PREDICTING TRANSIT RIDERSHIP AT THE STOP LEVEL: THE ROLE OF SERVICE AND URBAN FORM

Jennifer Dill (corresponding author)
Nohad A. Toulan School of Urban Studies & Planning
Portland State University
506 SW Mill Street, Suite 350
Portland, OR 97201
E-mail: jdill@pdx.edu
Phone: 503-725-5173, Fax: 503-725-8770

Marc Schlossberg
Department of Planning, Public Policy & Management
University of Oregon
1209 University of Oregon
Eugene, OR 97403-1209
E-mail: schlossb@uoregon.edu

Liang Ma
Nohad A. Toulan School of Urban Studies & Planning
Portland State University
506 SW Mill Street, Suite 320
Portland, OR 97201
E-mail: liangm@pdx.edu

Cody Meyer
University of Oregon
Department of Planning, Public Policy & Management
1209 University of Oregon
Eugene, OR 97403-1209
E-mail: codemeyer@gmail.com

Submitted for Presentation at the 92nd Annual Meeting of the Transportation Research Board

Word Count: 5,916 words + 7 Tables = 7,666 total

ABSTRACT

This research aims to better understand the relative and combined influence of transit service characteristics and urban form on transit ridership at the stop level. Three metropolitan regions in Oregon were included in the analysis, representing different types of communities. We use stop-level ridership data from 7,214 TriMet stops in the Portland region, 1,400 Lane Transit District (LTD) stops in the Eugene-Springfield and 350 Rogue Valley Transit District (RVTD) stops in Jackson County (Medford-Ashland area) as the dependent variable for regression models. Categories of independent variables tested include: (1) socio-demographics; (2) transit service characteristics (e.g. headways, hours of service, transfer stops, bus vs. light rail, etc.); (3) land use (employment, population, land use type, pedestrian destinations, etc.); and (4) transportation system (e.g. street connectivity, bike lanes, etc.). The final model results indicate that the TriMet model does a better job explaining the variation in ridership at the stop-level; the adjusted-R^2 is 0.69, compared to 0.61 for the LTD model, and 0.53 for the RVTD model. Land use characteristics around transit stops do have significant effects on transit ridership, though these effects are much smaller than the effects of transit level of service. Socio-demographic characteristics seem to have a larger effect on ridership in the large urban area than small urban areas (TriMet: 24% vs. LTD and RVTD: 11%). The land use characteristics have much smaller effect in large urban area than small urban area (TriMet: 5% vs. RVTD: 18%).

INTRODUCTION

This research aims to better understand the relative and combined influence of transit service characteristics and urban form on transit ridership at the stop level. Most previous work in this area has looked at these issues separately. On the one hand, there has been work on the system performance of transit (e.g. on-time performance, cost, etc.) and on the other hand there has been a recent flurry of research exploring the connection between urban form and transit or pedestrian travel. This project seeks to synthesize these disparate approaches, recognizing that while transit service characteristics (e.g. frequency, travel time, etc.) are important, most transit users are pedestrians at the beginning and end of any transit trip. Therefore, focusing also on the walkable zone around each transit stop is critically important.

Three metropolitan regions in Oregon were included in the analysis, representing different types of communities. TriMet serves the largest (approximately 1.8 million population) metropolitan area in the state, Portland. Lane Transit Distrist (LTD) serves the medium-sized Eugene-Springfield area, with a population of about 250,000. Rogue Valley Transit District (RVTD) is in the smaller urbanized area of Medford and Ashland, with a population about 150,000. In addition, there are very different built environment conditions within each metropolitan area.

We use stop-level ridership data from 7,214 TriMet stops in the Portland, OR region, 1,400 Lane Transit District (LTD) stops in the Eugene-Springfield, OR, and 350 Rogue Valley Transit District (RVTD) stops in Jackson County, OR as the dependent variable for regression models. Categories of independent variables tested include: (1) socio-demographics; (2) transit service (headways, hours of service, transfer stops, park-and-ride lots, bus vs. light rail, etc.); (3) land use (employment, population, land use type, land use mix, pedestrian destinations, parks, etc.); and (4) transportation system (e.g. street connectivity, bike lanes, etc.). The remainder of the paper is structured as follows: literature on linking urban form and transit ridership will be reviewed first, and then the research methodology and data will be introduced. The final section discusses and explains the model results and implications for public transit and land use policy.

RESEARCH LINKING URBAN FORM AND TRANSIT RIDERSHIP

Many previous empirical studies focus on transit ridership at the route-level and segment-level and largely assume homogeneous service levels and land use along each route *[1]*. However, these assumptions are not valid, especially for the routes that cross areas with dramatic changes in land use as well as social-demographic characteristics, for example, from central business districts (CBD) to suburban areas. Therefore, stop level transit demand models are needed to take into account stop-level land use characteristics, such as the surrounding pedestrian environment. Stop-level models are particularly useful to connect transit demand with demographic, service and land use characteristics *[2]*. Previous research linking land use and transit ridership at the stop level is somewhat limited. TABLE 1 lists the stop-level studies we identified. The following section focuses on the built environment and level of service variables used in these studies.

TABLE 1 Existing Research with Stop-level Transit Ridership Models

Sources	Title	Transit Type	Location of Study
Banerjee, Myers, and Irazabal [4]	Increasing Bus Transit Ridership: Dynamics of Density, Land Use, and Population Growth	Rapid Bus	Los Angeles, California
Cervero, Murakami, and Miller [12]	Direct Ridership Model of Bus Rapid Transit in Los Angeles County, California	Bus Rapid Transit (BRT)	Los Angeles County, CA
Cervero [5]	Alternative Approaches to Modeling the Travel-Demand Impacts of Smart Growth	Heavy Rail; Light Rail	San Francisco Bay Area; St. Louis
Chu [1]	Ridership Models at the Stop Level	Bus	Jacksonville, Florida
Estupinan and Rodriguez [9]	The Relationship Between Urban Form and Station Boardings for Bogota's BRT	Bus Rapid Transit (BRT)	Curitiba, Bogota
Lin and Shin [6]	Does Transit-Oriented Development Affect Metro Ridership? Evidence from Taipei, Taiwan	Heavy rail	Taipei, Taiwan
Pulugurtha and Agurla [14]	Assessment of Models to Estimate Bus-Stop Level Transit Ridership using Spatial Modeling Methods	Bus	Charlotte, NC
Ryan and Frank [13]	Pedestrian Environments and Transit Ridership	Bus	San Diego, California

Built Environment Variables

Researchers have often used the 3Ds to describe the built environment: density, diversity and design *[3]*. The findings with respect to 3Ds variables from the studies examined appear in Table 2.

Several aspects of density around transit stops are commonly used, including population density, employment density, housing density, and building density. Density is generally assumed to have positive correlation with transit ridership, and several empirical studies did find this relationship was significant *[1, 4, 5, 6]*. However, density itself may be too broad to capture the micro-scale built environment factors which may be more essential to the transit ridership.

Land use mix refers to the level of diversity of land uses in a given area. The relationship between the land use mix around transit stops and transit ridership is not clear. Even though many studies have shown that residents living in a mixed land use environment would be more likely to use transit than residents in a primarily residential neighborhood (e.g. *[7]*), few stop-level studies examined the relationship between the land use mix and transit ridership. Jobs-housing balance, entropy, and the proportion of each type of land use are common ways to measure land use diversity in a model. Among the studies reviewed, Lin and Shin *[6]* and Cervero *[5]* did not find a significant relationship between land use mix and transit ridership. By contrast, Banerjee et al. *[4]* found significant and positive relationship between percentage of non-residential land use and rapid bus ridership. They also found that land use diversity was significant, having a positive relationship with rapid transit ridership when tested alone. However, in a model testing the effects of both population density and land use mix, land-use mix or diversity had no significant effect. One of the reasons for the insignificant relationship between land use mix and transit ridership may be the methods these studies used to create the land use mix variables. Variables that use entropy as a measure, which is common, may not be measuring land use types at the right scale or level. Entropy measures are typically calculated at an aggregate level, e.g. residential, commercial, industrial, etc. There are a wide variety of uses within each of those categories that likely have differing effects on transit ridership. Consider, for example, the difference between a big-box home improvement store and an office building, both of which fall into the commercial land use category. Moreover, the impact of land use mix on transit use was found to be greater at employment destinations than at residential origins *[8]*. Having a mix of uses in close proximity to an employment destination facilitates people who use transit to commute to be able to walk to lunch or to run errands.

Design features may also affect ridership by making the accessibility conditions of station/stop area more or less attractive. Estupinan and Rodriguez *[9]* found that street connectivity had significantly positive relationship with transit ridership, while a negative correlation was found by Lin and Shin *[6]*. A research team from Department of City and Regional Planning at University of North Carolina *[10]* evaluated the micro accessibility environment, road design, pedestrian/bicycle environment, and architecture design at the stop level though auditing. They concluded that: bus stop amenities, such as having signs, shelters, schedules, lighting, and paved landing areas were significantly and positively correlated with increased ridership; pedestrian/bicycle friendly design was positively associated with ridership; and buildings designed with interesting features are likely to encourage ridership. Estupinan and Rodriguez *[9]* also employed an audit score to evaluate the design around BRT stations and concluded that walk/bike friendly design around station contributed positively to BRT ridership.

TABLE 2 Built Environment Variables Found in Existing Research

Built Environment Variables	Method to Create the Variable (Sources)	Relationship with Transit Ridership
Density		
Population Density	Number of population within the buffer area [5, 12, 15]	+
Employment Density	Number of employees/area of working floor space within the buffer area [6, 12, 15]	+; ns
Building Density	Area of floor space within the buffer area ([6])	+
Housing densities	Number of dwelling units within the buffer area [5, 13]	+; ns
Total Density	Total employment plus population within the buffer area [5, 12]	+
Diversity		
Residential Area	Residential land use area within the walkable distance from a bus stop ([14])	-
Industrial Area	Industrial land use area within the walkable distance from a bus stop ([14])	-
Commercial Area	Commercial land use area within the walkable distance from a bus stop ([14])	+
Institutional Area	Institutional land use area within the walkable distance from a bus stop ([14])	+
Land Use Mix	Proportion of seven land use types within station area (Ryan and Frank, 2009); Land use index (0-100) Audit ([9]); Entropy (Cervero, 2006 [5]); Land Use Diversity = 1- [Sum (Ia_1 , Ia_2 , Ia_3 , …….Ia_n)] . : area of each type of land use, A: total land area ([4])	ns; - ; +
Pedestrian Amenities	Index of amenities (0-100) Audit ([9])	+
Job-Housing Balance	Job-Housing balance= 1-[absolute value (Total employment-1.5 x Total housing units)/(Total employment+1.5 x Total housing units) ([15])	ns; -
Percentage of Retail and Service Floor Space	Area of retail and service floor space/area of total floor space ([6])	ns
Design		
Walkability Index	2x[Z(Land use mix]+Z(Residential Density)+Z(Retail FAR)+Z(Intersection Density)] ([13])	+
Street Connectivity	Number of blocks ([6])	ns
	Number of intersections/number of links ([12])	ns
Walking Support	Factor analysis of Bike Path, Sidewalk, Traffic Control, Sidewalk Continuity, Sidewalk Width, Sidewalk Quality, Amenities, Street Connectivity, Road Density ([9])	+
Barriers to Car Use		+
Safety and Security		+
Street Connectivity		+
Sidewalks	Length of sidewalk ([6])	ns
	Percentage of arterials and collectors with sidewalk in quarter mile around bus stops in a TAZ ([15])	+
	Percentage of street lengths with sidewalk in the quarter mile buffer around bus stop ([15])	+
Parking	Number of parking spaces/area of floor space ([6])	ns
Pedestrian Factor	Traffic signal in immediate vicinity; Median type; Number of lanes on street; Pedestrian street-crossing delay; TLOS pedestrian adjustment factor; P.M. peak hour traffic volume; Presence of continuous sidewalk in stop vicinity. ([1])	+

Notes: +: significantly positive relationship; -: significantly negative relationship; ns: no significant relationship was found

Transit Level of Service Variables

In the studies examined, transit level of service was primarily assessed by transit frequency, transit alternatives, and route density, which all proved to have significant and positive relationships with ridership (TABLE 3). Mishra et al. *[11]* estimated the connecting power of a transit line at a node by a function of the average vehicle capacity of the transit line, the frequency on the transit line, the daily hours of operation of the transit line, the speed of the transit line, and the distance of the node to the destination. Cervero *[12]* developed a Direct Ridership Model to predict the average daily boardings of 69 BRT bus stops in Los Angeles County. His model found that service quality (e.g. number of daily buses, number of feeder connections) positively contributed to ridership. Ryan and Frank *[13]* developed a measure of level of service to capture the level of transit accessibility to multiple destinations as well as the amount of waiting time between buses, and found that places with more routes and shorter wait times had higher bus ridership. Estupinan and Rodriguez *[9]* predicted BRT ridership using five LOS variables: 1) number of bus transit alternatives to BRT; 2) presence of a feeder bus; 3) number of routes, 4) types of station defined by size; and 5) number of vehicles per day per station. All five were significantly and positively correlated with BRT ridership. Cervero *[5]* estimated the peak-hour rail station boardings at San Francisco Bay Area, and found that train frequency and feeder bus service were positively and significantly associated with station boardings. Banerjee et al. *[4]* used the number of transit linkages with the availability of metro rail at a bus stop as measures of level of service to predict rapid bus ridership. The study found that these two variables had significant, positive effects on bus ridership.

TABLE 3 Variables Measuring Transit Level of Service Found in Existing Research

Sources	Level of Services Variables	Relationship with Transit Ridership
Cervero, Murakami, and Miller [12]	Number of daily metro rapid buses (both directions)	+
	Number of perpendicular daily feeder bus lines (both directions)	+
	Number of perpendicular daily rail feeder trains	+
Ryan and Frank [13]	Numbers of bus routes serving a bus stop divided by the mean wait time of all route serving the bus stop	+
Estupinan and Rodriguez [9]	Transit Supply—number of bus transit alternatives available different from BRT; Presence of feeder bus; number of Routes; Types of Station defined by size; Number of vehicles per day per station	+
Cervero [5]	Service Frequency: number of train cars in one direction	+
	Feeder Bus Service: number of feeder buses arriving at station	+
Chu [1]	LOS within one-minute walking	+
	LOS within two-five minutes walking	+
	Number of other TLOS stops in catchment area	-
Zhao et al [15]	Composite average peak hour headway	-
	Average number of bus runs per stop	+
	Percentage of TAZ area served by transit based on quarter mile buffers around bus stops	+
	Bus Route Density in feet per acre in a TAZ	+
	Number of Bus Routes in a TAZ	+
Banerjee, Myers, and Irazabal [4]	Number of transit linkages	+
	Availability of metro rail	+

Notes:
+: significantly positive relationship
-: significantly negative relationship
ns: no significant relationship was found
Blank cell means the variable was not included into the final model.

METHODOLOGY

Model Specification

Multivariate linear regression was employed to estimate the relative effects of socio-demographics, land use, transportation infrastructure, and transit service characteristics in predicting transit ridership at each stop. Because boardings (getting on transit) and alightings (getting off transit) are "count" data, and the distribution of count data can be skewed toward the origin (zero), it is not reasonable to use ridership data directly as the dependent variable in linear model due to the violation of a major assumption of OLS. Therefore, a logarithm transformation of ridership data was used. We also tested count data models, such as Poisson and Negative Binomial Regression models. The results of these models were very similar to the results of the linear models using the logarithm transformation, and we did not find any advantages to use count data model to predict transit ridership in this case.

We estimated separate models for each region. All the variables we created were entered into the model at the beginning, and different combinations of these variables were tested before we determined the final models based upon goodness-of-fit statistics (adjusted R^2). We eliminated variables that were highly correlated with one another, as well as variables that were not significant in any of the models. However, for comparison purposes, if a variable was significant in one model, we kept it in the other models. With a few exceptions, all of the variables were based on 2008 data (TABLE 4). In both the TriMet and LTD areas, network and circular-based buffers at quarter-mile and half-mile distances were developed around each stop. Network buffers differ from circular buffers in that they measure the distance away from each stop along the street network. The resulting polygon is often irregular-shaped due to the non-uniform street network pattern, thereby encompassing some aspect of the urban form within the spatial unit of analysis. After comparing the results across all four methods (circular and network buffers at both quarter- and half-mile distances), and with an eye toward keeping analysis approaches as simple as possible for easy replication, we settled on using quarter-mile circular buffers in the analysis of RVTD. Pulugurtha and Agurla (2012) also tested different buffer sizes and concluded that one-quarter mile was the best predictor of ridership. In addition, one of the independent variables, street connectivity, is the spatial characteristic that makes the network-based buffer different than a circular buffer. Therefore including both street connectivity and network buffers may be unduly repetitive.

TABLE 4 Variable statistics

	TriMet		LTD		RVTD	
	Mean	**s.d.**	**Mean**	**s.d.**	**Mean**	**s.d.**
Dependent Variables						
Total Riders	187	768	324	1562	22	125
Log Transformation of Total Rider	3.3	2.1	4.3	1.6	2.2	1.2
Socio-Demographic Variables						
% of female population	50.2%	5%	50.8%	5%	51.1%	6%
% of white population	81.1%	11%	87.4%	7%	91.4%	5%
% of population below 17	20.8%	7%	18.9%	8%	21.4%	7%
% of population aged 18-25	9.0%	5%	18.3%	16%	10.6%	6%
% of population aged 65 or older	10.8%	5%	12.9%	7%	15.0%	7%
% of population with college degree	26.7%	15%	19.3%	11%	13.1%	8%
Median family income (annual, $000)	70.2	25.9	55.2	16.8	47.7	11.5
% of households without vehicle available	10.5%	11%	11.4%	11%	8.9%	7%
% of households with annual HH income below the poverty level	12.8%	8%	21.0%	15%	16.8%	9%
Transit Service Variables						
Rail transit/BRT stations (0=bus stop)	1.6% of stops		0.7% of stops			
Transfer stop (1=yes)	21.9% of stops		53.9% of stops		3.3% of stops	
Transit center (1=yes)	1.3% of stops		2.9% of stops		0.3% of stops	
Average headway (minutes)	28	15	36	18	34	9
Maximum coverage time (minutes)	1,036	234	818	287	766	62
Total bus stops within buffer	16	21	8	6	5	2
Total light rail stations within buffer	0	1				
Park & Ride for bus and LRT/BRT (1=yes)	0.4% of stops					
Park & Ride for bus only (1=yes)	1.3% of stops		3.7% of stops		2.2% of stops	
Transportation Infrastructure Variables						
Street Connectivity (number of 3+-way intersections)	30	17	32	23	21	14
Miles of regional multi-use paths	0.1	0.2				
Miles of bike lanes	0.4	0.4	0.5	0.5	0.3	0.3
Land Use Variables						
Job Accessibility (000)	50.9	61.0	16.0	16.2	8.6	7.2
Total Employment (000)	1.1	2.9	0.8	1.4	0.6	0.7
Total Population (000)	1.0	0.5	0.8	0.5	0.6	0.4
% of SFR land use	35.9%	22%	34.9%	23%	43%	21%
% of MFR land use	5.6%	7%	4.3%	6%	7%	6%
% of COM land use	15.1%	15%	15.3%	16%	20%	18%
Total parks	1	2	1	1	1	1
Pedestrian Destinations	10	19	9	14	12	14
Land use mix index (Entropy index, 0-1)	0.4	0.1	0.4	0.1	0.5	0.1
Stop located: (1) in downtown Portland; (2) near Univ. of Oregon; (3) near So. Oregon Univ.	1.9% of stops		5.1% of stops		1.7% of stops	
Distance to city center (miles)	8.6	4.5	4.6	6.4	4.6	4.1

Transit Ridership (Dependent variable)

The data we used for TriMet were from a three-month weekday average from Fall 2008, collected using automated passenger counters on each bus or light rail car, are linked to stops via an automatic vehicle location (AVL) system. We aggregated the total "ons" and total "offs" for each stop location to create the dependent ridership variable. Ridership data for LTD is from one week in October 2008, also collected using automatic counters. We aggregated the "ons" and "offs" of the five weekdays by transit stop ID. RVTD's 2008 ridership data are based upon a hand-count. RVTD has since begun collecting data through an automatic counting system, but we wanted to use data across the three metropolitan areas from the same year. RVTD collected the ridership data by sampling transit trips for each transit route at different days from December 2007 to December 2008, and then aggregated the "ons" and "offs" during the sampling days by stop. The daily ridership was calculated by dividing the aggregated ridership by number of sampling days. As mentioned above, due to the skewed distribution of ridership data, we used logarithm form of ridership data as the dependent variable for models of both areas.

Independent Variables

The socio-demographic makeup of each stop buffer area was obtained using available United States Census data from the 2005-2009 American Community Survey (ACS). ACS data from block groups around each stop buffer were compiled to determine the age, employment, gender, income, population, poverty, and race of the residents surrounding each transit stop. A proportional split methodology was used that assigns block group attributes at the same proportion of that block group that falls within the transit stop buffer area. For example, if 42% of the area of a block group falls within the stop's buffer, 42% of the block group's population would be assigned to the stop area.

Transit service characteristics were measured in a variety of ways. Maximum coverage time (in minutes) is the difference in time between the first and last route of the day. However, for some routes there were large gaps of time without service. For example, some routes only operate during the peak commute times. If the gap was more than four hours, those gap times were eliminated from the coverage time. The coverage time was then used to calculate average headways – the number of minutes between each vehicle – for the route. If more than one route served a stop, the headway for most frequent route was assigned to the stop.

Each transit stop was also coded as to its transfer availability or the number of transfer opportunities between routes available at each stop. The presence of high capacity transit such as light rail or bus rapid transit (BRT) within each stop area was also noted. Park and ride lots were characterized as one of two types: (1) for bus only; or (2) for both bus and MAX light rail in Portland or for both bus and Bus Rapid Transit (BRT) in Lane County. There were no such lots in the RVTD area.

Transportation infrastructure was characterized by the street pattern and bicycle facilities. Our measure of street connectivity is the number of three- or more-way intersections within the buffer, or a measure of intersection density, since the buffers are consistent. Bicycling may be a complementary or competitive mode for transit. Bicycle infrastructure was measured as the miles of bike lanes and multi-use paths within the buffer. Multi-use paths are separated from the street and include access for pedestrians. Path data were only available for the TriMet area.

The land use variables tested in our models tried to reflect a variety of uses that could positively or negatively affect ridership (TABLE 4). The variables for total employment and total population within the buffer act as density measures, since the circular buffer sizes are constant. Employment data from Oregon Employment Department quarterly reports were geocoded to taxlots within the study areas. The data includes such information as salary, North American Industrial Classification System (NAICS Codes), and total number of employees. An improved 2008 dataset was available for both Lane and Metro, but not for Jackson County. The most important improvement in the data was the increased employment accuracy that resulted from employment data that was spatially disaggregated from a corporate headquarters to its regional outlets. In addition to the total number of jobs within the stop's buffer area, we measured job accessibility for each stop using the multi-modes network analysis tool in ArcGIS. The variable is defined as the total jobs that can be accessed by transit (plus walking) within 15 minutes. This measure is assumed to have a positive association with transit ridership.

We tested three other variables that might capture pedestrian destinations other than employment: commercial land use; land use mix; and pedestrian destinations. An entropy land use mix measure was created utilizing a variety of land use types, including institutional, industrial, recreational, commercial, multi-family, and single family housing land uses. The number of "Pedestrian Destinations" within the buffer area was derived using the address or tax lot-based employment data. This was intended to provide a measure of possible pedestrian-oriented destinations in close proximity to each transit stop and included the following NAICS codes, along with parks and libraries (identified through GIS files):

Convenience stores (445120, 447110)
Supermarkets and other grocery stores (445110)
Hardware stores (444130)
Fruit and Vegetable Markets (445230)
Dry cleaning and laundry (812320)
Clothing stores (448110, 448120, 448130, 448140, 448150, 453310)
Postal service (491110)
Schools and colleges (611110, 611210, 611310, 611410)
Bookstores (451211)
Used merchandise stores (453310)
Restaurants & bars (722211, 722213, 722110, 722211)
Banks (522110)
Video/Disc rental (532230)
Pharmacies and drug stores (446110)
Beauty salons (812112)
Fitness/sports centers, recreation centers (713940, 624110)
Child day care services (624410)
Religious organizations, including churches (813110)
Services for elderly and persons with disabilities (624120)
Medical and dental offices (621111, 621112, 621210, 621310, 621320, 621330, 621391)

In addition to total population, residential land use was measured as the share of buffer area used for single-family or multi-family residential land uses. To account for major destinations that might have more of a regional draw and characteristics not accounted for with the other land use variables, we created variables for downtown Portland and the University of Oregon and Southern Oregon University campus areas. Stops were coded as either being within (1) or outside (0) these areas. In addition, for each region, the distance to downtown (Portland, Eugene, or Medford) was measured and used to reflect the relative position of each stop with the downtown employment center. A list of the variables and their basic descriptive statistics are summarized in TABLE 4.

FINDINGS

The final model results for TriMet, LTD and RVTD are summarized in TABLE 5 and shown in detail in TABLE 6. The TriMet model does the best job explaining the variation in ridership at the stop-level; the adjusted-R^2 is 0.69, indicating that the independent variables explain 69% of the variance in the dependent variable. The adjusted-R^2 for the LTD and RVTD model are 0.62 and 0.53 respectively. Because the dependent variable is a logarithmic form of ridership data, the estimated coefficients should be interpreted as the percentage change in ridership associated with one unit change in the independent variable. After developing the final models, we entered the variables into each model in groups (socio-demographic, transit service, transportation infrastructure, and urban land use) to estimate the relative contribution of each of those sets of characteristics (TABLE 5). As expected, transit level of service characteristics are the most important factors in determining ridership at the stop level. For the Portland region and Lane County, socio-demographic factors are second in importance, followed by land use variables. For Rogue Valley, land use variables explain more than the socio-demographic variables. The discussion below discusses the statistically significant variables, including differences among the three models.

TABLE 5 Contribution of Variables to Overall Model Explanatory Power

	Portland (TriMet)	Lane County (LTD)	Rogue Valley (RVTD)
Adjusted R^2	0.69	0.62	0.53
Socio-Demographic Variables	24%	11%	14%
Transit Service Variables	41%	46%	24%
Transportation Infrastructure Variables	1%	1%	1%
Land Use Variables	4%	5%	17%
Unexplained by the model	31%	38%	47%

Note: The contribution of the variables as a group to the overall model is estimated using the change in the adjusted R^2 after each group of variables is entered into the model, starting with socio-demographic variables. The percentages do not add up to the final adjusted R^2 due to rounding.

TABLE 6 Model Results

	Portland (TriMet)		Lane County (LTD)		Rogue Valley (RVTD)	
	Coeff.	**p**	**Coeff.**	**p**	**Coeff.**	**p**
Socio-Demographic Variables						
% of white population	-1.163	.00	-.314	.56	-.491	.68
% of population with aged under 17	.662	.03	1.076	.08	-1.072	.27
% of population aged 65 or older	.058	.85	-2.603	.00	3.953	.00
% of population with college or above degree	-.799	.00	-.299	.40	-1.437	.21
% of households without vehicle available	-.788	.00	-1.474	.01	-5.102	.00
% of households with annual HH income below the poverty level	.920	.00	-.113	.80	2.838	.02
Transit Service Variables						
Rail transit or BRT station (0=bus stop)	2.814	.00	1.962	.00		
Transfer Stop	.577	.00	.177	.01	.170	.64
Transit Center	2.297	.00	2.807	.00	4.003	.00
Average headway (minutes)	-.041	.00	-.025	.00	-.042	.00
Maximum Coverage Time (minutes)	.003	.00	.002	.00	.077	.41
Total bus stops (within buffer)	-.012	.00	-.016	.07	-.077	.02
Total light rail stations (within buffer)	-.239	.00				
Park & Ride for Bus and LRT (or BRT)	.944	.00	.553	.00	1.287	.00
Park & Ride for bus only	.328	.01				
Transportation Infrastructure Variables						
Street Connectivity	.020	.00	.007	.00	.011	.07
Miles of regional multi-use paths	.300	.00				
Miles of bike lanes	.182	.00	-.102	.25	.374	.03
Land Use Variables						
Job Accessibility (natural log*, 000)	.057	.00	.010	.01	.174	.05
Total Employment (000)	.091	.00	-.071	.12	.113	.33
Total Population (000)	.303	.00	-.139	.23	.754	.00
% of SFR land use	.099	.36	.289	.19	-.617	.38
% of MFR land use	2.339	.00	4.089	.00	2.739	.01
% of COM land use	1.882	.00	.432	.18	2.152	.00
Total parks (area)	-.031	.00	-.048	.07	.001	.99
Pedestrian destinations	.013	.00	.024	.00	.013	.03
Land use mix index	.160	.12	.741	.04	-.077	.89
Stop located: (1) in downtown Portland; (2) near Univ. of Oregon; (3) near So. Oregon Univ.	.921	.00	-.185	.35	-.098	.85
Distance to city center (miles)	-.017	.01	.032	.00	.065	.02
Model Statistics						
Adjusted R^2	.69		.62		.53	
N	7214		1400		350	

*Natural-log form from was used for TriMet and RVTD models.

Socio-Demographic Variables

The socio-demographic variables explain about 24% of the variance in the TriMet model and 11% in LTD and 14% in RVTD models. The sign, magnitude and significance level of the coefficients for socio-demographic variables among the three models do share several similar characteristics but differences exist as well. Within the Portland area, three demographic variables had a significant negative effect on ridership: the share of the population that was white, was college-educated, and did not have a vehicle. The first two are consistent with other research. These variables were not significant in the LTD or RVTD models, though the signs of the coefficients were consistent. This may be due to the relative lack of variation of these two variables in those two areas. The third relationship is unexpected. The model predicts that as the share of households without vehicles increases, ridership at that stop will decrease. A similar relationship was found in the LTD and RVTD models. As expected, the TriMet and RVTD models do predict that as the share of households below poverty increases, ridership will increase. The unexpected coefficient for vehicle ownership indicates that when the model controls for income (poverty) and other demographics, zero vehicle households have a negative effect on ridership. This may indicate that zero-vehicle households that are not in poverty are not riding transit at a particularly high rate. It may also be due to geography and where zero-vehicle households are concentrated. In the Portland region, most of the stops with high concentrations of zero vehicle households are in or near downtown. In the LTD area, the stops with concentrations of zero-vehicle households were in or near downtown and the University of Oregon campus. It may be that these residents are walking or bicycling to many destinations, rather than using transit.

The final two demographic variables included in the models were the shares of the population under 17 and 65 and older. For both the TriMet and LTD models the share of population under 17 had a positive relationship with ridership. At the time in Portland and Eugene, students were eligible for free transit passes and the public transit buses were often used in place of school bus service, particularly at the high school level. One interesting variable is the share of population 65 years or older, which had a non-significant relationship with ridership in TriMet model, a negative relationship in LTD model, and a positive relationship in RVTD model. Rogue Valley has a higher portion of its stops with a relatively high share of the population over 65. About one-quarter of the RVTD stops have a surrounding population that is at least 20% older adults. This fits Rogue Valley's reputation as an attractive retirement community. In contrast, only about five percent of TriMet's stops have that high of a share. With more stops having a concentration of older adults in Rogue Valley and Lane County (about 13% of stops), there is a greater possibility that ridership at those stops can influence the model coefficients, either positively or negatively. The direction of the relationship might be due to unique characteristics of older adult communities in the two areas. For example, it may be that there are some older adult communities in Rogue Valley that are particularly well-served by transit and do not provide their own competing transportation services.

Transit Service Variables

The transit service variables explain about 41% of the variance in the TriMet model, 46% in the LTD model, and 24% in the RVTD model. All of the variables were significant in the TriMet and LTD models, with coefficients in the expected direction. In the RVTD model, two variables, transfer stop and transit coverage time, were not significant even though their coefficients have the expected sign. This is not surprising when considering the relatively small sample size of the RVTD model (350) compared with LTD (1400) and TriMet (7214). There was less variation within the variables in the RVTD service area. For example, there are only 11 transfer stops, and all the stops have a coverage time ranging from 10.5 to 13.5 hours.

In general, transit ridership was higher at transfer stops, transit centers and stops with park and ride lots, however it was lower as the number of nearby stops increased. This makes sense, as a greater number of stops nearby (for the same route or other routes) can disperse riders. Longer headways decreased ridership, and longer coverage time increased ridership. The magnitude of the variables was similar among the three models, with a few exceptions. Transfer stops had a greater effect on ridership in the Portland region; all else being equal, ridership at a transfer stop was 58% higher than at other stops. This likely reflects the larger transit network. Longer headways appear to have a slightly larger effect on RVTD and TriMet ridership than LTD ridership. Each extra minute of headway is associated with a four to five percent drop in ridership for RVTD and TriMet, compared to a two percent drop for LTD. The larger effect for RVTD might be explained by the limited range of values: 30, 45, and 60 minutes (based upon schedules). Riders may be even more sensitive to waiting times in this range. For TriMet, where the headways ranged from 11 to 76 minutes (based upon on-board data), riders overall might be more time sensitive, indicating that they are more likely to be "choice" riders. Proximity to a park and ride lot had a significant and positive association with ridership, and this is consistent among the three models. Finally, ridership at rail and BRT stations is about three times and two times higher, respectively, than ridership at bus stops.

Transportation Infrastructure Variables

The three transportation infrastructure variables explain about one-percent of the variation in each model. Street connectivity is positively associated with ridership in three models (though it is not significant in the RVTD model), indicating that the shorter walking distances afforded by increased connectivity likely improve accessibility. While a small overall percentage, this result confirms earlier work by Ryan and Frank ([13]). The presence of multi-use pedestrian and bicycle paths was associated with increased transit ridership in Portland, while the presence of nearby bike lanes was associated with increased transit ridership in both Portland and Rogue Valley. This may be capturing both direct and indirect relationships. All TriMet buses are equipped with bike racks, allowing for easy transfer between the modes. Therefore, the two types of infrastructure (bike facilities and transit) may be synergistic. On the other hand, bike lanes or paths may be located along corridors that exhibit some other characteristic that is associated with transit ridership – a variable that we have not otherwise accounted for in our models.

Land Use Variables

The land use variables explain about 4-5% of the variance in the TriMet and LTD models and 17% in the RVTD model. The reasons for this large difference are not immediately apparent and are worth further exploration. The significant effects of the individual variables are generally consistent with theory, though the models are not consistent with respect to which variables are significant. As expected, the better the job accessibility of the stop, the higher the ridership; this is found in all three models. As the total employment around a stop increases, so does ridership – but only in the Portland region. In both Portland and Rogue Valley, as the total population near a stop increases, so does ridership. This variable is not significant for Lane County; moreover the coefficient is negative.

The portion of land used for multi-family residential (MFR) is significantly and positively associated with higher ridership in all three locations. Commercial land use is also positively associated with ridership in all three areas, but only significant in Portland and Rogue

Valley. The effect of MFR is somewhat higher in Lane County, while the effect of commercial land uses is larger in Portland and Rogue Valley. The proportion of acreage in single-family housing is not significantly related to ridership in any of the models. It is included because it does help control for other relationships.

The proximity to possible pedestrian-oriented destinations is consistently significant in all three models; for each additional destination within the quarter-mile buffer, ridership goes up by 1-2%. The significance of this variable may explain why the land use mix entropy index is not significant in the TriMet and RVTD models; the pedestrian destination measure may have more power for predicting transit ridership. However, land use mix remains significant in the LTD model even after controlling for proximity of pedestrian destinations.

Stops located in downtown Portland have higher ridership, even after accounting for density, other land use factors, and transit service characteristics. This indicates that there is something else, not explicitly captured in our model, about downtown Portland that attracts transit riders. On the other hand, there was no significant relationship between ridership and a stop being located near the University of Oregon and Southern Oregon University campus, which might be expected to be major transit destinations. Distance to downtown is negatively associated with ridership in Portland, indicating that ridership goes down at stops farther away from the city center. However, the opposite relationship was found in Lane County and Rogue Valley – ridership increases further from downtown. Finally, the presence of parks is associated with lower transit ridership. This makes sense, in that parks are not a common transit origin or destination.

Combined Effect of Service and Land Use

Does the combination of having a high level of service and high proximity density or pedestrian friendly design contribute to a proportionally greater effect on ridership than the sum of these two individual effects? In order to test this hypothesis, we added interactive terms to the three models. For simplicity, only the results of statistically significant interactive terms are shown in TABLE 7. The negative signs of the significant interactive terms indicate that density or pedestrian design immediately around the transit stop or station could have a larger effect on ridership when the headway is low, or the impact of headway on ridership (negative effect) is greater at a stop or station with higher density or better pedestrian design.

TABLE 7 Model Results of Interactive Terms

	Portland (TriMet)		Lane County (LTD)		Rogue Valley TD	
	Coeff.	**p**	**Coeff.**	**p**	**Coeff.**	**p**
Population * Headway	-.0216	.0000	-.0186	.0002	-.0443	.0032
Employment * Headway	-.0018	.0453				
Pedestrian Destination * Headway	-.0002	.0111			-.0007	.0367
Street Connectivity * Headway					-.0014	.0104

CONCLUSION AND POLICY IMPLICATIONS

This study developed stop-level transit ridership models with relatively comprehensive transit service and built environment variables using data from three different urban regions in Oregon. The adjust R^2 values and significance of most the chosen variables suggest that the models do a good job of explaining transit ridership. Results of three models indicate that transit service plays the most important role in predicting transit ridership, but that the built environment characteristics around the stop or station also matter. The built environment not only has a direct influence on ridership, but also may interact with transit service to deliver additive effects on ridership.

The results provide important implications for transit policy and how to promote more livable communities through public transit. Five primary policy implications could be drawn from this analysis:

1. Improving level of service of transit is an important tool to leverage transit ridership. This not only includes shortening headways and extending service coverage time, but also improving multi-modal connections and providing transfer opportunities.
2. Promoting a pedestrian-friendly built environment around transit stops or stations can contribute to ridership. This includes enhancing street or pedestrian-path connectivity and encouraging more pedestrian-oriented business development around transit stops.
3. Better integrating land use development with transit investments, in particular focusing on multi-family housing and pedestrian-oriented commercial land use is important for transit ridership. Focusing such efforts around stops and stations with higher levels of service will be most effective.
4. Focusing further research, as well as transit planning, at the transit stop level is important as it is the spatial scale by which transit users experience transit. While regional connectivity of the transit system is obviously important (does transit go where it needs to), so too is the local built environment around individual transit stops as most transit users are pedestrians at their origin or destination or both. Policy, planning, development, and research would do well to focus at this spatial scale.
5. There may be further aspects of the urban design or "quality" of the local built environment that are important for ridership but are not captured in this study. For example, is the transit stop adjacent to a street crossing, are there pedestrian paths from the transit stop to the commercial areas or does one need to walk through large parking lots, and do the scale of buildings, quality of sidewalks, presence of street trees, etc. support the feeling of comfort and safety for pedestrians?

ACKNOWLEDGEMENTS

We would like to thank the Oregon Department of Transportation (ODOT), Federal Transit Administration (FTA), and Oregon Transportation Research and Education Consortium (OTREC) for their support of this research and interest in thinking about the context of urban form as an important aspect of transit service. The findings and conclusions, including any errors, are those of the authors alone.

REFERENCES

1. Chu, X., *Ridership models at the stop level*, in Report No. BC137-31 prepared by National Center for Transit Research for Florida Department of Transportation. 2004.
2. Peng, Z.-R., et al., A simultaneous route-level transit patronage model: demand, supply, and inter-route relationship. *Transportation*, 1997. **24**(2): p. 159.

3. Ewing, R. and R. Cervero, Travel and the built environment. *Journal of the American Planning Association*, 2010. **76**(3): p. 265-294.
4. Banerjee, T., D. Myers, and C. Irazabal, Increasing Bus Transit Ridership: Dynamics of Density, Land Use, and Population Growth. METRANS report No. 03-24. Los Angeles, CA, 2005.
5. Cervero, R., Alternative Approaches to Modeling the Travel-Demand Impacts of Smart Growth. *Journal of the American Planning Association*, 2006. **72**(3): p. 285.
6. Lin, J.J. and T.Y. Shin, Does Transit-Oriented Development Affect Metro Ridership? Evidence from Taipei, Taiwan. *Transportation Research Record*, 2008(2063): p. 149-158.
7. Cervero, R., Mixed land-uses and commuting: evidence from the American Housing Survey. *Transportation Research: Part A, Policy and practice.* 1996. **30**(5): p. 361.
8. Cervero, R., Built environments and mode choice: toward a normative framework. *Transportation Research Part D Transport and Environment,* 2002. **7**(4): p. 265-284.
9. Estupinan, N. and D.A. Rodriguez, The relationship between urban form and station boardings for Bogota's BRT. *Transportation Research. Part A, Policy and Practice,* 2008. **42**(2): p. 296-306.
10. Brown, S., et al., *Understanding How the Built Environment Around TTA Stops Affects Ridership: A Study for Triangle Transit Authority.* 2006, Department of City and Regional Planning, University of North Carolina, Chapel Hill.
11. Mishra, S., T.F. Welch, and M.K. Jha, Performance indicators for public transit connectivity in multi-modal transportation networks. *Transportation Research Part A,* 2012. **46**(7): p. 1066-1085.
12. Cervero, R., J. Murakami, and M. Miller, Direct Ridership Model of Bus Rapid Transit in Los Angeles County, California. *Transportation Research Record: Journal of the Transportation Research Board*, 2010. **2145**(-1): p. 1-7.
13. Ryan, S. and L.F. Frank, Pedestrian Environments and Transit Ridership. *Journal of Public Transportation,* 2009. **12**(1): p. 39-58.
14. Pulugurtha, S.S. and M. Agurla, Assessment of Models to Estimate Bus-Stop Level Transit Ridership using Spatial Modeling Methods. *Journal of Public Transportation*, 2012. **15**(1): p. 33-52.
15. Zhao, F., et al., *A Transit Ridership Model Based on Geographically Weighted Regression and Service Quality Variables*, in Final Report, Prepared for the Florida Department of Transportation, Tallahassee, FL. 2005.