How to Increase Rail Ridership in Maryland: Direct Ridership Models for Policy Guidance

Chao Liu, Ph.D.[1]; Sevgi Erdogan, Ph.D.[2]; Ting Ma[3]; and Frederick W. Ducca[4]

Abstract: The State of Maryland aims to double its transit ridership by the end of 2020. The Maryland Statewide Transportation Model (MSTM) has been used to analyze different policy options at a system-wide level. Direct ridership models (DRMs) estimate ridership as a function of station environment and transit service features rather than using mode-choice results from large-scale traditional models. They have been particularly favored for estimating the benefits of smart growth policies such as transit-oriented development (TOD) on transit ridership and can be used as complements to the traditional four-step models for analyzing smart-growth scenarios at a local level. They can also provide valuable information that a system-level analysis cannot provide. For example, DRMs can provide ridership estimate when densifying either households or employment at certain stations to test the effectiveness of TOD strategies at target stations. In this study, the authors developed DRMs of rail transit stations, namely light rail, commuter rail, Baltimore metro, and Washington DC Metro for the State of Maryland. Data for 112 rail stations were gathered from a variety of sources and categorized by transit service characteristics, station built-environment features, and social-demographic variables. The results suggest that impacts of the built environment differ for light rail and commuter rail. For light rail stations, employment at 0.8 km (half-mile) station areas, service level, feeder bus connectivity, station location in the Central Business District (CBD), distance to the nearest station, and terminal stations are significant factors influencing ridership. For commuter rail stations, only feeder bus connections are found to be significant. The DRM results have implications for agencies aiming to increase transit ridership through leveraging the land use near transit-station areas. The results show that stations with higher employment were observed to have higher transit boardings. Thus, agencies desiring to increase transit ridership should consider zoning regulations and site-design requirements that allow denser development around transit stations. However, increasing densities must be in conjunction with improved transit service levels, parking, and feeder bus services to take full advantage of rail transit. Some of the limitations of the current study and future research directions Include, first, that the DRMs were developed for the State of Maryland only. Future research should be carried out for surrounding regions to get generalized conclusions. Second, other important factors should be incorporated in the model, such as safety and transit reliability attributes. Third, more-detailed parking information, such as parking costs and supply, should be included in the model. **DOI:** 10.1061/(ASCE)UP.1943-5444.0000340.

Author keywords: Direct ridership models (DRMs); Rail transit ridership; Land use; Transit-oriented development (TOD); Zoning policy.

Introduction

Transportation is a critical issue for Maryland, both as a foundation for the state's economy and for meeting the travel needs of Maryland residents. Maryland's transportation system features extensive intraurban travel within two major metropolitan areas, Washington DC, and Baltimore, as well as interurban travel that can traverse the Appalachian Mountains and the Chesapeake Bay. The state expects significant impacts on its transportation system through changes in population, economy, and environment over time (MDOT 2009). By 2030, the population of the state is expected to increase to 6.7 million (about one million increase), exacerbating the pressure on a transportation system that is already experiencing congestion. The state prepares itself to tackle these challenges with a range of policy options such as smart growth, travel-demand management, strategies that target reducing demand for transportation, and operational strategies that target use of existing system efficiently (MDOT 2009).

Maryland is one of the pioneers of progressive land-use policies. These policies intend to prevent further sprawl in the state as well as its negative impacts on the transportation system. Now known as smart growth policies, these land-use policies were initiated in 1992 with the Economic Growth, Resource Protection and Planning Act, a policy that outlined seven (later eight) visions for the future growth in Maryland (MDP 2013).

Later in 1997, the Smart Growth and Neighborhood Conservation initiative was launched to use state funds as incentives to direct growth. Since then, the state has adopted a variety of Smart Growth laws and policies namely the Priority Funding Areas Act of 1997, legislation on Planning in 2006, and Sustainable Communities in 2010, and finally the Sustainable Growth & Agricultural Preservation Act of 2012 (MDP 2013).

The most-recent initiative added to Maryland's smart growth agenda is the Smart, Green and Growing Initiative, a multiagency, statewide initiative that aims to achieve a more-sustainable future

[1]Faculty Research Associate, National Center for Smart Growth Research and Education, Univ. of Maryland, 1219A Preinkert Field House, College Park, MD 20742 (corresponding author). E-mail: cliu8@umd.edu

[2]Faculty Research Associate, National Center for Smart Growth Research and Education, Univ. of Maryland, 1112J Preinkert Field House, College Park, MD 20742.

[3]Ph.D. Student, National Center for Smart Growth Research and Education, Univ. of Maryland, 1219B Preinkert Field House, College Park, MD 20742.

[4]Ph.D. Senior Research Scientist, National Center for Smart Growth, Univ. of Maryland, 1112N Preinkert Field House, College Park, MD 20742.

Note. This manuscript was submitted on November 5, 2014; approved on February 9, 2016; published online on May 20, 2016. Discussion period open until October 20, 2016; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Urban Planning and Development***, © ASCE, ISSN 0733-9488.

for Maryland through community revitalization, transportation improvements, economic development, smart growth, and environmental restoration efforts (MDOT 2009). As part of this initiative and as a transportation policy, the state established transit-oriented development (TOD) to strengthen coordination between land use and transportation planning. By encouraging development around existing and planned transit stations, the state aims to maximize the value of its investment in transit, reduce congestion, reduce greenhouse gas (GHG) emissions and pollution, and provide an alternative to sprawl (MDOT 2009). The state aims to double its transit ridership by the end of 2020, while reducing its greenhouse gas by 25%. Among other strategies to achieve these goals, TOD has a significant role. This is partially due to the high potential of Maryland for TOD application with over 75 rail, light rail, and metro stations, and many more planned ones in the next 20 years (CMRT 2009). However, challenges exist such as higher upfront infrastructure costs, necessity to use public land, and complex community-related issues (NCSG 2012a). Therefore, identifying the factors that influence the success of the TOD program is a critical issue for Maryland and other states.

The state utilizes the Maryland Statewide Transportation Model (MSTM), a three-tier integrated land-use and transportation model, to analyze different policy options including TOD at a system-wide level (NCSG 2012a, b). However, the state agencies also need more-detailed, local-level policy guidance to meet this goal. The current MSTM operates on a regional scale, predicting travel behavior at an aggregate level, and is useful in guiding highway and transit network capital investment priorities but cannot estimate the travel impacts of neighborhood-level development (Cervero 2006; Duduta 2013). This presents an issue for regional travel-demand models given the relatively low transit usage in the United States. Even minor model imprecisions can cause significant changes in the location-specific ridership estimates, yielding unreliable transit share forecasts (Fehr & Peers 2013). The four-step models also cannot adequately reflect the built environment's impact on boosting transit ridership, a point that was consistently found in a large quantity of empirical studies (Cervero 2006; Duduta 2013). Using the four-step models for transit ridership forecasting also faces institutional and financial barriers as developing and maintaining them require staff resources and interagency or consultant involvement (Marshall and Grady 2006). Given the nature of the travel-demand modeling that is not suitable for the local level transit policy testing, a method that can estimate transit ridership is needed for policy guidance.

Direct ridership models (DRMs) have emerged in the United States as a low cost, quick alternative to traditional four-step travel-demand models to forecast transit ridership at the station or corridor level (Cervero 2007; R. Cervero, et al., "Direct ridership model of bus rapid transit in Los Angeles county," working paper; Gutiérrez et al. 2011). DRMs estimate transit ridership based on multiple regression analyses of the built-environment characteristics of station areas, transit features such as transit service and station facilities, and sociodemographic characteristics of riders. DRMs complement conventional four-step travel-demand models.

DRMs respond to changes in the built environment and transit services in the immediate station area (Cervero 2006; Fehr & Peers 2013). Utilizing multivariate regression analyses based on empirical data, DRMs quantify how land-use factors at the local level influence transit ridership in a direct and immediate way. Thus, it is a low-cost prediction approach compared to complex travel-demand models because it only requires data inputs associated with immediate station areas. These areas are called catchment areas, large enough geographic buffers to capture neighborhood attributes (Cervero and Murakami 2009; Adams et al. 2009; Pulugurtha et al. 2006; Antipova et al. 2011; Mitra and Buliung 2012). The catchment size for the built-environment measurement is flexible in the literature. For example, in analyzing the Bay Area Rapid Transit (BART) system in the San Francisco Bay Area, it was defined as the contiguous area that historically captured 90% of all access trips to and egress trips from a station. An application in Charlotte, North Carolina defined the catchment size as the distance to the nearest adjacent station (Cervero 2006; Duduta 2013). Comparing different catchment sizes using a sample of about 1,500 stations, Guerra et al. ("The half-mile circle: Does it best represent transit station catchments?" working paper, Center for Future Urban Transport, UC Berkeley) found that station catchment size is not a significant factor in predicting ridership. They suggest that the use of available data related to catchment area would be sufficient when estimating DRMs.

DRMs are built upon understanding of factors at the end points of possible trips that contribute to transit ridership. Literature suggests three categories of independent variables in their DRMs model: built-environment characteristics (of neighborhoods), socioeconomic factors, and transit/station attributes (Cervero and Murakami 2009; R. Cervero, et al., "Direct ridership model of bus rapid transit in Los Angeles county," working paper; Gutiérrez et al. 2011; Kuby et al. 2004; Li et al. 2015).

In terms of the built-environment characteristics, population density and employment density have been included in almost all DRMs empirical studies as key built-environment factors affecting ridership (Cervero 2006; Fehr & Peers 2013). As research on the built environment's effect on travel behavior evolved, new findings suggested additional variables reflecting smart growth to be added in, such as distance to central business district (CBD) and mixed-land-use index (Cervero 2006; Cervero and Murakami 2009). Kuby et al. (2004) found that the numbers of employment, population, and renters within walking distance are significantly associated with the average weekday boardings of light rail stations. However, the CBD dummy variable was not significant in their study (Kuby et al. 2004). Land-use mix was found to be positively associated with transit ridership in Gutierrez et al. (2011) and Filion (2001). Chu (2004) found the number of jobs and pedestrian environment positively associated with weekday total boardings (Chu 2004). Li et al. (2015) found that distance weighted population is a significant variable in affecting ridership (Li et al. 2015).

For socioeconomic features, Kuby et al. (2004) found the percentage of households who rent instead of own is significantly associated with light rail ridership, specifically that a 1% increase in the number of renters will generate 6.24 daily boardings (Kuby et al. 2004). Gutierrez et al. (2011) found that the size of the foreign population, the number of workers, and specifically, the number of jobs in commercial and educational sectors, are associated with transit monthly boardings in the Madrid Metro system (Gutiérrez et al. 2011). Chu (2004) found that the share of households without a car, percentage of the people in workforce, share of women, and share of the Hispanic population are positively associated with weekday total boardings in Jacksonville, Florida, while median household income, percentage of youth (under 18), and share of White population are negatively associated (Chu 2004) with boardings.

Regarding transit service/station attributes, transit service variables include frequency of service and operating speeds, as well as the number and frequency of feeder bus lines (Cervero 2006; R. Cervero, et al., "Direct ridership model of bus rapid transit in Los Angeles county," working paper; Duduta 2013). Cervero (2006) found in a Charlotte, North Carolina case study that the inclusion of service level greatly improved model accuracy. Using a DRM to predict Bus Rapid Transit ridership in Los Angeles County, Cervero and Murakami (2009) found that service frequency strongly influences Bus Rapid Transit (BRT) ridership estimates, as well as high intermodal connectivity (both feeder bus routes and rail-transit services). Kuby et al. (2004) found that bus connections and park and ride spaces are significantly associated with the average weekday light-rail ridership, using data from 268 light-rail stations in the United States (Kuby et al. 2004). Station attributes including bus shelters, bus benches, park-and-ride lot capacity, and availability of information systems have been found significant in influencing transit ridership. For example, using DRMs to predict BRT ridership in Los Angeles County, Cervero and Murakami (2009) found that park and ride lot capacity is a significant factor (positive) in increasing ridership. Kuby et al. (2004) also found that dummy variables for terminal and transfer stations are all significant (positive) (Kuby et al. 2004). Gutierrez et al. (2011) found that nodal accessibility at stations, number of transit lines, and bus feeders are significant characteristics (positive) (Gutiérrez et al. 2011). Chu (2004) found that the Transit Level of Service (TLOS), an indicator developed by the Florida Department of Transportation (FDOT) to capture transit availability and mobility features [includes three separate measures: (1) service frequency, (2) hours of service, and (3) spatial service coverage] within walking distance positively affects the total weekday boardings (Chu 2004).

DRMs have been demonstrated to predict ridership impacts of smart-growth policies such as transit-oriented development (TOD). Cervero used a direct ridership model to specify the transit ridership bonus that TOD design elements contribute. In a case study analyzing Hong Kong, he compared coefficients in three ridership models: one without TOD elements as a baseline model, one with additional park-and-ride lot dummy variable, and the other one with TOD dummy variable. The results showed that the model with a park-and-ride lot did not increase transit ridership significantly, either for weekends or weekdays, but TOD did increase station boardings for both weekdays and weekends (Cervero and Murakami 2009). This empirical study illustrates that the DRMs helps to quantify TOD's positive effects on transit ridership with more-accurate estimates. DRMs also improve understanding of conventional four-step model results and extend analysis capability by providing additional information. For example, Cervero's Charlotte, North Carolina case study findings showed that the DRMs can add further detail into the conventional four-step model results. He found that a higher combination of population and employment density implies a jobs-–ousing balance, and thus reduces the car ownership rate of traffic analysis zones (TAZs) with TOD (Cervero 2006).

There are also challenges in applying DRMs. Direct ridership models generally have small sample sizes since observations consist of transit stations or stops [from a high of 261 stations in a nationwide Transit Cooperative Research Program (TCRP) model to a low of 27 for a model of St. Louis MetroLink Stations] (Duduta 2013). Thus, degree-of-freedom constraints often limit the number of variables that can be included as well as their specifications (e.g., inclusion of interaction terms) (Cervero 2006; Cervero and Murakami 2009). DRMs are useful policy tools since they can express the ridership's elasticity with respect to independent variables (Duduta 2013); however, they do not consider some variables that four-step models utilize, typically in mode-choice models, such as comparative travel times and prices of transit versus automobiles (Cervero and Murakami 2009). Thus DRMs should be seen as sketch-planning tools that complement conventional travel-demand models, well suited for producing order-of-magnitude estimates of patronage and for probing the sensitivity to key input variables in smart-growth scenarios, but not as a replacement of a fully specified travel-demand models (Cervero 2006; Cervero and Murakami 2009).

In this study, the authors developed DRMs for rail transit using 112 stations in the Baltimore–Washington Metropolitan area. These models are developed considering as many factors as possible based on the literature. The model results are analyzed with specific emphasis on supplementing four-step model analysis results regarding rail transit ridership. The purpose is to provide policy guidance to the State of Maryland in achieving its transit ridership and GHG reduction goals by providing station level analysis. The study is expected to help improve effectiveness of Maryland's TOD policies and shed new insights on further opportunities.

This paper is organized as follows: the next section describes the study area. The analysis framework and modeling approach are in "Methodology" section. The results of the analysis are presented in "Results" section, and policy implications are discussed in "Policy Implications" section. Finally, the conclusions are given in the last section.

Study Area

The study area encompasses the State of Maryland, which consists of 23 counties and the City of Baltimore. It had an estimated population of 5.8 million in the year 2012 (U.S. Census 2013). The state operates a transit system that includes urban bus and metro rail transit in Baltimore and Washington DC. The Maryland Transit Administarvie (MTA) operates service including the Maryland Area Regional Commuter (MARC) intercity train service, light rail train service (LRT), an extensive commuter bus service, and 25 locally operated transit systems (LOTS) (MDOT 2009). The authors included Washington Metropolitan Area Tranist Authority (WMATA) service in the study area since it serves part of the Baltimore–Washington Metropolitan area, specifically Montgomery County and Prince George's County of Maryland. The authors identified 112 rail stations in the area where ridership data are available (Fig. 1).

Methodology

Key Variables and Modeling Approach

In order the develop DRMs, the authors identified available data in three main categories: transit service, station built environment, and sociodemographics (Table 1). Transit service data included daily boardings, availability of park-and-ride facilities and feeder bus services, transit service frequency at the station, station catchment size (defined by the distance to the nearest adjacent station), whether it is a terminal station or not, and the station connectivity (defined by a composite index including variables such as transit routes, coverage, speed, capacity, and urban form (Mishra et al. 2012). The data regarding station built environment are generated for three station areas to investigate the impact of station areas on the ridership. The built-environment data included population and employment density, land-use mix index (defined by entropy land-use mix, which is explained in the next section), street network connectivity (defined by the number of intersections around station area), regional accessibility (defined by the jobs that can be reached in 30 min by auto and transit), distance to central business district (CBD) (since the study area includes two CBDs, the authors included both Baltimore and Washington DC downtowns), and walk score at census tract level. Finally, the authors used vehicle ownership and income level as sociodemographic data.

The authors developed a DRM to estimate the effects of the built-environment variables together with transit service operational and

J. Urban Plann. Dev., 2016, 142(4): 04016017

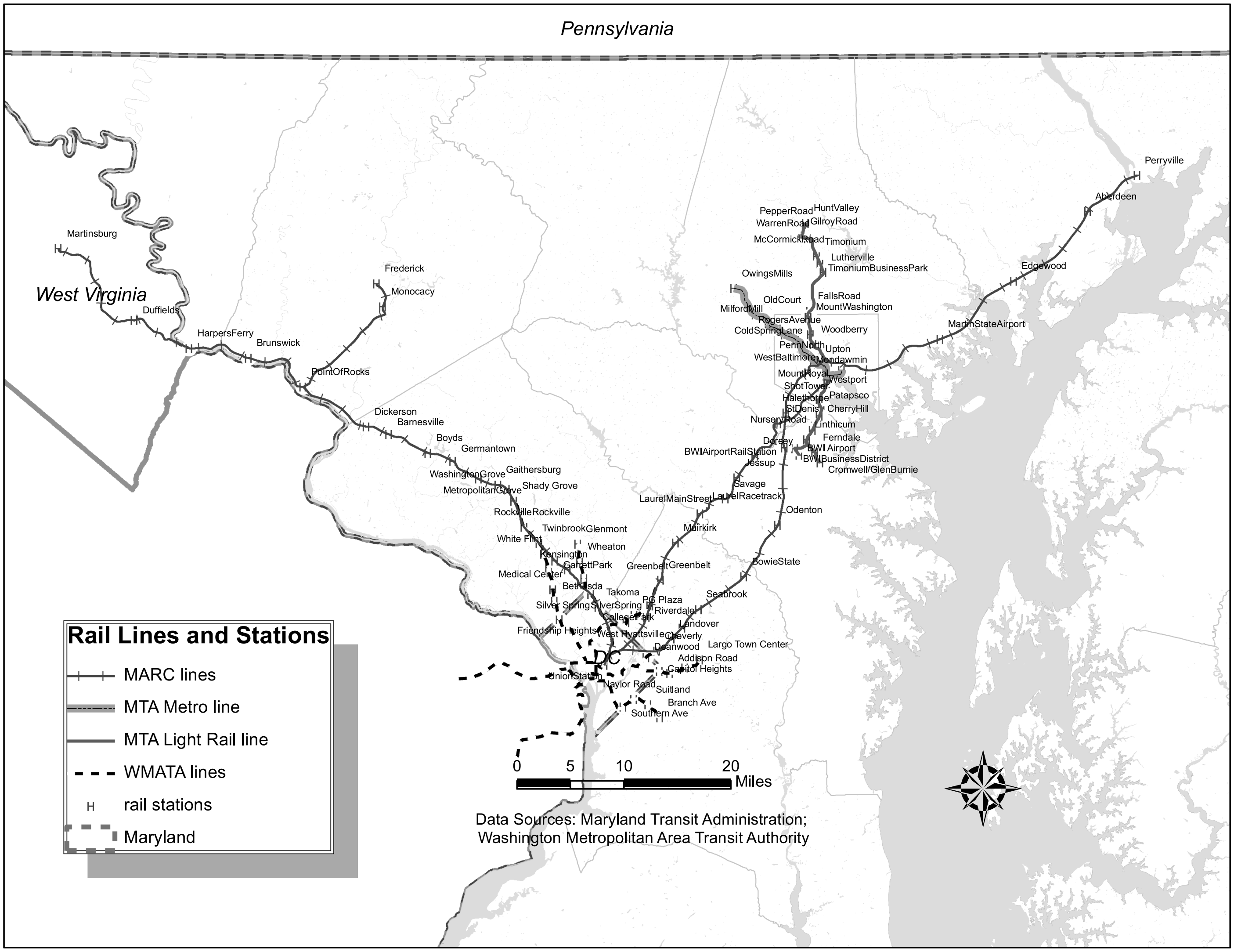

Fig. 1. Study area and rail system in Maryland (data from MTA 2011; WMATA 2011)

design features, and sociodemographic characteristics on transit ridership. Station built-environment data were gathered from multiple data sources (Table 2) for light rail (LRT), metro (Baltimore and Washington DC), and commuter rail (MARC) stations in the Baltimore–Washington metropolitan area and were merged to generate composite data for 112 station proximity areas for both light rail stations and commuter rail stations in the Baltimore–Washington metropolitan areas. The authors utilized six built-environment variables in this research. Each measurement is discussed in detail. In order to test the sensitivity of transit ridership in response to the built environment in different transit catchment areas, a series of station buffer areas were created: [0.4 km (quarter-mile) buffer, 0.8 km (half-mile) buffer, and 1.6 km (one-mile) buffer]. The variables used in DRMs and model specification are presented next.

Direct ridership models estimate boardings at a station for a predefined period of time (e.g., daily or weekly) as a function of three variable sets. In this study, the authors used average daily station boardings in 2011 as the dependent variable.

The model uses a combination of variables constructed from other sources. The transit service variables tend to be directly input while data on population, employment, and station characteristics are converted to constructed variables such as density and accessibility. The details of variables and model specification are presented next.

Transit Service Variables

MTA provided average daily boardings for (1) three MARC lines (the Brunswick Line, the Camden Line, and the Penn Line), (2) Baltimore metro lines and (3) Baltimore light rail. WMATA provided station boardings data for five metro lines (Green, Orange, Red, Yellow, and Blue Lines). Multiple routes can use a single transit station. Thus, boardings were modeled separately for each line in the analysis since different rail systems may behave differently. Level of service (LOS) was captured by using the number of trains in one direction during morning peak period (from 5:50 to 9:00 a.m., which covers the morning period for all the systems considered). Parking and connecting feeder bus service availability were the key variables used to estimate boardings at a given station, as they capture the access and egress to and from other modes. Dummy variables were utilized to indicate whether parking lots and feeder bus service were available at a given rail station. Terminal stations tend to have more boardings than other stations (Cervero 2006). It is also treated as a dummy variable in the model. To measure the nodal connectivity of the station within the network, a node-level transit connectivity index was incorporated in the analysis (Mishra et al. 2012; Hadas and Ranjitkar 2012; Welch and Mishra 2013; Kaplana et al. 2014; Mishra et al. 2015; Sarker et al. 2015). Finally, catchment area, a standard measurement in DRMs

Table 1. Variable Descriptions and Sources

Category	Variable	Description	Data source
Transit service	Boardings	Daily boardings (MARC, LRT, Baltimore Metro, and WMATA Metro)	MTA (2011), WMATA (2011)
	Park and ride	Park-and-ride at station (0,1)	MTA (2011), WMATA (2011)
	Feeder bus services	Bus connection at station (0,1)	MTA (2011), WMATA (2011)
	Service level	Number of trains in one direction in morning peak	MTA (2011), WMATA (2011)
	Catchment size	Distance to nearest adjacent station	MTA (2011), WMATA (2011)
	Terminal station	Station is a terminal station (0,1)	MTA (2011), WMATA (2011)
	Station connectivity index	Composite index including transit routes, coverage, speed, capacity, and urban form	NCSG (2010)
Station built environment [(0.4 km) quarter-mile, (0.8 km) half-mile, and (1.6 km) one-mile buffer]	Density	Population density, employment density	U.S. Census (2010), LEHD (2008)
	Land-use mix index	Entropy land-use mixture	MDP (2010)
	Street network connectivity	Number of intersections around station	U.S. Census (2010)
	Regional accessibility	Accessible jobs within 30 min by auto and transit	Maryland State wide model, LEHD (2008)
	Distance to CBD	Distance to Baltimore and Washington, DC downtown	U.S. Census (2010)
	Walk score	Walk score at census tract level	NCSG (2010)
Socio-demographics [(0.4 km) quarter-mile, (0.8 km) half-mile, and (1.6 km) one-mile buffer]	Vehicle ownership	Mean vehicle per occupied housing unit within catchment area of station	U.S. Census (2010)
	Income	Income level within catchment area of station	U.S. Census (2010)
	Age	Median age within catchment area of station	U.S. Census (2013)
	Ethnic groups	Percent of different racial groups (including White, Black, Hispanic, and other) within catchment area of station	U.S. Census (2013)
	Homeownership	Percent of owner-occupied houses within catchment area of station	U.S. Census (2013)

Table 2. Descriptive Statistics of MARC Stations ($N = 39$)

Variable	Minimum	Maximum	Mean	Standard deviation
Boardings	1.0	3,458.0	530.3	744.0
Park and ride	0.0	1.0	1.00	0.0
Bus	0.0	1.0	0.9	0.4
Level of service	3.0	11.0	7.9	3.2
Distance to nearest station	0.0	8.8	2.8	2.2
Terminal	0.0	1.0	0.1	0.3
Connectivity index	0.0	0.6	0.1	0.1
Population density within 0.8 km (half-mile) buffer	3	16,056	3,502	3,934
Employment density within 0.8 km (half-mile) buffer	3	110,483	10,083	18,065
Land use mixed index within 0.8 km (half-mile) buffer	0.0	1.0	0.5	0.3
Street connectivity within 0.8 km (half-mile) buffer	0.0	264	65	64
Accessibility of auto within 0.8 km (half-mile) buffer	50,749.5	1,752,354.3	769,705.8	486,419.1
Accessibility of transit within 0.8 km (half-mile) buffer	94.4	1,369,992.1	504,985.8	418,944.5
Distance to CBD	9.58 km	108.83 km	41.79 km	24.69 km
Walk score	0.0	180.5	23.1	35.0
Number of vehicles owned by household within 0.8 km (half-mile) buffer	0.7	2.7	1.7	0.4
Household income within 0.8 km (half-mile) buffer	\$25,783	\$1,110,961	\$100,976	\$168,758
Percent of White within 0.8 km (half-mile) buffer	0.0	100%	38.9%	35.6
Percent of Black within 0.8 km (half-mile) buffer	0.0	97.9%	18.8%	24.2
Percent of Hispanic within 0.8 km (half-mile) buffer	0.0	43.3%	8.6%	12.0
Percent of Others within 0.8 km (half-mile) buffer	0.0	42.0%	8.3%	10.9
Median age within 0.8 km (half-mile) buffer	0	70	30	18.6
Percent of owner-occupied housing within 0.8 km (half-mile) buffer	0.0	97.7%	34.2%	32.6

defined by the distance to the nearest station on the same line, is used.

Built-Environment Variables

Six main variables are used to describe the built environment: density, street network connectivity, land-use mix, accessibility, distance to CBD and walk score.

Density. Population density and employment density at the block level were utilized. Block-level population data collected from the U.S. Census (2010) *Summary file 1* (SF1) were merged to station buffer areas. Employment data at block-level were obtained from the Longitudinal Employer and Household Dynamics (LEHD) (LEHD 2008, 2013). Densities were then calculated by dividing population and employment data by the area of buffer

J. Urban Plann. Dev., 2016, 142(4): 04016017

zones. The authors applied log transformation to MARC boardings, population density, and employment density to reduce the skewness in the distribution of the dependent and two independent variables. The authors used interactive terms population × population density, and employment × employment density to avoid the nonlinear relationship of station boardings and density.

Street Network Connectivity. Street network connectivity is measured by the number of intersections (except cul-de-sacs) within (0.4 km) quarter-mile, (0.8 km) half-mile, and (1.6 km) one-mile buffer zones of each station. Connectivity measurement is a variable indicating the connectivity of streets. To obtain the connectivity measure, a street network layer was overlaid on the buffer layer and the number of intersections within the buffer zones was calculated. Street network data were obtained from the U.S. Census Tiger 2000 files (https://www.census.gov/geo/maps-data/data/tiger.html). The connectivity of a station increases as the number of intersections within station buffer zone increases.

Land-Use Mix Index. Three land-use types, namely residential, commercial, and industrial, are considered in this study. A land-use mix index is used to capture how evenly the square footage of commercial, residential, and industrial floor area is distributed within station buffer zones [(0.4 km) quarter-mile, (0.8 km) half-mile, and (1.6 km) one-mile buffers]. The land-use mix index is calculated as follows:

$$\text{Land-use mix} = [(-1)/\ln n] \times \sum_{i=1}^{n} p_i \ln p_i \tag{1}$$

where p_i = percentage of land use type i of the total land area; and n = total number of different land-use types. The land-use mix ranges from 0 (homogeneous land use, such as in rural areas or suburban subdivisions) to 1 (most mixed, such as diverse city centers) (Frank et al. 2004). Land-use data were originally acquired from the 2010 Maryland Property View data set (http://www.mdp.state.md.us/OurProducts/PropertyMApProducts/MDPropertyViewProducts.shtml), which are point-based data that include X, Y coordinates of properties, land acres, and land-use types including residential, commercial, and office of each property.

Accessibility. A gravity-based accessibility measure is used to define accesibility from one zone to all other zones. The gravity-based accessibility measure provides accurate estimates of the accessibility of zone i to opportunities in all other zones j in the region, where fewer and/or more-distant opportunities provide diminishing influences (Geurs and Wee 2004). The accessibility measure for zone i in a region with n TAZs ($i = 1,2, \ldots, n$), A_i, is represented as a function of number of opportunities in zone j ($j = 1,2, \ldots, n$) and impedance function between zones i and j as follows:

$$A_i = \sum_{i=1}^{n} O_j f(C_{ij}) \tag{2}$$

where A_i = accessibility for TAZ i; O_j = number of relevant opportunities in TAZ j; C_{ij} = travel time or monetary cost for a trip from TAZ i to TAZ j; and $f(C_{ij})$ = impedance function measuring the spatial separation between TAZ i and TAZ j;

The impedance function, $f(C_{ij})$ is an indicator of the difficulty of travel between TAZ i and TAZ j. A commonly used mathematical formula of the impedance function $f(C_{ij})$ is based on the theoretical work of Wilson (1971), and is expressed as $f(C_{ij}) = \exp(-\beta C_{ij})$, where β is an empirically calibrated parameter. Employment data used to represent the opportunities in TAZ_j were obtained from LEHD (2008).

Distance to CBD. The central city remains the main trip attractor of the Baltimore–Washington Metropolitan region. The authors would expect that stations that are closer to the Central Business District (CBD) would have higher ridership.

Walk Score. There are many walkability indices cited in the literature. Based on the area scale, level of detail, and data availability for this study, the authors adapted the Walk Score method to calculate the walkability of each census tract in the study area (Walk Score 2011). A distance decay function was applied to calculate walk score of a variety of amenities that were weighted by the importance. Then the walk scores were aggregated to census tract for further analysis. The walk score uses data from a number of sources, such as amenity data from Quarterly Census of Employment and Wages (QCEW) 2008 (https://www.dllr.state.md.us/lmi/emppay/) and road network data from U.S. Census Tiger file 2000.

Sociodemographic Variables

Median household income and mean vehicle per occupied housing unit within the catchment area of a given station were used. Previous literature suggests that low-income households and households without access to vehicles tend to rely more on transit than higher-income households and households owning one or more vehicles. The authors also included median age, percent of different racial groups, and homeownership in the dataset.

Tables 2 and 3 present the descriptive statistics for the dependent and independent variables. Since people who are taking commuter rail behave differently from the people who are taking light rail, the model was split into different modes. Due the small sample size of the Baltimore metro stations, the authors decided to combine different light rail stations together to compare the results.

Model Specification

Ordinary least-squares (OLS) regression was used to estimate the direct ridership model based on Maryland data. Equation (3) shows the linear relationship between station boardings Y_i and all the independent variables. Transit services, land-use characteristics, and sociodemographic variables are represented by X_1, X_2, and X_3, respectively; α = the constant term; β_1, β_2, and β_3 = the coefficients estimated from the linear regression, and ε_i = the unobserved random error

$$Y_i = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon_i \tag{3}$$

Estimating Direct Ridership Model

Several model specifications were tested using different combinations of the independent variables. The authors also tested the sensitivity of transit ridership in response to the built environment in different transit catchment areas and found that there is no significant difference on model's predictive power when applying different catchment areas. Nevertheless, the authors found that land-use variables performed better when using the 0.8 km (half-mile) buffer as the catchment benchmark in the model. Variables with multicollinearity issue were eliminated by inspection of the Variance Inflation Factor (VIF) and the correlation matrix. All insignificant independent variables (p-value > 0.01) were systematically eliminated from the final model. Variables with p-value close to 0.01 were kept in the model if they are important explanatory variables intended to test for policy guidance or they showed considerable explanatory power. Several tests were conducted to assure all the OLS regression assumptions are satisfied. Scatter plots were created to check the linearity of the rail boardings and individual independent variables. Heteroscedasticity was examined by generating plots of the residuals versus predicted values, and residuals versus independent variables.

For the final models, elasticities ρ_i were evaluated to quantify the impact of the various built-environment variables, transit service

J. Urban Plann. Dev., 2016, 142(4): 04016017

Table 3. Descriptive Statistics of Light Rail, Baltimore Metro, and WMATA Stations ($N = 73$)

Variable	Minimum	Maximum	Mean	Standard deviation
Boardings	113.0	13,865.0	2,938.7	2959.6
Park and ride	0.0	1.0	0.6	0.5
Bus	0.0	1.0	0.8	0.4
Level of service	20.0	60.0	41.9	20.0
Distance to nearest station	0.0	3.2	0.9	0.6
Terminal	0.0	1.0	0.2	0.4
Connectivity index	0.0	4.3	0.6	1.1
Population density within 0.8 km (half-mile) buffer	0.0	18,754	6,105	4,940
Employment density within 0.8 km (half-mile) buffer	259.0	13,703	22,626	35,352
Land use mixed index within 0.8 km (half-mile) buffer	0.0	1.0	0.5	0.3
Street connectivity within 0.8 km (half-mile) buffer	4.0	347.0	113.5	104.2
Accessibility of within 0.8 km (half-mile) buffer	539,932.2	1,951,165.5	1,074,789.2	310,931.4
Accessibility of transit within 0.8 km (half-mile) buffer	155,979.4	1,722,464.6	777,796.9	359,076.2
Distance to CBD	6.91 km	72.65 km	41.14 km	21.96 km
Walk score	2.3	237.3	54.8	66.3
Number of vehicles owned by household within 0.8 km (half-mile) buffer	0.6	2	1.4	0.4
Household income within 0.8 km (half-mile) buffer	\$22,302	\$801,584	\$68,220	\$89,951
Percent of White within 0.8 km (half-mile) buffer	0.0	98.1%	38.7%	32.5
Percent of Black within 0.8 km (half-mile) buffer	0.0	97.4%	40.7%	35.5
Percent of Hispanic within 0.8 km (half-mile) buffer	0.0	44.0%	6.0%	9.0
Percent of Others within 0.8 km (half-mile) buffer	0.0	41.7%	9.2%	9.3
Median age within 0.8 km (half-mile) buffer	0	70	31	13.6
Percent of owner-occupied housing within 0.8 km (half-mile) buffer	0.0	100%	39.7%	27.6

factors, and sociodemographic characteristics. The calculation of the elasticities for all the significant variables (including both discrete and continuous variables) is shown in Eqs. (4) and (5), respectively, where σ_i = coefficient; and $\bar{X}$ = mean of the corresponding variable

$$\rho_{\text{idiscrete}} = \frac{\exp(\sigma i) - 1}{\exp(\sigma i)} \times 100\% \qquad (4)$$

$$\rho_{\text{icontinuous}} = \sigma_i \times \overline{X}_i \times 100\% \qquad (5)$$

A power analysis was conducted to assure the validity of the OLS regression analysis when sample size is considered as small. A power analysis is used to statistically examine the probability of a regression analysis correctly rejects a false null hypothesis (i.e., Type II errors). A widely cited measurement of power to justify the effect of sample size on the analysis results was applied (Cohen 1988; VanVoorhis and Morgan 2007; Maxwell and Kelly 2008).

The effect size is defined as f^2 and power is defined as λ

$$f^2 = \frac{R^2}{(1 - R^2)} \qquad (6)$$

where R^2 = unadjusted R^2

$$\lambda = f^2 \times (p + \nu + 1) \qquad (7)$$

where f^2 = effect size; p = number of independent variables; n = number of observations; and $\nu = n - p - 1$, where v is the degree of freedom for error.

Applying the model results from Table 4, the authors computed $\lambda = 30$. Then the authors checked Cohen's suggested power value when $\lambda = 30$ and found 0.76. This means there is a very high probability of correctly rejecting a false null hypothesis with the current regression model specifications and parameters.

Table 4. Regression Model Predicting MARC Station Boardings ($N = 39$)

Variable	Coefficient	Significance level[a]	Elasticity
Constant	−7.815	—	—
Park-and-ride	[b]	—	—
Bus	3.302	[c]	96.32
ln (distance to CBD)	0.592	—	—
ln [population density within 0.8 km (half-mile) buffer]	0.153	—	—
ln [employment density within 0.8 km (half-mile) buffer]	0.056	—	—
ln [percent of White within 0.8 km (half-mile) buffer]	−0.383	—	—
ln [median age within 0.8 km (half-mile) buffer]	0.648	—	—
ln [percent homeownership within 0.8 km (half-mile) buffer]	−0.074	—	—
R^2	0.360	—	—

[a]Indicates coefficient that is statistically significant at the 0.001 level.
[b]Indicates that park-and-ride is a singular variable.
[c]Indicates coefficient that is statistically significant at the 0.05 level.

Results

Tables 4 and 5 present the results of the final regression models of commuter rail stations and light rail stations. The results from commuter rail stations suggest that bus connections are statistically significant in the model. Also, the R^2 shows that roughly 36% of the variation in the boarding of commuter rail stations is explained by all the variables in combination. Among all the independent variables, only feeder bus connections are statistically significant and have a positive sign. The effect of feeder bus services is dominant and the elasticity is 96.32%. While comparing with the noncommuter rail stations, feeder buses have a stronger impact on rail boardings.

Noncommuter rail stations were grouped into three samples to show the differences: light rail (LRT) alone, LRT and the Baltimore

Table 5. Regression Model Predicting LRT, Baltimore Metro, and WMATA Station Boardings

Variable	LRT model			LRT and Baltimore Metro model			LRT, Baltimore Metro, and WMATA model		
	Coefficient	Significance	Elasticity	Coefficient	Significance	Elasticity	Coefficient	Significance	Elasticity
Constant	6.540	—	—	4.783	—	—	4.666	—	—
Terminal	0.615	—	—	0.48	a	38.12	0.528	b	41.02
Park-and-ride	0.103	—	—	0.195	—	—	0.147	—	
Bus	0.515	a	40.24	0.520	a	40.54	0.537	a	41.55
ln (distance to the nearest station)	0.049	a	2.59	0.034	—	—	0.042	a	3.91
ln (distance to CBD)	0.901	—	—	−0.158	—	—	−0.244	a	−28.2
Level of service	c	—	—	0.980	b	24.5	0.132	b	39.6
Number of employment	0.00001	b	13.02	0.00001	b	18.11	0.000009	b	20.36
Number of population × population density	−0.000000002	—	—	0.0000000001	—	—	0.000000001	—	—
N		32			46			73	
R^2		0.566			0.723			0.812	

Note: NA indicates that this variable is a singular variable.
[a]Indicates coefficient are statistically significant at the 0.05 level.
[b]Indicates coefficient is statistically significant at the 0.01 level.
[c]All the LRT stations have the same level of service.

metro, and the combination of LRT, Baltimore metro, and WMATA stations. The authors investigated whether there is a significant difference among different rail systems. Three models were developed to address different policy questions. On the other hand, the authors would like to see whether general policy guidance should be applied across the entire region. Two variables are significantly associated with station boardings across all three samples: bus connections and employment. Both variables have expected signs. The results suggest that bus connections and higher employment will increase the station boardings for all rail lines. Population is not significantly associated with the station boardings. This suggests that employment plays a more important role in promoting transit use through affecting the built environment around the station area, even though the magnitude of the coefficient of the employment variable is modest. The rail station catchment area shows a significant and expected relationship in the LRT model only. The higher level of service is statistically significant when the Baltimore metro and WMATA are included with the LRT but not for LRT alone. The R^2 shows that roughly 56.6% of the variation in the boardings of LRT stations is explained by all the variables, 72.3% of the variation in the boardings of LRT and Baltimore metro stations is explained by all the variables, and 81.2% of the variation in the boardings of LRT, Baltimore Metro, and WMATA stations is explained by all the variables. The variable for terminal stations is statistically significant and has a positive sign for LRT and Baltimore Metro model and the LRT, Baltimore Metro, and WMATA Metro model. Terminal stations often attract more boardings since end-line stations attract riders from further areas. The elasticities of terminal station for the model of LRT and Baltimore Metro, and the model of LRT, Baltimore Metro and WMATA stations are 38.12 and 41.02%, respectively. Distance to the nearest station has a positive sign in both LRT model and model with LRT, Baltimore Metro, and WMATA stations. In the downtown areas, stations are located closer to each other, which will trigger higher numbers of boardings. Distance to CBD is only significant in the LRT, Baltimore Metro, and WMATA model. Most of the LRT stations are located in downtown Baltimore. One of the major purposes of the LRT is to serve the residents in suburbs who go downtown for work. So, as the distance increases from CBD, the ridership goes up. This is intuitive. On the other hand, the other two models included both Baltimore Metro and WMATA Metro stations. Since these two rail systems serve people across a large scale, the centrality of stations still play a significant role in terms of attracting higher ridership. So, as distance increases from CBD, the ridership goes down.

Parking at stations bears further discussion. The initial intent was to determine the significance of parking as a factor in transit station boardings. One only needs to view parking at the stations to understand it is a significant factor. Parking was treated as a dummy variable, and every commuter rail station has parking, eliminating the possibility of statistical tests.

Policy Implications

The DRMs provide useful insights into how land-use and transportation attributes interact to influence transit boardings at the station level. The results suggest that these influences have significant impacts on transit ridership but that the impacts differ depending on whether light rail or commuter rail is being analyzed. For light rail stations, employment within a 0.8 km (half-mile) buffer area of the station, transit service level, feeder bus connectivity, station location in the CBD, distance to the nearest neighboring station, and whether or not the station is a terminal station all affect transit boardings. For commuter rail, (MARC stations in this study), the impact of feeder bus connections is found to be significant. This suggests that commuter rail behaves differently from light rail. For light rail, employment is the most significant predictor of station boardings. Increasing the employment in a transit station area by 1% can result in a 13.02% increase in boardings. For light rail, Baltimore metro, and WMATA Metro stations, the employment elasticity is 20.36%, which suggests that employment plays a critical role in affecting transit ridership for the whole region.

The DRMs results have implications for areas wishing to increase transit ridership. They show that increasing employment can impact transit ridership and increase transit boardings at certain transit stations. However, policy-makers should use this link with caution. Increasing density will likely induce increase in demand for transit usage. Thus, any policy related with densification should also be combined with necessary changes in transit service level, parking capacity, and feeder bus services to meet increased demand and to take full advantage of the transit system. Areas desiring to increase transit ridership should consider zoning regulations and site-design requirements that allow for denser development around transit stations.

The DRMs cannot substitute for traditional mode choice models, which are essential for determining system-level characteristics of transit and for developing forecasts on ridership on individual transit lines. At the same time, the DRMs can complement traditional mode choice models by estimating the effects of changes in station boardings, and thus transit ridership, resulting from changes in urban form.

The conclusions in this research also have implications beyond transportation. They can indicate how housing decisions can affect the transportation system and, in turn, can be supported by the system. Employment and economic development policies would benefit from an understanding of the factors that affect transit boardings and station utilization. Companies may consider access to transit as part of their location decision. These conclusions are particularly important for retailers who rely on local shopping to support their business.

Conclusions

The DRMs model can be useful to provide insights on how different land-use and transit attributes can affect boardings. To date, the DRMs have been used to estimate station-level transit ridership for rail investments and expansion proposals for many places (R. L. Banks & Associates 2001; Blash et al. 2005; LS Transit Systems 1998). However, no DRM model has been estimated for the State of Maryland. This study has contributed to the state of practice by applying DRMs for all the rail stations in Maryland. This study used the readily available data, a low-cost and fast-action approach that can be used as complementary to the current MSTM by providing useful order-of-magnitude estimates of travel demand effects of TOD scenarios. In addition, the study provides a platform that can be accessed by a large number of transportation planners and policy-makers in prioritizing areas for development.

The results suggest that impacts of the built environment show differences for light rail and commuter rail. For light rail stations, employment at 0.8 km (half-mile) buffer areas, service level, feeder bus connectivity, stations located in the CBD, and terminal stations are significant factors affecting ridership. Among these variables, employment is the most significant predictor of station boardings ($p < 0.001$) and the results are consistent across all the models. Increasing the employment by 1% can result in a 13.02% increase in boardings. For MARC stations (commuter rail), only feeder bus connections are found to be significant, which suggests that commuter rail behaves differently from light rail. Adding more bus access to the commuter rail station will be a good policy for improving the commuter rail ridership. Parking was not included in the MARC model since all the MARC stations have parking available. Therefore, the authors need additional variables to test the significance of parking of MARC stations.

In summary, DRMs can provide estimates of station boardings without relying on complicated transportation demand models and an extensive data-collection process. It can capture the key features of the built-environment characteristics, transit service attributes, and their relationships with station boardings, which can provide timely policy guidance on how to improve the transit usage. The current DRMs model also has several limitations. First, the DRMs were developed for the State of Maryland only. Future research should be carried out for the surrounding regions (like Washington DC or Virginia) to get a general pattern. Second, other important factors should be incorporated in the model, such as safety and transit reliability attributes. Third, more-detailed parking information, such as parking fees, should also be included in the model.

Acknowledgments

The authors are thankful to Maryland Department of Transportation (MDOT) for the continued support in the development of models used in this study. The opinions and viewpoints expressed are entirely those of the authors, and do not necessarily represent policies and programs of any agency. The authors would like to thank the two anonymous reviewers for their very constructive and insightful comments. The authors would also like to thank Daniel Engelberg for his proofreading of the manuscript.

References

Adams, M. A., et al. (2009). "Validation of the neighborhood environment walkability scale (NEWS) items using geographic information systems." *J. Phys. Act. Health*, 6(1), S113.

Antipova, A., Wang, F., and Wilmot, C. (2011). "Urban land uses, socio-demographic attributes and commuting: a multilevel modeling approach." *Appl. Geogr.*, 31(3), 1010–1018.

Blash, L., Duncan, M., Hranac, R., and Khazai, B. (2015). "AC transit telegraph MIS project: Integrating GIS and travel model evaluation methodologies." ⟨http://www-laep.ced.berkeley.edu/classes/cp255.people/cp255-bx/Project/final.html⟩ (Nov. 10, 2005).

Cervero, R. (2006). "Alternative approaches to modeling the travel-demand impacts of smart growth." *J. Am. Plann. Assoc.*, 72(3), 285–295.

Cervero, R. (2007). "Transit-oriented development's ridership bonus: A product of self-selection and public policies." *Environ. Plann. A*, 39(9), 2068–2085.

Cervero, R., and Murakami, J. (2009). "Rail and property development in Hong Kong: Experiences and extensions." *Urban Stud.*, 46(10), 2019–2043.

Chu, X. H. (2004). "Ridership models at the stop level." *NCTR-473-04, BC137-31 Final Rep.*, Center for Urban Transportation Research, Univ. of South Florida.

CMRT (Central Maryland Transportation Alliance). (2009). "Central Maryland TOD strategy: A regional action plan for transit-centered communities." ⟨http://ctod.org/pdfs/2009CentralMarylandTODStrategy.pdf⟩ (Mar. 11, 2013).

Cohen, J. (1988). *Statistical power analysis for the behavioral science*, 2nd Ed., Erlbaum, Hillsdale, NJ.

Duduta, N. (2013). "Direct ridership model of Mexico City's BRT and metro systems." *Transportation Research Board 92nd Annual Meeting*, Washington, DC.

Fehr & Peers. (2013). "DRM (Direct Ridership Models)." ⟨http://www.fehrandpeers.com/drm-direct-ridership-models/⟩ (Jul. 17, 2013).

Filion, P. (2001). "Suburban mixed-use centres and urban dispersion: What difference do they make?" *Environ. Plann. A*, 33(1), 141–160.

Frank, L., Martin, M., and Schmid, T. (2004). "Obesity relationships with community design, physical activity, and time spent in cars." *Am. J. Preventive Med.*, 27(2), 87–96.

Geurs, K. T., and Wee, B. V. (2004). "Accessibility evaluation of land-use and transport strategies: Review and research directions." *J. Transp. Geogr.*, 12(2), 127–140.

Gutiérrez, J., Cardozo, O. D., and García-Palomares, J. C. (2011). "Transit ridership forecasting at station level: an approach based on distance-decay weighted regression." *J. Transp. Geogr.*, 19(6), 1081–1092.

Hadas, Y., and Ranjitkar, P. (2012). "Modeling public-transit connectivity with spatial quality-of-transfer measurements." *J. Transp. Geogr.*, 22, 137–147.

Kaplana, S., Popoksa, D., Pratoa, C. G., and Ceder, A. (2014). "Using connectivity for measuring equity in transit provision." *J. Transp. Geogr.*, 37, 82–92.

Kuby, M., Barranda, A., and Upchurch, C. (2004). "Factors influencing light-rail station boardings in the United States." *Transp. Res. Part A: Policy Pract.*, 38(3), 223–247.

LEHD (Longitudinal Employer-Household Dynamics). (2008). ⟨http://lehd.ces.census.gov/⟩ (Mar. 10, 2013).

LEHD (Longitudinal Employer-Household Dynamics). (2013). ⟨http://lehd.ces.census.gov/⟩.

Li, J., Yang, G., Co, J., and Li, L. (2015). "Forecasting method for urban rail transit ridership at the station-level using a weighted population variable and genetic algorithm back propagation neural network." *5th Int. Conf. on Transportation Engineering*.

LS Transit Systems. (1998). "Around the Bay rail study." ⟨http://sccrtc.org/wp-content/uploads/2011/07/980700-AroundTheBayRailStudy.pdf⟩.

Marshall, N., and Grady, B. (2006). "Sketch transit modeling based on 2000 census data." *Transp. Res. Rec.*, 1986(1), 182–189.

Maxwell, S. E., and Kelley, K. (2008). "Sample size planning for statistical power and accuracy in parameter estimation." *Annu. Rev. Psychol.*, 59, 537–563.

MDOT (Maryland Department of Transportation). (2009). "2009 Maryland transportation plan." ⟨http://www.mdot.maryland.gov/Office_of_Planning_and_Capital_Programming/Plans_Programs_Reports/Reports/MTP/09MTP.pdf⟩.

MDP (Maryland Department of Planning). (2010). "Property view data." ⟨http://www.mdot.maryland.gov/Office_of_Planning_and_Capital_Programming/Plans_Programs_Reports/Reports/MTP/09MTP.pdf⟩.

MDP (Maryland Department of Planning). (2013). "Smart growth planning topics." ⟨http://www.mdp.state.md.us/OurWork/smartGrowth.shtml⟩ (Jul. 30, 2013).

Mishra, S., Welch, T., Torrens, P., Fu, C., Zhu, H., and Knaap, E. (2015). "A tool for measuring and visualizing connectivity of transit stop, route and transfer center in a multimodal transportation network, public transport." *Public Transp.*, 7(1), 77–99.

Mishra, S., Welch, T. F., and Jha, M. K. (2012). "Performance indicators for public transit connectivity in multi-modal transportation networks." *Transp. Res. Part A: Policy Pract.*, 46(7), 1066–1085.

Mitra, R., and Buliung, R. N. (2012). "Built environment correlates of active school transportation: neighborhood and the modifiable areal unit problem." *J. Transp. Geogr.*, 20(1), 51–61.

MTA (Maryland Transit Administration). (2011). "Transit ridership.".

NCSG (National Center for Smart Growth). (2010). "Research and education. Data inventory." (May 20, 2013).

NCSG (National Center for Smart Growth). (2012a). "Maryland scenarios project draft final year 3 report.".

NCSG (National Center for Smart Growth). (2012b). "Maryland statewide travel demand model." ⟨http://smartgrowth.umd.edu/statetransportationmodel⟩ (Mar. 20, 2013).

Pulugurtha, S. S., Nambisan, S. S. and Maheshwari, P. (2006). "Estimating pedestrian counts in urban areas for transportation planning and safety analyses." *Applications of Advanced Technology in Transportation, 9th Int. Conf.*

R.L. Banks & Associates. (2001). "Kansas City region commuter rail study: Initial corridor screening. Mid-America regional council." ⟨http://www.marc.org/Transportation/Plans-Studies/Transportation-Plans-and-Studies/Special-studies-and-projects/special-studies-pdfs/commuterrail_2002/Assessment1.aspx⟩.

Sarker, A., Mishra, S., Welch, T., Golias, M., and Torrens, P. (2015). "Model framework for analyzing public transit connectivity and its application in a large-scale multimodal transit network." *Compendium of Papers in 94th Annual Board Meeting of Transportation Research Board*, National Research Council, Washington, DC.

U.S. Census. (2010). "American fact finder." ⟨http://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml⟩ (Mar. 20, 2013).

U.S. Census. (2013). "American fact finder." ⟨http://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml⟩ (Mar. 20, 2013).

VanVoorhis, C. R. W., and Morgan, B. L. (2007). "Understanding power and rules of thumb for determining sample sizes." *Tutorials Quant. Methods Psychol.*, 3(2), 43–50.

Walk Score. (2011). "Most walkable cities in the United States, Canada, and Australia on Walk Score." ⟨http://www.walkscore.com/cities-and-neighborhoods/⟩ (Jul. 30, 2013).

Welch, T., and Mishra, S. (2013). "A measure of equity for public transit connectivity." *Transp. Geogr.*, 33, 29–41.

Wilson, A. G. (1971). "A family of spatial interaction models, and associated developments." *Environ. Plann.*, A3.1, 1–32.

WMATA (Washington Metropolitan Area Transit Authority). (2011). "Transit ridership.".