to visualize socio-demographic data, agency ridership data, and demand modeling. Additionally, small and medium size market area models require local knowledge. For example, NJTransit staff provided input on the behavior of local shuttle (i.e., Jitneys in Atlantic City) and long haul service. The reason for this, is the basic model structure is for residentially generated AM bus to work trips. The AM bus to work trip model can be leveraged to generate PM Peak trips and full day trips with several modifications.

All of the market area models underestimate full day ridership, despite often over-estimating peak-time ridership. This is expected due to the microsimulation model algorithm only accounting for work trips. It also points to how Peak ridership behaves differently from market area to market area. Farebox data indicates fairly steady ridership throughout the day with a more concentrated AM Peak than PM Peak.

Due to assumptions made in trip table generation about an 8-hour workday and the lack of information about work to home trips, the microsimulation modeling algorithm shows overly concentrated peaks, as well as a PM Peak that generally start later, than farebox data indicates.

In dense urban areas with two tracts in downtown and a number of busses going between the two tracts, the microsimulation does not distribute the trips accurately. This is another example of attempting to forecast cross-town ridership using a residentially generated AM bus to work trips. While service levels are included in the microsimulation modeling algorithm, we still find that while the overall market area is accurate in the peaks, there are a number of trips that show up in farebox attributed to one route showing up on a different route. The modeling tool doesn’t differentiate well between two routes competing for ridership in overlapping census tracts this could be improved by using Transportation Analysis Zone (TAZ) census geographies and/or using a three-stage-least-squares estimation method developed by Peng (1994), for competing routes.

Lastly, the microsimulation modeling algorithm requires additional research to better predict bus-route level ridership. The algorithm currently distributes riders randomly throughout their home and work census tracts within a mile of a bus-stop to increase the likelihood that they will find a bus in the Open Trip Planner microsimulation. This likely decreases the accuracy of model generated route level ridership. They are declared bus riders according to the census, but they are being distributed to routes inaccurately. There are a number of approaches that could be explored for improving origin/destination location generation, such as using parcel data, point-based establishment and employment data, and expanded survey data.

Given these route level problems associated with ridership distribution, it is encouraging to see that the microsimulation modeling algorithm can locate latent demand. The example of the Route 655 case study scenario demonstrates that there were potential riders in market area that were not being served. The model Run 119 overestimates AM Peak ridership on the 655 by almost exactly the same amount as the number of missing riders in Run 120, once the route is removed. It is possible, then, to hypothesize that the 80 riders on the 655 that were predicted by Run 119, were either randomly placed close enough to the 655 to find their way onto the 655 through microsimulation, or they are located in the route 655 commute-shed but didn't show up in the farebox data as "actual" 655 riders due to previously formed habits of commuting.

Conclusion and Recommendations

The Integration of Bus Stop Counts Data with Census Data for Improving Bus Service research project produced an open source transit market data visualization and analysis tool suite, *Transit Market Analyst*.

The *Transit Market Analyst* combines the rich source of archived transit operations data (e.g., automatic farebox data), with new open data resources, particularly GTFS and US Census. It is web-based and open source, allowing for easy deployment and consistent non-proprietary upgrades. The data visualization and informatics tool suite offers an enhanced perspective on NJTransit's own transportation assets, and is capable of:

- Identifying the key demographic factors that influence transit ridership;
- Assessing the social, economic and systemic determinants that exist within the 2010 census;
- Developing solutions that provide persistent competitive advantage for NJTransit; and
- Providing the tools for continued success subsequent to the completion of this research project.

The microsimulation modeling algorithm employed by *Transit Market Analyst* requires additional research to better predict bus-route level ridership. There are a number of approaches that could be explored for improving origin/destination location generation, such as using parcel data, point-based establishment and employment data, and expanded survey data.

Even with these bus-route level ridership inaccuracies, it appears that the microsimulation modeling algorithm can locate latent demand. In the example of the Route 655 case study scenario, there were potential riders in market area that were not being served.

The most promising future research should address the use of farebox data at the stop level and the landmarks in the vicinity of the stop to clarify trip purpose. This will reduce the need for surveying and provides a monitoring and validating data strategy going forward. The ability to develop hypothetical scenarios using the agency generated GTFS could improve accuracy, cost effectiveness and service efficiency.

IMPLEMENTATION AND TRAINING

The research team devised an implementation and training plan to introduce the Transit Market Analyst software to NJ TRANSIT staff. As part of the implementation plan, two NJ TRANSIT employees working in demand modeling were identified as the shepherds, tasked with transferring the knowledge, deploying the technology, and trained as internal points of reference to support users of the software within the agency.

In order to accommodate this implementation and training plan, the research team acquired a no-cost extension to the project, effectively extending the end date from December 31, 2015 to April 30, 2016. During the months of January, February and March of 2016, the research team held weekly trainings with the NJ TRANSIT employees on understanding the software and the accompanying report. The research team simultaneously organized a hands-on training seminar for mid-April 2016 for other employees at NJ TRANSIT.

Additionally, the software will be hosted on servers located at the University at Albany for at least one year beyond the project end date.

APPENDIX A: OVERVIEW ON BUILDING REGRESSION MODELS

The following is an overview of processes followed in building the regression models derived from the census data. At this time, the regression models show high sensitivity to changes in geographies.

Extract Census Tracts from Web Tool

Our first step is to extract the GeoID's of the Census Tracts for a given market area from the database of our web-tool. For Atlantic City, the GeoID's are shown below:


```
"34001010102","34001010104","34001010105","34001001400","34001002500","34001002400","34001001900","340
01001200","34001001500","34001001100","34001000300","34001002300","34001000500","34001000400","3400100
0100","34001000200","34001013201","34001013202","34001013302","34001013301","34001013000","34001013101
","34001013102","34001001300","34001010101","34001010505","34001010200","34001011900","34001010300","34
001012000","34001012100","34001012200","34001012401","34001012302","34001012402","34001012501","340010
13500","34001012602","34001012701","34001012801","34001011803","34001012502","34001011702","3400101170
1","34001010506","34001010503","34001983400","34001010403","34001011802","34001011804","34001011404","3
4001012702","34001011805","34001011403","34009020201","34009020101","34009020102","34009020203","34001
012802","34001011600","34001011500","34001011401","34011030100","34009020301","34009020205","340090208
00","34009020206","34001010401","34001010501","34009021001","34009020902","34009020901","34009021100","
34009022102","34009022101","34009021804","34009021803","34009021701","34009021702","34009021805","3400
9021806","34009021900","34009022000","34009021600","34009021500","34009021400","34009021300","34009021
002","34009020302","34009020700","34011041000","34011040800","34011040700","34011040600","34011040500",
"34011041100","34011040300","34011040400","34001011202","34001011100","34001011000","34001010900","3400
1010800","34029736101","34029736102","34029737000","34029736002","34029736001"
```


Creating a Data Rich GIS File

Now that we have a subset of Census Tracts, we slice this group out of our database that contains New Jersey Census Tracts, and query the Census API to acquire a selection of ACS data for these tracts. Once completed, this data is exported from our database as a JavaScript Object Notation (JSON) object that contains the census geographies and with the ACS data as attributes, this JSON object is then converted to a shapefile for use in GIS applications.

geoid	total_popu	employment	unemployme	travel_to_	car_to_wor	public_tra	bus_to_wor
34009020901	828	318	27	308	244	0	0
34009021002	3623	1930	141	1897	1735	16	16
34009022102	5533	2367	356	2299	2140	37	37
34009021001	2600	848	89	835	738	20	20
34009022101	1887	649	57	649	545	0	0
34009020301	4202	2136	185	2070	1969	10	0
34009021500	2150	1116	206	1084	944	43	24
34009020102	2413	1220	47	1210	867	19	9
34009021400	3650	1687	202	1657	791	104	101
34009020101	3307	1703	172	1703	1190	48	38
34009021804	5603	2789	343	2633	2448	79	79
34009021300	4111	1670	196	1580	1411	25	12

Figure 29. Data Rich GIS FIle

Data Interrogation and Regression Methodology

Then we look at the data to see which variables correlate with Bus to Work and we build the regression models by adding and removing highly correlative census variables one at a time.

Meta Data

The data used in this study was obtained from the US Census Application Programming Interface (API). The data set is called the American Community Survey Five-Year Data 2006-2010 (ACS). The ACS is an ongoing survey that provides data every year -- giving communities the current information they need to plan investments and services. The ACS covers a broad range of topics about social, economic, demographic, and housing characteristics of the U.S. population.[1] Employment Density (EMP_DEN) and Population Density (POP_DEN) were derived by dividing the employment at tract level by polygon tract area and population at tract level by polygon tract area.

Descriptive Statistics

Descriptive Statistics include, among others, the mean, median and standard deviation of each of the variables. The **Mean** and **Median** are measures of central tendency. The **Mean** is the numerical value found by summing the values and dividing by the number of cases. The **Median** is the numerical value separating the higher half of a data sample from the lower half. The **Median** can be found by arranging all observations from lowest value to highest value and picking the middles. The **Standard Deviation** measures the amount of variation or dispersion from the average and is equal to the square root of the sample variance. The **Sum (Frequency)** is the total cases for each variable. The **Percent of Category** is the ratio of the **Sum** to the parent **Category Sum**.

Correlations

A correlation coefficient is the measure of strength of the linear association between two variables (-1 to +1). This table contains only variables that have a statistically significant correlation with the bus_to_wor variable.

Regression Methodology

The model used in this analysis is a linear regression model that assumes a linear relationship between the dependent variable (bus_to_wor) and a set of independent variables. A regression model fits a straight line to a set of observed data and provides the statistical significance of the included variables. [2]

The regression model will produce a number of parameters and model fitting indicators such as the coefficient of determination (R Squared). The R Squared is defined as the percent of the variation of the dependent variable (bus_to_wor) explained by the set of independent variables. The percent of bus riders from each census tract will be explained by the regression model's set of independent variables. Therefore the higher the R Squared the more explanatory power the model provides.

The regression model output also provides a constant (intercept) which is the average value of the dependent variable when the independent variables equal zero.[3]

---

[1]United States Census Bureau, American Community Survey, http://www.census.gov/data/developers/data-sets/acs-survey-5-year-data.html

[2] Rogerson, Peter A., 2006, *Statistical Methods for Geography 2nd Edition,* London: Sage Publications

[3] Lewis-Beck, Michael S., 1980, *Applied Regression, An Introduction,* Newbury Park: Sage Publications

Slope coefficient indicate the average change in the dependent variable with a one unit change in the independent variable.

For the purposes of this modelling effort statistical significance is defined as a p-value of <.05 or a t-value >2.5.

APPENDIX B: ATLANTIC CITY DESCRIPTIVE STATISTICS AND CORRELATIONS

Atlantic City Descriptive Statistics

Table 7. Atlantic City Descriptive Statistics

Atlantic City Descriptive Statistics											
	Description	Category	N	Mean	Median	Std. Deviation	Variance	Min	Max	Sum (Freq)	% of Category
total_popu			110	3,900	3,503	2,216	4,910,855	929	15,676	428,946	
total_hous			110	2,243	2,114	1,317	1,734,157	7	8,573	246,770	
unemployme	Unemployed Population	Labor Force	110	179	151	127	16,219	0	714	19,664	8.87%
public_tra	Journey to Work by Public Transportation Total	Journey to Work	110	105	45	141	19,848	0	795	11,533	5.91%
bus_to_wor	Journey to Work by Public Transportation by Bus or Trolley Bus	Journey to Work, Public Trans.	110	96	41	138	19,135	0	774	10,580	91.74%
informatio	Employment in Information	Labor Force	110	25	17	29	817	0	135	2,732	1.23%
arts	Employment in Arts	Labor Force	110	411	339	346	119,507	26	1,879	45,198	20.39%
under_1000	Annual Income Under $10,000	House-holds	110	92	64	79	6,221	0	354	10,144	4.11%
10000_1499	Annual Income $10,000-$14,999	House-holds	110	84	62	73	5,316	0	453	9,210	3.73%
15000_1999	Annual Income $15,000-$19,999	House-holds	110	73	65	57	3,201	0	321	8,015	3.25%
25000_2999	Annual Income $25,000-$29,999	House-holds	110	76	64	56	3,121	0	293	8,379	3.40%
30000_3499	Annual Income $30,000-$34,999	House-holds	110	80	62	62	3,901	0	326	8,803	3.57%
35000_3999	Annual Income	House-holds	110	76	57	60	3,643	0	249	8,316	3.37%

	$35,000-$39,999										
125000_149	Annual Income $125,000-$149,999	House-holds	110	72	56	64	4,140	0	360	7,950	3.22%
150000_199	Annual Income $150,000-$199,999	House-holds	110	72	65	58	3,417	0	399	7,890	3.20%
200000+	Annual Income Greater than $200,000	House-holds	110	51	41	52	2,752	0	258	5,634	2.28%
poverty_st	Poverty Status	House-holds	110	421	302	374	139,642	3	1,803	46,312	18.77%
no_high_sc	No High School Education	House-holds	110	248	166	245	59,841	0	1,521	27,239	11.04%
bachelors	Bachelors degree	Labor Force	110	519	440	360	129,431	0	2,511	57,066	25.75%
foreign_bo	Foreign Born	Population	110	486	299	502	251,963	11	2,109	53,457	12.46%
spanish_sp	Spanish Speaking	House-hold	110	219	93	339	114,801	0	2,396	24,102	9.77%
other_lang	Other Language Speaking	House-hold	110	135	65	179	32,212	0	905	14,814	6.00%
	Description	Category	N	Mean	Median	Std. Deviation	Variance	Min	Max	Sum (Freq)	% of Category
age25_29	Age 25 to 29 Total	Population	110	220	187	161	25,970	0	711	24,254	5.65%
age30_34	Age 30 to 34 Total	Population	110	213	145	179	32,219	0	991	23,478	5.47%
race_white	Race White	Population	110	2,826	2,451	1,816	3,296,559	194	11,359	310,871	72.47%
race_black	Race Black	Population	110	499	205	693	479,619	0	3,885	54,858	12.79%
race_asian	Race Asian	Population	110	200	75	319	102,059	0	1,406	21,945	5.12%
race_other	Race Other	Population	110	270	97	385	148,082	0	2,171	29,661	6.91%
race_two	Bi-racial	Population	110	90	57	107	11,427	0	491	9,943	2.32%
1_unit_det	Housing 1 Unit Detached	House-holds	110	1,255	1,123	818	669,284	7	4,676	138,027	55.93%
5_9units	Housing 5-9 Units	House-holds	110	110	54	154	23,654	0	800	12,143	4.92%
20_49units	Housing 20-49 Units	House-holds	110	55	24	88	7,779	0	572	6,014	2.44%
50+_units	Housing 50+ Units	House-holds	110	140	38	251	63,157	0	1,424	15,403	6.24%
occupanc_1	Tenure, Occupancy Status, Renter Occupied	House-holds	110	435	371	326	106,117	0	1,340	47,852	19.39%
car_0	Zero Vehicles Available by Worker	Labor Force, Employed	110	112	60	147	21,632	0	895	12,277	N/A

car_1	One Car Available by Worker	Labor Force, Employed	110	394	363	241	58,093	0	1,499	43,375	N/A
car_3	Three Cars Available by Worker	Labor Force, Employed	110	315	251	303	92,014	0	2,106	34,622	N/A
car_4	Four Cars Available by Worker	Labor Force, Employed	110	111	79	112	12,517	0	483	12,156	N/A
car_0_hous	Households, Zero Vehicles Available	House-holds	110	185	116	196	38,382	0	1,075	20,375	8.26%
car_1_hous	Households, One Car Available	House-holds	110	535	502	268	71,890	7	1,274	58,878	23.86%
car_3_hous	Households, Three Cars Available	House-holds	110	168	136	148	22,030	0	946	18,477	7.49%
car_4_hous	Households, Four Cars Available	House-holds	110	49	37	49	2,376	0	203	5,397	2.19%
emp_den	Employment/ Area	N/A	110	1,469	857	2,313	5,352,048	32	14,801	161,627	N/A
pop_den	Population/Area	N/A	110	3,278	1,677	5,118	26,192,401	68	33,671	360,527	N/A

Atlantic City Correlations

Table 8. Atlantic City Correlations

Variable	Description	Pearson Correlation
unemployme	Unemployed Population	.240*
public_tra	Journey to Work by Public Transportation Total	.993**
bus_to_wor	Journey to Work by Public Transportation by Bus or Trolley Bus	1
constructi	Employment in Construction	-.208*
arts	Employment in Arts	.558**
under_1000	Annual Income Under $10,000	.400**
10000_1499	Annual Income $10,000-$14,999	.191*
15000_1999	Annual Income $15,000-$19,999	.234*
20000_2499	Annual Income $20,000-$24,999	.235*
25000_2999	Annual Income $25,000-$29,999	.266**
30000_3499	Annual Income $30,000-$34,999	.326**
35000_3999	Annual Income $35,000-$39,999	.328**
125000_149	Annual Income $125,000-$149,999	-.256**
150000_199	Annual Income $150,000-$199,999	-.287**
200000+	Annual Income Greater than $200,000	-.294**
poverty_st	Poverty Status	.459**
no_high_sc	No High School Education	.453**
bachelors	Bachelors Degree	-.190*
foreign_bo	Foreign Born	.605**
spanish_sp	Spanish Speaking	.429**
other_lang	Other Language Speaking	.507**
age25_29	Age 25 to 29 Total	.304**
age30_34	Age 30 to 34 Total	.219*
race_white	Race White	-.288**
race_black	Race Black	.538**
race_asian	Race Asian	.403**

race_other	Race Other	.485**
race_two	Two or More Races	.210*
1_unit_det	Housing Unit Detached	-.338**
5_9units	Housing 5-9 Units	.216*
20_49units	Housing 20-49 Units	.206*
50+_units	Housing 50+ Units	.389**
car_0	Zero Vehicles Available by Worker	.716**
car_1	One Car Available by Worker	.499**
car_3	Three Cars Available by Worker	-.208*
car_0_hous	Households, Zero Vehicles Available	.583**
car_1_hous	Households, One Vehicle Available	.222*
car_2_hous	Households, Two Vehicles Available	-.193*
car_3_hous	Households, Three Vehicles Available	-.242*
car_4_hous	Households, Four Vehicles Available	-.245**
car_5+_hou	Households, Five or More Vehicles Available	-.224*
emp_den	Employment/Area	.594**
pop_den	Population/Area	.575**

Atlantic City Regression Model

Table 9. Atlantic City, Regression Model

Model		**Unstandardized Coefficients**		**Standardized Coefficients**	**t**	**Sig.**	**95.0% Confidence Interval for B**	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	-41.505	14.316		-2.899	.005	-69.887	-13.123
	car_0_hous	.230	.051	.326	4.542	.000	.130	.331
	arts	.163	.025	.406	6.432	.000	.112	.213
	emp_den	.019	.004	.321	4.442	.000	.011	.028

APPENDIX C: PATERSON CASE STUDY

Paterson Descriptive Statistics

Table 10. Paterson Descriptive Statistics

PT Descriptive Statistics											
	Description	Category	N	Mean	Median	Std. Deviation	Variance	Min	Max	Sum (Freq)	% of Category
total_popu			128	4,839	4,770	1,603	2,569,208	406	8,409	619,366	
total_hous			128	1,767	1,713	585	342,754	0	3,334	226,184	
unemployme	Unemployed Population	Labor Force	128	167	148	102	10,319	0	489	21,362	6.83%
public_tra	Journey to Work by Public Transportation Total	Journey to Work	128	238	221	154	23,733	0	779	30,439	10.73%
bus_to_wor	Journey to Work by Public Transportation by Bus or Trolley Bus	Journey to Work, Public Trans.	128	192	171	138	19,067	0	755	24,614	80.86%
informatio	Employment in Information	Labor Force	128	59	50	53	2,800	0	248	7,586	2.43%
arts	Employment in Arts	Labor Force	128	146	125	95	9,042	0	565	18,703	5.98%
under_1000	Annual Income Under $10,000	House-holds	128	127	105	103	10,571	0	538	16,211	7.17%
10000_1499	Annual Income $10,000-$14,999	House-holds	128	81	72	51	2,552	0	230	10,335	4.57%
15000_1999	Annual Income $15,000-$19,999	House-holds	128	81	71	57	3,252	0	251	10,327	4.57%
25000_2999	Annual Income $25,000-$29,999	House-holds	128	80	70	51	2,559	0	264	10,272	4.54%
30000_3499	Annual Income $30,000-$34,999	House-holds	128	74	66	48	2,301	0	201	9,417	4.16%
35000_3999	Annual Income $35,000-$39,999	House-holds	128	64	57	45	2,032	0	188	8,197	3.62%
125000_149	Annual Income $125,000-$149,999	House-holds	128	98	88	84	7,063	0	375	12,607	5.57%
150000_199	Annual Income $150,000-$199,999	House-holds	128	108	75	108	11,699	0	466	13,822	6.11%

200000+	Annual Income Greater than $200,000	House-holds	128	90	44	131	17,182	0	728	11,499	5.08%
poverty_st	Poverty Status	House-holds	128	662	421	614	376,628	11	3,060	84,789	37.49%
no_high_sc	No High School Education	House-holds	128	367	273	312	97,446	12	1,541	46,983	20.77%
bachelors	Bachelors degree	Labor Force	128	788	690	605	365,580	0	2,664	100,866	32.26%
foreign_bo	Foreign Born	Population	128	1,449	1,315	764	583,559	14	3,593	185,422	29.94%
spanish_sp	Spanish Speaking	House-hold	128	677	436	706	498,677	0	3,107	86,592	38.28%
other_lang	Other Language Speaking	House-hold	128	304	247	253	64,214	0	1,193	38,892	17.19%
age25_29	Age 25 to 29 Total	Population	128	332	311	166	27,721	0	999	42,492	6.86%
age30_34	Age 30 to 34 Total	Population	128	332	307	162	26,398	14	787	42,529	6.87%
race_white	Race White	Population	128	3,062	2,909	1,746	3,048,866	72	7,069	391,947	63.28%
race_black	Race Black	Population	128	557	234	773	597,145	0	3,708	71,282	11.51%
race_asian	Race Asian	Population	128	310	196	347	120,210	0	1,753	39,725	6.41%
race_other	Race Other	Population	128	818	418	935	873,989	0	4,114	104,767	16.92%
race_two	Bi-racial	Population	128	85	62	78	6,090	0	371	10,934	1.77%
1_unit_det	Housing 1 Unit Detached	House-holds	128	713	544	620	384,885	0	2,578	91,232	40.34%
5_9units	Housing 5-9 Units	House-holds	128	105	69	111	12,323	0	489	13,498	5.97%
20_49units	Housing 20-49 Units	House-holds	128	68	26	97	9,375	0	487	8,720	3.86%
50+_units	Housing 50+ Units	House-holds	128	117	31	245	59,865	0	2,035	14,943	6.61%
occupanc_1	Tenure, Occupancy Status, Renter Occupied	House-holds	128	770	807	426	181,738	0	1,656	98,542	43.57%
car_0	Zero Vehicles Available by Worker	Labor Force, Employed	128	211	110	274	75,189	0	1,466	27,048	N/A
car_1	One Car Available by Worker	Labor Force, Employed	128	599	591	314	98,633	0	1,565	76,673	N/A
car_3	Three Cars Available by Worker	Labor Force, Employed	128	341	321	266	70,654	0	1,186	43,652	N/A
car_4	Four Cars Available by Worker	Labor Force, Employed	128	151	122	152	23,035	0	770	19,367	N/A
car_0_hous	Households, Zero	House-holds	128	257	196	228	51,969	0	1,119	32,920	14.55%

	Vehicles Available										
car_1_hous	Households, One Car Available	House-holds	128	626	589	284	80,783	0	1,624	80,128	35.43%
car_3_hous	Households, Three Cars Available	House-holds	128	153	136	123	15,235	0	521	19,520	8.63%
car_4_hous	Households, Four Cars Available	House-holds	128	51	41	48	2,346	0	235	6,469	2.86%
emp_den	Employment/ Area	N/A	128	5,661	4,877	3,899	15,203,719	45	23,315	724,650	N/A
pop_den	Population/Area	N/A	128	12,919	10,903	9,410	88,543,374	1,221	48,091	1,653,656	N/A

Paterson Correlations

Table 11. Paterson Correlations

Variable	Description	Pearson Correlation
manufactur	Employment in Manufacturing	.367**
arts	Employment in Arts	.271**
under_1000	Annual Income Under $10,000	.184*
10000_1499	Annual Income $10,000-$14,999	.255**
15000_1999	Annual Income $15,000-$19,999	.254**
20000_2499	Annual Income $20,000-$24,999	.374**
25000_2999	Annual Income $25,000-$29,999	.368**
30000_3499	Annual Income $30,000-$34,999	.397**
35000_3999	Annual Income $35,000-$39,999	.274**
50000_5999	Annual Income $50,000-$59,999	.222*
125000_149	Annual Income $125,000-$149,999	-.178*
poverty_st	Poverty Status	.347**
no_high_sc	No High School Education	.437**
foreign_bo	Foreign Born	.501**
spanish_sp	Spanish Speaking	.515**
age22_24	Age 22 to 24 Total	.288**
age25_29	Age 25 to 29 Total	.417**
age30_34	Age 30 to 34 Total	.451**
age35_39	Age 35 to 39 Total	.279**
age70_74	Age 70 to 74 Total	-.176*
age75_79	Age 75 to 79 Total	-.235**
race_white	Race White	-.197*
race_black	Race Black	.177*
race_other	Race Other	.498**
race_two	Two or More Races	.217*
1_unit_det	Housing Unit Detached	-.256**
3_4units	Housing 3-4 Units	.381**
5_9units	Housing 5-9 Units	.249**
10_19units	Housing 10-19 Units	.320**
20_49units	Housing 20-49 Units	.366**
50+_units	Housing 50+ Units	.354**
car_0	Zero Vehicles Available by Worker	.569**
car_1	One Car Available by Worker	.498**
car_3	Three Cars Available by Worker	-.282**
car_4	Four Cars Available by Worker	-.230**
car_0_hous	Households, Zero Vehicles Available	.438**
car_1_hous	Households, One Vehicle Available	.342**
car_3_hous	Households, Three Vehicles Available	-.283**
car_4_hous	Households, Four Vehicles Available	-.246**
car_5+_hou	Households, Five or More Vehicles Available	-.179*
emp_den	Employment/Area	.627**

pop_den	Population/Area	.546**

Paterson Regression Model

Table 12. Paterson Regression Model

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	emp_den	.023	.002	.670	12.127	.000	.019	.027
	arts	.395	.075	.291	5.268	.000	.247	.543

APPENDIX D: TRENTON/PRINCETON CASE STUDY

Princeton/Trenton Descriptive Statistics

Table 13. Princeton/Trenton Descriptive Statistics

PR Descriptive Statistics											
	Description	Category	N	Mean	Median	Std. Deviation	Variance	Min	Max	Sum (Freq)	% of Category
total_popu			69	4,609.30	4,502.00	1,851.93	3,429,644.89	579	8,897	318,042	
total_hous			69	1,818.93	1,829.00	745.08	555,138.01	0	3,502	125,506	
unemployme	Unemployed Population	Labor Force	69	203.492754	190	112.372362	12,627.55	0	487	14,041	8.41%
public_tra	Journey to Work by Public Transportation Total	Journey to Work	69	175.289855	116	164.13867	26,941.50	0	727	12,095	8.16%
bus_to_wor	Journey to Work by Public Transportation by Bus or Trolley Bus	Journey to Work, Public Trans.	69	72.1014493	54	92.1547843	8,492.50	0	565	4,975	41.13%
informatio	Employment in Information	Labor Force	69	61.7391304	45	53.1412248	2823.98977	0	169	4,260	2.55%
arts	Employment in Arts	Labor Force	69	144.072464	134	90.4502331	8,181.24	0	471	9,941	5.95%
under_1000	Annual Income Under $10,000	House-holds	69	96.2463768	66	103.065286	10,622.45	0	469	6,641	5.29%
10000_1499	Annual Income $10,000-$14,999	House-holds	69	68.4057971	55	58.3098516	3,400.04	0	249	4,720	3.76%
15000_1999	Annual Income $15,000-$19,999	House-holds	69	59.7101449	54	49.1606315	2,416.77	0	254	4,120	3.28%
25000_2999	Annual Income $25,000-$29,999	House-holds	69	61.1304348	50	49.2161837	2,422.23	0	207	4,218	3.36%
30000_3499	Annual Income $30,000-$34,999	House-holds	69	68.4057971	63	54.2321586	2,941.13	0	232	4,720	3.76%
35000_3999	Annual Income $35,000-$39,999	House-holds	69	54.826087	47	36.7179058	1,348.20	0	155	3,783	3.01%
125000_149	Annual Income $125,000-$149,999	House-holds	69	112.376812	91	97.5806004	9,521.97	0	370	7,754	6.18%
150000_199	Annual Income $150,000-$199,999	House-holds	69	133.043478	81	137.045767	18,781.54	0	554	9,180	7.31%
200000+	Annual Income Greater than $200,000	House-holds	69	173.072464	46	240.685071	57,929.30	0	1061	11,942	9.52%
poverty_st	Poverty Status	House-holds	69	458.913043	326	403.42321	162,750.29	0	1,708	31,665	25.23%

no_high_sc	No High School Education	House-holds	69	257.202899	165	270.752306	73,306.81	0	1,427	17,747	14.14%
bachelors	Bachelors degree	Labor Force	69	1091.21739	800	1014.29301	1,028,790.32	24	3,529	75,294	45.07%
foreign_bo	Foreign Born	Population	69	984.797101	794	705.629164	497,912.52	23	2,939	67,951	21.37%
spanish_sp	Spanish Speaking	House-hold	69	246.826087	67	412.170187	169,884.26	0	2,333	17,031	13.57%
other_lang	Other Language Speaking	House-hold	69	235.014493	160	232.532825	54,071.51	0	924	16,216	12.92%
age25_29	Age 25 to 29 Total	Population	69	300.927536	256	200.437413	40,175.16	8	1108	20,764	6.53%
age30_34	Age 30 to 34 Total	Population	69	304.130435	264	188.743205	35,624.00	38	933	20,985	6.60%
race_white	Race White	Population	69	2,821.55	2,842.00	1,828.56	3,343,619.81	76	6,937	194,687	61.21%
race_black	Race Black	Population	69	993.391304	613	1011.94515	1,024,032.98	7	3,785	68,544	21.55%
race_asian	Race Asian	Population	69	511.985507	177	744.10573	553,693.34	0	3,060	35,327	11.11%
race_other	Race Other	Population	69	181.304348	88	232.233611	53,932.45	0	1,062	12,510	3.93%
race_two	Bi-racial	Population	69	86.7971014	62	79.6659533	6,346.66	0	401	5,989	1.88%
1_unit_det	Housing 1 Unit Detached	House-holds	69	839.67	810.00	645.527592	416,705.87	0	2,681	57,937	46.16%
5_9units	Housing 5-9 Units	House-holds	69	106.710145	57	140.741325	19,808.12	0	685	7,363	5.87%
20_49units	Housing 20-49 Units	House-holds	69	52.7826087	19	77.3781527	5,987.38	0	369	3,642	2.90%
50+_units	Housing 50+ Units	House-holds	69	106.768116	54	146.886325	21,575.59	0	687	7,367	5.87%
occupanc_1	Tenure, Occupancy Status, Renter Occupied	House-holds	69	571.478261	500	396.263786	157,024.99	0	1,588	39,432	31.42%
car_0	Zero Vehicles Available by Worker	Labor Force, Employed	69	127.434783	64	170.555794	29,089.28	0	963	8,793	N/A
car_1	One Car Available by Worker	Labor Force, Employed	69	456.608696	392	288.068988	82,983.74	0	1,200	31,506	N/A
car_3	Three Cars Available by Worker	Labor Force, Employed	69	402.710145	407	283.242367	80,226.24	0	1,212	27,787	N/A
car_4	Four Cars Available by Worker	Labor Force, Employed	69	145.333333	102	152.892952	23,376.25	0	614	10,028	N/A
car_0_hous	Households, Zero Vehicles Available	House-holds	69	198.811594	145	196.134653	38,468.80	0	797	13,718	10.93%
car_1_hous	Households, One Car Available	House-holds	69	552.028986	502	307.646855	94,646.59	0	1,257	38,090	30.35%
car_3_hous	Households, Three Cars Available	House-holds	69	194.347826	185	137.82073	18,994.55	0	561	13,410	10.68%
car_4_hous	Households, Four Cars Available	House-holds	69	56.4637681	38	56.5582936	3,198.84	0	248	3,896	3.10%
emp_den	Employment/Area	N/A	69	2,814.07	1823.46091	3,033.62	9,202,854.26	0	15,049	194,171	N/A

pop_den	Population/Area	N/A	69	7,480.41	4,224.65	11,902.98	141,680,886.94	272.966919	89,976	516,148	N/A

Princeton/Trenton Correlations

Table 14. Princeton/Trenton Correlations

Variable	Description	Pearson Correlation
other_to_w		.315**
bus_to_wor	Journey to Work by Public Transportation by Bus or Trolley Bus	1
agricultur	Employment in Agriculture	.295*
wholesale	Employment in Wholesale	.339**
informatio	Employment in Information	-.280*
arts	Employment in Arts	.294*
under_1000	Annual Income Under $10,000	.332**
10000_1499	Annual Income $10,000-$14,999	.401**
25000_2999	Annual Income $25,000-$29,999	.416**
35000_3999	Annual Income $35,000-$39,999	.243*
75000_9999	Annual Income $75,000-$99,999	-.244*
100000_124	Annual Income $100,000-$124,999	-.237*
125000_149	Annual Income $125,000-$149,999	-.257*
150000_199	Annual Income $150,000-$199,999	-.253*
poverty_st	Poverty Status	.580**
no_high_sc	No High School Education	.630**
high_schoo	High School Education	.378**
foreign_bo	Foreign Born	.285*
spanish_sp	Spanish Speaking	.640**
age22_24	Age 22 to 24 Total	.306*
age25_29	Age 25 to 29 Total	.443**
age30_34	Age 30 to 34 Total	.239*
age60_62	Age 60 to 62 Total	-.320**
age62_64	Age 62 to 64 Total	-.244*
age70_74	Age 70 to 74 Total	-.254*
age80_84	Age 80 to 84 Total	-.257*
race_black	Race Black	.360**
race_other	Race Other	.553**
1_unit_det	Housing Unit Detached	-.396**
2_units	Housing 2 Units	.594**
3_4units	Housing 3-4 Units	.369**
car_0	Zero Vehicles Available by Worker	.790**
car_1	One Car Available by Worker	.298*
car_3	Three Cars Available by Worker	-.285*
car_0_hous	Households, Zero Vehicles Available	.537**
car_2_hous	Households, Two Vehicles Available	-.280*
car_3_hous	Households, Three Vehicles Available	-.323**
car_5+_hou	Households, Five or More Vehicles Available	-.273*
emp_den	Employment/Area	.557**

Princeton/Trenton Regression Model

Table 15. Princeton/Trenton Regression Model

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	car_0_hous	.199	.043	.475	4.639	.000	.113	.284
	age25_29	.124	.033	.385	3.756	.000	.058	.190

APPENDIX E: MPO FORECASTS FOR SJTPO REGION

Table 16. MPO Forecasts, Employment and Population 2010 to 2020

SJTPO –EMPLOYMENT and POPULATION PROJECTIONS						
Municipality	Employement			Population		
	2010		2020	2010		2020
SJTPO REGION	259,782	9.51%	284,483	594,795	6.15%	631,396
Atl (non-farm)	136,800	8.77%	148,796	274,549	7.25%	294,449
Absecon city	3,670	11.74%	4,101	8,411	6.47%	8,955
Atlantic City city	56,263	7.02%	60,213	39,558	-0.79%	39,246
Brigantine city	1,592	6.51%	1,695	9,450	-8.32%	8,664
Buena borough	1,260	6.51%	1,342	4,603	12.05%	5,158
Buena Vista township	1,350	11.74%	1,509	7,570	1.15%	7,657
Corbin City city	514	6.51%	548	492	3.28%	508
Egg Harbor City city	3,125	6.51%	3,329	43,323	18.35%	51,274
Egg Harbor township	14,404	10.65%	15,938	4,243	-2.21%	4,149
Estell Manor city	239	6.51%	254	1,735	6.05%	1,840
Folsom borough	872	11.74%	975	1,885	-1.47%	1,857
Galloway township	8,901	11.74%	9,946	37,349	12.58%	42,048
Hamilton township	10,554	10.65%	11,679	26,503	18.73%	31,467
Hammonton town	8,838	11.74%	9,876	14,791	11.10%	16,432
Linwood city	2,803	6.51%	2,986	7,092	-0.37%	7,066
Longport borough	160	6.51%	170	895	-5.03%	850
Margate City city	1,680	11.74%	1,877	6,354	-7.48%	5,879
Mullica township	615	6.51%	655	6,147	2.37%	6,293
Northfield city	4,077	6.51%	4,342	8,624	7.44%	9,266
Pleasantville city	7,720	11.74%	8,626	20,249	4.16%	21,091
Port Republic city	86	6.51%	92	1,115	4.81%	1,169
Somers Point city	6,137	11.74%	6,857	10,795	-2.35%	10,541
Ventnor City city	1,733	6.51%	1,846	10,650	-5.84%	10,029
Weymouth township	180	6.51%	191	2,715	12.98%	3,067
	Employment			Population		
	2010		2020	2010		2020
Cape May County	41,500	12.80%	46,812	97,265	2.74%	99,928
Avalon borough	1,333	4.12%	1,388	1,334	-9.44%	1,208
Cape May city	5,115	21.80%	6,231	3,607	-2.65%	3,512
Cape May Point borough	163	4.12%	170	291	10.49%	322
Dennis township	1,884	4.12%	1,962	6,467	-0.10%	6,461
Lower township	3,012	12.50%	3,389	22,866	-0.09%	22,846
Middle township	10,741	21.80%	13,083	18,911	15.66%	21,872
North Wildwood city	1,307	4.12%	1,361	4,041	-4.53%	3,858
Ocean City city	5,717	4.12%	5,952	11,701	-5.98%	11,002
Sea Isle City city	1,190	4.12%	1,239	2,114	-6.36%	1,980
Stone Harbor borough	924	4.12%	962	866	-5.81%	816
Upper township	2,970	12.50%	3,342	12,373	6.99%	13,237
West Cape May borough	163	4.12%	170	1,024	-1.62%	1,007
West Wildwood borough	56	21.80%	68	603	17.52%	709
Wildwood city	3,589	4.12%	3,737	5,325	-0.51%	5,298
Wildwood Crest borough	1,361	4.12%	1,417	3,270	-4.46%	3,124
Woodbine borough	1,974	21.80%	2,404	2,472	-2.25%	2,416
	Employment			Population		
	2010	0.50%	2020			
Cumberland County	59,330	8.05%	64,107	156,898	5.99%	166,302
Bridgeton city	10,235	8.82%	11,138	25,349	9.17%	27,674

Commercial township	390	5.95%	413	5,178	-0.51%	5,151
Deerfield township	923	9.84%	1,013	3,119	5.31%	3,285
Downe township	455	9.84%	500	1,585	-0.94%	1,570
Fairfield township	1,021	5.95%	1,081	6,295	0.15%	6,305
Greenwich township	60	5.95%	64	804	-1.69%	790
Hopewell township	105	5.94%	111	4,571	2.50%	4,685
Lawrence township	687	5.95%	727	3,290	16.94%	3,847
Maurice River township	2,544	9.84%	2,794	7,976	12.25%	8,953
Millville city	10,354	5.95%	10,970	28,400	4.69%	29,731
Shiloh borough	88	5.95%	93	516	-1.12%	510
Stow Creek township	325	5.95%	345	1,431	0.11%	1,433
Upper Deerfield township	1,898	5.95%	2,011	7,660	1.11%	7,745
Vineland city	30,245	8.82%	32,913	60,724	6.41%	64,616
	Employment			Population		
	2010	0.50%	2020			
Salem County	22,152	11.81%	24,768	66,083	7.01%	70,717
Alloway township	524	8.76%	570	3,467	14.65%	3,975
Carneys Point township	3,022	16.70%	3,527	8,049	8.36%	8,722
Elmer borough	1,594	8.76%	1,734	1,395	1.48%	1,416
Elsinboro township	152	16.70%	177	1,036	-1.71%	1,018
Lower Alloways Creek township	978	16.70%	1,142	1,770	-0.75%	1,757
Mannington township	1,428	16.70%	1,667	1,806	9.14%	1,971
Oldmans township	525	8.76%	571	1,773	0.70%	1,785
Penns Grove borough	1,119	8.76%	1,217	5,147	6.48%	5,480
Pennsville township	3,526	8.76%	3,835	13,409	11.09%	14,896
Pilesgrove township	1,500	16.70%	1,751	4,016	4.24%	4,186
Pittsgrove township	1,685	8.76%	1,832	9,393	9.74%	10,307
Quinton township	291	16.70%	340	2,666	0.02%	2,666
Salem city	3,164	8.76%	3,441	5,146	-1.51%	5,068
Upper Pittsgrove township	688	8.76%	748	3,505	3.23%	3,618
Woodstown borough	1,886	16.70%	2,201	3,505	8.33%	3,797

APPENDIX F: DATA FORMATS

Survey Data

Survey Data .csv Format											
Columns 1-30			**Columns 31-61**			**Columns 62-92**			**Columns 93-123**		
Col	**Column Name**	**Example Data**	**Col**	**Column Name**	**Example Data**	**Col**	**Column Name**	**Example Data**	**Col**	**Column Name**	**Example Data**
1	ID	1	**31**	EGRESS_OTHER		**62**	AGE_MIDPT	58	**93**	ON_MAT_TYPE	
2	SURVEYNAME	0	**32**	DESTSTREET	723 SOUTH BROAD ST	**63**	INCOME_MIDPT	42500	**94**	OFF_MAT_LAT	40.20391
3	DIRECTION	2	**33**	DESTTOWN	TRENTON	**64**	O_MAT_LAT	40.27388	**95**	OFF_MAT_LONG	-74.7426
4	MILITARYSTARTTIME	8:30:00 PM	**34**	DESTSTATE	NJ	**65**	O_MAT_LONG	-74.7251	**96**	OFF_FIPS_ST	34
5	AMPM	2	**35**	DESTZIP	8611	**66**	O_FIPS_ST	34	**97**	OFF_FIPS_CTY	21
6	BUSROUTE	603	**36**	DESTPLACE	1	**67**	O_FIPS_CTY	21	**98**	OFF_CEN_TRCT	4
7	ORIGINPLACE	8	**37**	CASINOWORKER	Err:512	**68**	O_CEN_TRCT	32.02	**99**	OFF_CEN_BLK	3041
8	ORIGINSTREET	MERCER MALL	**38**	CASINOVISITOR	Err:512	**69**	O_CEN_BLK	7004	**100**	OFF_FIPS_PLC	74000
9	ORIGINTOWN	LAWRENCEVILLE	**39**	CAPTIVITY	1	**70**	O_FIPS_PLC		**101**	OFF_FIPS_MCD	74000
10	ORIGINSTATE	NJ	**40**	VEHICLEAVAIL	2	**71**	O_FIPS_MCD	39510	**102**	OFF_FIPS_MSA	8480
11	ORIGINZIP	8648	**41**	TICKETTYPE	1	**72**	O_FIPS_MSA	8480	**103**	OFF_MAT_CENT	0
12	ACCESSMODE	1	**42**	TRIPFREQ	6	**73**	O_MAT_CENT	X	**104**	OFF_CBSA	
13	ACCESS_BUS		**43**	TRIPTENURE	6	**74**	O_CBSA		**105**	OFF_METROD	
14	ACCESS_TRAIN		**44**	QUALSERVCHG	3	**75**	O_METROD		**106**	OFF_MICROF	
15	ACCESS_PATCO		**45**	GENDER	1	**76**	O_MICROF		**107**	OFF_MAT_STAT	B3
16	ACCESS_SEPTA		**46**	AGE	6	**77**	O_MAT_STAT	15	**108**	OFF_MAT_TYPE	
17	ACCESS_OTHER		**47**	RACE	1	**78**	O_MAT_TYPE		**109**	D_MAT_LAT	40.20811
18	BUS_ONSTREET	MERCER MALL	**48**	RACE_OTHER		**79**	ON_MAT_LAT	40.27388	**110**	D_MAT_LONG	-74.7527
19	BUS_ONTOWN	LAWRENCE	**49**	HISPANIC	2	**80**	ON_MAT_LONG	-74.7251	**111**	D_FIPS_ST	34
20	BUS_ONSTATE	NJ	**50**	OCCUPATION	2	**81**	ON_FIPS_ST	34	**112**	D_FIPS_CTY	21
21	BUS_ONZIP	8648	**51**	HOUSEHOLDSIZE	1	**82**	ON_FIPS_CTY	21	**113**	D_CEN_TRCT	8
22	BUS_OFFSTREET	S BROAD ST AND LIBERTY AVE	**52**	HOUSEHOLDEMPLOYED	1	**83**	ON_CEN_TRCT	32.02	**114**	D_CEN_BLK	2026
23	BUS_OFFTOWN	TRENTON	**53**	HOUSEHOLDCARS	999	**84**	ON_CEN_BLK	7004	**115**	D_FIPS_PLC	74000
24	BUS_OFFSTATE	NJ	**54**	INCOME	4	**85**	ON_FIPS_PLC		**116**	D_FIPS_MCD	74000
25	BUS_OFFZIP	8611	**55**	CONTACTSTREET	723 SOUTH BROAD ST	**86**	ON_FIPS_MCD	39510	**117**	D_FIPS_MSA	8480
26	EGRESSMODE	1	**56**	CONTACTAPT		**87**	ON_FIPS_MSA	8480	**118**	D_MAT_CENT	0
27	EGRESS_BUS		**57**	CONTACTTOWN	TRENTON	**88**	ON_MAT_CENT	X	**119**	D_CBSA	
28	EGRESS_TRAIN		**58**	CONTACTSTATE	NJ	**89**	ON_CBSA		**120**	D_METROD	
29	EGRESS_PATCO		**59**	CONTACTZIP	8611	**90**	ON_METROD		**121**	D_MICROF	
30	EGRESS_SEPTA	Err:512	**60**	WEIGHT	4.5	**91**	ON_MICROF		**122**	D_MAT_STAT	B1
			61	FREQUENTRIDER	0	**92**	ON_MAT_STAT	15	**123**	D_MAT_TYPE	

Farebox Data Format

Column Name	Example Data
LINE	603
RUN	34
TRIP	1
PATTERN	37
TIME_PERIOD	236p-357p
BOARDING_ZONE	2
ALIGHTING_ZONE	2
TOTAL_TRANSACTIONS	35
COUNTER	2
CASH	19
TICKET	1
TEN_TRIP	0
CONTINUE_TRIP	0
CASH_RT	0
MONTHLY_PASS	10
TRANSFER1_RECEIVED	1
TRANSFER1_ISSUED	2
SPECIAL	0
OVERRIDE	0
TRANSFER2	0
RAIL_PASS_CH	0
NONE	0
ADULT	30
CHILD	0
SENIOR_HANDICAPPED	2
STUDENT	1
EMPLOYEE	0
FAMILY	2
FOREIGN	0
TEN_TRIP_CLASS	0
DATE	7/8/2013

APPENDIX G: LITERATURE REVIEW

Annotated Bibliography

Kimpel, Thomas. (2001). *Time Point-level Analysis of Transit Service Reliability and Passenger Demand.* An unpublished dissertation. [http://www.pdx.edu/sites/www.pdx.edu.cus/files/SR036.pdf]

Archived Tri-Met Bus Dispatch System data, bus transit performance, passenger activity, socioeconomic data, and land use information were used to analyze transit service reliability and passenger demand at the time point level.

- Study site
    - Portland, Oregon (Tri-Met)
- Relevant findings
    - Crosstown demand models behave differently from radial line routes.
    - The only explanatory variables that appear to matter for crosstown routes were number of zero auto-ownership households and the particular segment of the route (*see* p. 121)

McKenzie, Brian. (2011). Transit Access and Labor Market Outcomes across Segregated Neighborhoods. An unpublished dissertation.

To understand the role of public transportation and low-income and residentially segregated residents, this research examined Latino and segregated black neighborhoods using transit access measures; variation in transit access across neighborhood types with respect to labor market outcomes; and a Geographic Information Systems (GIS) transit access measure was developed and applied in Portland, Oregon.

- Study site
    - Portland, Oregon (Tri-Met)
- Relevant findings
    - Including data on sidewalk networks yielded improvement in transit analysis (p.175)

Peng, Zhongren. (1994). *A Simultaneous Route-level Transit Patronage Model: Demand, Supply and Inter-route Relationship.* An unpublished dissertation. [http://www.pdx.edu/sites/www.pdx.edu.cus/files/SR027.pdf]

To better understand the relationship between transit patronage and service levels, a quantitative model was developed that incorporates the interactions of this relationship using a simultaneous system approach including: a demand equation, a supply equation and an equation for competing routes, using a three-stage-least-squares estimation method.

- Study site
    - Portland, Oregon (Tri-Met)
- Relevant findings
    - Inter-relationships among transit routes: indifference, complementary, competing and synergistic
    - It is important to understand how routes relate to each other when forecasting at the route and stop level (*see* p.103)

Transit Technologies

Puchalsky, C., Joshi, D., and W. Scherr. (2012). *Development of a Regional Model Based on Google Transit Feed Specification.* A paper presented at the 13th TRB Planning Application Conference, May 2011, Reno, NV.

- Study site
    - Delaware Valley Regional Planning Commission (DVRPC) nine county area including Bucks, Chester, Delaware, Montgomery and Philadelphia in Pennsylvania; and Burlington, Camden, Gloucester and Mercer in New Jersey.
- Modeling advancement
    - Use of Google Transit Feed Specifications (GTFS), OpenStreetMaps, and other "open" data sources and technologies in forthcoming TIM 2.0

Wong, J. (2013). *Leveraging the General Transit Feed Specification (GTFS) For Efficient Transit Analysis.* A paper presented at the 92nd Transportation Research Board Annual Meetings, January 13 – 17, 2013, in Washington D.C.

- Study site
    - Fifty largest transit agencies with GTFS and Southeastern Pennsylvania Transit Authority (SEPTA)
- Methodological contribution
    - Provides a framework for suggested analyses of agency operations at the stop, route, and system level using GTFS

Wong, J., Reed, L., Watkins, K., and R. Hammond. (2013). *One Transit Data: State of the Practice and Experiences from Participating Agencies in the United States. .* A paper presented at the 92nd Transportation Research Board Annual Meetings, January 13 – 17, 2013, in Washington D.C.

- Study site
    - Case studies: Southeastern Pennsylvania Transit Authority (SEPTA), Chicago Transit Authority (CTA), Bay Area Rapid Transit (BART), the New York Metropolitan Transportation Authority (NYMTA), the Massachusetts Bay Transportation Authority (MBTA) and the Metropolitan Atlanta Rapid Transit Authority (MARTA)
- Findings
    - Use of Open Transit Data has positive results for agencies, but no quantitative or formal studies regarding impact on ridership or customer satisfaction.

Research Theories in Practice

Brown, J., Thompson, G., Bhattacharya, T., and M. Jaroszynski. (2013) Understanding Transit Ridership Demand for the Multi-Destination, Multi-Modal transit Network in Atlanta, Georgia: Lessons for Increasing Rail Transit Choice Ridership While Maintaining Transit-Dependent Bus. A paper presented at the 92nd Transportation Research Board Annual Meetings, January 13 – 17, 2013, in Washington D.C.

- Study site
    - Atlanta, Georgia (MARTA)
- Research findings
    - Bus rider came from zones with lower incomes, lower vehicle access and higher minority populations
    - Rail riders also came from disproportionately from zones with high minority populations, but these zones had greater vehicle access and income variables were not significant, except where zones included more dispersed destinations, then low income status is significant

Dill, J., Scholossberg, M., Ma, L., and C. Meyer. (2013). *Predicting Transit Ridership at the Stop Level: The Role of Service and Urban Form*. A paper presented at the 92nd Transportation Research Board Annual Meetings, January 13 – 17, 2013, in Washington D.C.

- Study site
    - Portland, Oregon (Tri-Met), Eugene, OR (Lane Transit District), and Jackson County, OR (Rogue Valley Transit District)

- Research Findings
    - Transit level of service characteristics are the most important factors for determining ridership at the stop level, followed by socio-demographics and land use variables , respectively for large and medium-sized transit systems
        - For a small transit system, land use variables explain more than socio-demographic variables

Lee, S., Hickman, M., and D. Tong. (2013). *A Time-Varying Route-level Transit Patronage Model*. A paper presented at the 92nd Transportation Research Board Annual Meetings, January 13 – 17, 2013, in Washington D.C.

- Study site
    - Minneapolis/St. Paul, MN (Metro Transit)
- Research findings
    - Commercial and institutional variables were statistically significant indicating important demand generators for transit services
        - Ridership increases over the day in commercial land uses (AM to Midday to PM)
        - Ridership increases over the day in institutional land uses, particularly in the PM

Background Research

Antrim, A., & Barbeau, S. J. (2013). The many uses of GTFS data–opening the door to transit and multimodal applications. *Location-Aware Information Systems Laboratory at the University of South Florida*.

- Authors expound upon the numerous and ever-growing uses of General Transit Feed Specification (GTFS) data. Of note are applications that assist in data visualization, planning and analysis, and real-time monitoring. History and function of GTFS offered; resources for additional information provided.

Frei, C. and H. Mahmassani. (2013). *Riding More Frequently: Disaggregate Ridership Elasticity Estimation for a Large Urban Bus Transit Network.* A paper presented at the 92nd Transportation Research Board Annual Meetings, January 13 – 17, 2013, in Washington D.C.

- Study site
    - Chicago, IL (Chicago Transit Authority)
- Research findings
    - When transit stops are located near medical facilities, ridership increase is higher in the medium and long term
    - With respect to stability over the course of the day, elasticities are less for industrial, medical, recreational and educational areas than for other types of land uses

Furth, P. (2000). *TCRP Synthesis 3: Data Analysis for Bus Planning and Monitoring*. Transportation Research Board, National Research Council, Washington, D.C. <http://onlinepubs.trb.org/onlinepubs/tcrp/tsyn34.pdf>

- Research findings
- Survey of 20 transit agencies regarding their use of automatic vehicle location (AVL) equipment, automatic passenger counters (APC), and trip time analyzers for data analysis and bus planning purposes

Lawson, C. T. (2006). Microsimulation for urban transportation planning: Miracle or mirage?. *Journal of Urban Technology*, *13*(1), 55-80.

- The author briefly reviews the history of travel demand modeling and shows how stricter federal requirements and advances in computing technology are driving modeling innovations. One such innovation is TRANSIMS, a traffic microsimulation tool. Numerous experimental studies using TRANSIMS are presented. Interwoven are explanations of TRANSIMS architecture. Comparison to traditional four-step model is drawn throughout. Author points to several features that make TRANSIMS a superior transportation planning tool-- high spatiotemporal resolution of individuals moving through simulated area, high degree of internal consistency and feedback, and open source code permitting pan-disciplinary review and revision. The barriers and opportunities for TRANSIMS implementation are noted.

Lee, S., Tong, D., and M. Hickman. (2013). *A Comparative Study of Alternative Methods for Generating Route-level Mutually Exclusive Service Areas*. A paper presented at the 92nd Transportation Research Board Annual Meetings, January 13 – 17, 2013, in Washington D.C.

- Study site

- Minneapolis-St. Paul, MN (Metro Transit)
- Research findings
  - On-board surveys were used to better understand willingness-to-walk using both Network Distance-based Service Area (NDSA) and Combination of Thiessen Polygon and Buffer (CTPB) methods
    - NDSA is not necessarily better for mutually exclusive service areas at the route-level

Liebig, T., Piatkowski, N., Bockermann, C., & Morik, K. (2014, March). Predictive Trip Planning-Smart Routing in Smart Cities. In *EDBT/ICDT Workshops* (pp. 331-338).

- Authors created a route planning architecture that uses an Open Trip Planner web interface and real-time processing of data from traffic sensors to generate traffic flows for "unobserved locations at future times." The intention is to create a routing system that can anticipate future traffic ebbs and flows and re-route users around areas of high traffic. Authors applied system to a use-case in Dublin, Ireland. Future areas of research were identified, including different kernels for flow estimation, additional transportation data sources, and hazard and weather data sources.

Sun, D.J., Peng, Z.R., Shan, X., Chen, W., & Zeng, X. (2011). Development of Web-Based Transit Trip-Planning System Based on Service-Oriented Architecture.

- The majority of transit trip planners exist as proprietary systems based on particular vendor products. With the incorporation of more functional components, system maintenance and regular transit information updates become burdensome tasks for transit agencies. In addition, the proprietary nature of the systems makes it difficult to take advantage of the rapid advancement of geospatial information and web technologies. The authors proposed an open and interoperable transit trip-planning system based on a service-oriented architecture, with the principle of reusing the existing modular resources, while providing user-friendly interfaces for expansion of functionality. The objective was to integrate geospatial services available online (such as Google Maps), open-source geospatial database technologies, and path-finding algorithms in a loosely coupled manner. The proposed system was developed with spatial and temporal transit data from Waukesha Metro Transit in Wisconsin. Research results were validated by comparing outputs from the existing South-East Wisconsin Transit Trip Planner and route schedule matching. Comparison results showed that the new service-oriented architecture provided a flexible, efficient mechanism for transit-trip planners. The architecture took advantage of rapidly changing online geospatial services, yet maintained the core functions of itinerary search that may be unique to each transit agency.

Verbas, I., Frei, C. Mahmassani, H., and R. Chan. (2013). *Stretching Resources: Sensitivity of Optimal Bus Frequency Allocation to Stop-Level Demand Elasticities.* A paper presented at the 92nd Transportation Research Board Annual Meetings, January 13 – 17, 2013, in Washington D.C.

- Study site
  - Chicago, IL (Chicago Transit Authority)
- Research Findings
  - Multiple scenarios illustrated ridership with respect to headways

- – Cross-elasticities for routes and alternative modes, spatio-temporal elasticities in the Transit Network Design Problem (TNDP) can capture unique attributes of individual locations as well as the relationships among trip destinations

Vij, A., and J. Walker. (2013). *You Can Lead Travelers to the Bus Stop, But You Can't Make Them Ride.* A paper presented at the 92nd Transportation Research Board Annual Meetings, January 13 – 17, 2013, in Washington D.C.

- Study sites
  - San Francisco, CA (Bay Area Travel Survey – BATS)
  - Karlsruhe, Germany (MOBI*DRIVE*)
- Research Findings
  - Mode share analysis found that incremental improvements in the transportation system, without corresponding shifts in "individual modality styles" will result in far smaller changes in travel behavior than traditional models would predict