5-2023

Enhancing station level Direct-Demand models with Multi-Scalar accessibility indicators

Luis Enrique Ramos-Santiago

Transportation Research Interdisciplinary Perspectives 19 (2023) 100834

Contents lists available at ScienceDirect

Transportation Research Interdisciplinary Perspectives

journal homepage: www.sciencedirect.com/journal/transportation-research-interdisciplinary-perspectives

Enhancing station level Direct-Demand models with Multi-Scalar accessibility indicators

Luis Enrique Ramos-Santiago

Clemson University, College of Architecture, Arts, and Construction, Nieri Department of Construction, Development, and Planning, 2-215 Lee Hall Clemson, SC 29634, USA

ARTICLE INFO

Keywords:
Transit
Forecasting
Direct demand
Accessibility
Station
TOD

ABSTRACT

Direct-demand models (DDM) are increasingly being used for a diversity of transit research and practice purposes. Yet few station-level DDM studies have explored the use of composite indicators of metropolitan accessibility in predicting demand. After all, provision of access to metropolitan destinations is one of the main goals of rapid-transit systems. Furthermore, to this author's knowledge no study has explored potential interactions with local-level accessibility indicators that are typically included in station level transit DDMs. This study explores these possibilities and uses Los Angeles multimodal rapid-transit network as a representative case study of a system that operates in a dispersed agglomeration where multiple sub-centers are linked. Multi-level generalized linear models were implemented where key predictors, including stations' metropolitan- and a local-accessibility indicators are regressed onto average weekday boardings. Furthermore, more general accessibility constructs were developed via EFA and implemented in models; and parameters non-stationarity was assessed via geographically weighted regressions. Results indicate that nodal metropolitan accessibility is a significant predictor of patronage in LA's rapid-transit network, and that its interaction with local-accessibility amplifies boardings and improves DDM models' explanatory power. More general constructs of accessibility at metropolitan and local-scale were derived via EFA and these resulted in a more parsimonious model with equal predictive power. Land-use and transit planners would benefit from including an accessibility lens in their DDM modeling. Practical applications of these type of models include TOD scenario planning, comparative route alignment studies, system expansion studies, and for didactic purposes given the ability of accessibility measures to capture land-use/transportation interactions.

1. Introduction

Direct-demand modeling (DDM) in transit studies has increased and diversified, becoming more nuanced and effective for a diversity of modes, outcomes, geographical scales, and research/practice purposes (Ramos-Santiago, 2021, Ramos-Santiago et al., 2022). DDM has been particularly useful in ridership and performance studies at station-, line-, and/or system-levels; for policy analyses; comparative performance studies; and basic research. Furthermore, DDMs can inform feasibility, scenario, and/or sketch planning studies, including route alternatives studies and Transit Oriented Development (TOD) scenario projection. It is usual to find a triad of vectors for organizing candidate predictor variables for DDM models. These are regressed onto various boarding count measures and other types of service consumption variables. Covariates are typically organized along *land-use* and *built-environment* characteristics, *socioeconomics*, and *transit service levels*. More recent studies include *network topology* attributes, among other context-relevant controls.

Less frequently used in station-level DDM studies are metropolitan-scale *accessibility indicators* (a.k.a. nodal cumulative opportunity measures). And seemingly no study has explored interaction between *metropolitan* and *local* accessibility indicators, excluding mode-choice studies that have addressed the independent effects of each on the propensity of users to use transit for the most part assessed within discrete-choice logit or probit models.

Accessibility is used in this study as 'the extent to which land-use and transport systems enable (groups of) individuals to reach activities or destinations (i.e., opportunities) by means of a (combination of) transport mode(s)' (Geurs and Van Wee, 2004). The role and contribution of accessibility on aggregate transit travel within the context of DDM modeling is explored. Understanding the influence, if any, of *metropolitan-* and *local-accessibility* indicators and their interaction onto a

E-mail address: eramoss@clemson.edu.

https://doi.org/10.1016/j.trip.2023.100834
Received 19 August 2022; Received in revised form 24 March 2023; Accepted 2 May 2023
Available online 16 May 2023

continuous outcome variable, average weekday boardings is the main focus. This framework is applied to a specific transit technology, rapid transit, using stations as the main unit of analyses. Schedule-based weekday transit service frames the temporal scale.

The main hypotheses are tested on a set of generalized linear regressions and the potential of more general latent constructs of accessibility is investigated in pursuit of theoretical insights and a more parsimonious DDM model. This relied on Exploratory Factor Analysis (EFA) and results indicate that a more ample accessibility construct is germane for transit systems that operate in more dispersed polycentric agglomerations where bus-rail interaction is critical (Ramos-Santiago, 2021, Ramos-Santiago et al., 2022). Finally, a spatial analysis modeling technique, geographically weighted regression (GWR) was applied to the data to explore the possibility of regression parameter's non-stationarity and consider the modeling and practical implications.

Special interest is placed on two coupled and interacting systems: 1- the *land use* structure, and 2- the *transportation system*, which includes all interacting multimodal infrastructures, modal characteristics, and service levels. Land-use professionals (e.g., city planners, urban designers, architects, developers) and transportation professionals (e.g., transportation planners, civil and traffic engineers) are not, however, necessarily trained under the same curricula or school and often rely on distinct theoretical frameworks and disciplinary traditions (Levinson et al., 2017). Yet, their planning and design decisions overlap and interact in the resulting urban plexus.

In this study an interdisciplinary approach is taken and geography sub-disciplines in *spatial analysis* and *accessibility* contribute key theoretical and methodological approaches that inform the operationalization of *local-* and *metropolitan-accessibility* indicators. Additionally econometric models derived from economics are fundamental to the DDM modeling approach, its specifications, interpretation, and data treatment. Together the different disciplines complement transportation science and transit planning, and inform hypotheses, methods, measures, and specification of DDM models in pursuit of theoretical, practical, policy and methodological insights for transit planning and station-level forecasting.

This paper is structured as follows: Section-1 summarizes a review of literature from various disciplines that includes transit ridership determinants; DDM modeling; accessibility measures; and recent accessibility/transit studies. Section-2 presents main propositions and hypotheses while Section-3 details methods and instruments. This is followed by presentation of results in Section-4. And Section-5 contains insights and conclusions together with a review of limitations and lines of possible future research.

2. Literature review

This literature review draws from several disciplines and is organized in five overlapping parts that relate to key theoretical and methodological issues. These inform model specifications, controls, and interpretation. The first part presents a general overview of transit ridership determinants, while the second discusses transit ridership studies that relied on DDM models. Part three reviews more recent studies where various accessibility measures are considered in the evaluation of transit systems performance and/or transit ridership (**Appendix A**). Parts four and five focus on the concepts of *metropolitan-* and *local*-accessibility, respectively; latent constructs of general accessibility; and reviews associated measures.

2.1. Transit ridership determinants

Taylor et al. (2009) and Taylor and Fink (2003) organize transit ridership determinants along two main classifications: *external* factors, which are outside direct control of transit planners and managers; and *internal* factors, which are susceptible to influence by transit planners and managers. Most transit scholarship has emphasized the importance of socioeconomic factors and built environment factors as key *external* influences on patronage and acknowledge the important roles of *internal* factors such as fare and service level decisions (Cervero and Duncan, 2002; Balcombe et al., 2004; Mees, 2009; Walker, 2012).

Other researchers have reported higher population and job densities; higher share of land-use mix; pedestrian-friendly environments; lower unemployment; and lower share of vehicle ownership are also associated with higher counts of transit trips (Cervero, 2001; Kim et al., 2007; Guerra and Cervero, 2011; Chakraborty and Mishra, 2013; Ramos-Santiago and Brown, 2016). Furthermore, lower fares; increased service frequency; and better service coordination have been identified as promoters of more transit usage (Ramos-Santiago, 2021; Taylor and Fink, 2003; Balcombe et al., 2004; Guerra and Cervero, 2011; Ramos-Santiago and Brown, 2016; Brown and Thompson, 2008; 2012; Mees et al., 2010; Currie et al., 2011). These results have emerged from studies done at a variety of geographic scales in locations throughout North America, Europe, and Australasia, and relying on a variety of analytical methods. More recent studies highlight the importance of bus-rail connectivity at station-level as a significant factor behind higher patronage, particularly in the context of large polycentric agglomerations (Ramos-Santiago, 2021).

2.2. Direct-Demand models

Transportation research and demand analyses rely in great part on discrete-choice models and more recently on direct-demand models (DDM). Compared to discrete-choice models, which rely on survey data and modeling individual decision-making via logit or probit methods, DDM works with cross-sectional aggregate travel data and multivariate regressions (Wardman et al., 1994). This flexible statistical modeling technique is considered effective; more appropriate; and a less costly approach to transit ridership forecasting and policy analyses when compared to the traditional 4-step urban transport modeling protocol considered more complex, timely, and costly (Chen and Zegras, 2016; Kuby et al., 2004; Cervero, 2006; Cervero et al., 2010; Zhao et al., 2014; Durning and Townsend, 2015).

The general equation for a station-level DDM ridership model is (**Equation (1)**):

$$Boardings^{station} = f(SE, LU.BE, TS, NT) \tag{1}$$

Where explanatory variables are organized in four main groups; SE: socioeconomic attributes of households and employment levels within stations' pedestrian service areas; LU.BE.: land-use/build-environment characteristics of stations' service areas; TS: transit service levels of both line-haul and bus feeder services; and NT: rapid-transit network attributes (Ramos-Santiago, 2021).

2.2.1. DDMs in Rapid-Transit ridership studies

Studies by Parsons Brinckerhoff et al. (1996) and Kuby et al. (2004) are often cited in DDM literature. Both extend Pushkarev & Zupan (1977) influential work on transit ridership and cost-effectiveness that focused on land-use characteristics (e.g., density), distance from CBD, among other factors. Cervero et al. (1995) study on access modes and catchment areas for BART's system is also often cited. These studies documented positive associations between transit patronage and employment density, Park & Ride facilities, feeder bus network connectivity, and the station's status as either a terminal or transfer hub.

Kuby et al. (2004) also found positive relationships between transit patronage and airports located within 800 m of the station and proximity to international borders; and noted the non-significance of stations located in the CBD district. Special attractors/generators near stations also have strong and positive associations with more boardings (Ramos-Santiago and Brown, 2016; Foletta et al., 2013); and reliable results are found across heavy-rail, light-rail, and bus rapid-transit systems (Foletta et al., 2013; Gutiérrez et al., 2011).

In general, transit DDM studies organize explanatory variables along socioeconomics, transit service levels, and land-use/built-environment vectors (Ramos-Santiago and Brown, 2016; Foletta et al., 2013). Other studies have included stations' network topological attributes in model specifications (NT; Chen and Zegras, 2016; Ramos-Santiago and Brown, 2016; Kuby et al., 2004; Zhao et al., 2014; Durning and Townsend, 2015; Foletta et al., 2013; Gutiérrez et al., 2011; Sohn and Shim, 2010); and inter-station spacing whilst controlling for modal differences (e.g., streetcar/light-rail/metro) that surfaced as significant in some studies (Chen and Zegras, 2016; Ramos-Santiago and Brown, 2016; Zhao et al., 2014; Foletta et al., 2013). More recently the key role and influence of bus-rail connectivity and bus feeder service characteristics was emphasized in Ramos-Santiago (2021) and Ramos-Santiago et al. 2022).

Factors found to be negatively associated with station ridership in DDM studies and that register high statistical significance and large effects are household automobile availability, higher fare, and travel time (Chen and Zegras, 2016; Ramos-Santiago and Brown, 2016; Zhao et al., 2014; Ramos-Santiago et al., 2015). The strong negative influence of automobile availability is to be expected in the U.S. context as transit travel is characterized as an 'inferior good' (McLeod et al., 1991) when considering the central role, infrastructural, policy, and land-use regulatory bias towards automobile mobility.

In summary, station, and station-to-station DDM models have diversified and reflect greater explanatory power and more nuanced modeling approaches. Several factors have been found to be consistently associated with higher or lower patronage. However, to this author's knowledge few station-level DDM studies, at least in the English-language literature, have incorporated metropolitan-scale accessibility measures in their model specifications, nor have explored potential interactions with local-accessibility measures (**Appendix A**).

2.3. Accessibility and its measures

In general, accessibility can be defined as 'the ease of reaching goods, services, activities, and destinations, which together are called opportunities' (Litman, 2021). As noted by Levinson and Wu (2020), there is robust evidence for accessibility explaining location decisions, including commuting time, employment rates, mode shares, real estate prices and density, income and productivity, and investment decision. These authors also note that urban agglomeration economies (e.g., *localization* and *urbanization* economies) reflect the tendency of people, resident, workers and/or firms to be in proximity and reduce travel distance to engage in routine activities. That is, in general people value locations with greater access to people and opportunities 'they care about' in order to reduce travel costs and to be more productive and increase earnings (all else equal; Levinson and Wu, 2020).

Geurs et al. (2004) reviewed a number of accessibility measurement approaches in evaluating land-use and transport strategies and developments and classify them along three main analytical lenses: location accessibility, individual accessibility, and economic benefits of accessibility. Connected to the purpose and interests in this study are *location-based* measures that focus on accessibility to spatially distributed activities, which are typically used in urban planning and geographical studies. In assessing their usefulness, he identified four main criteria towards a more appropriate analysis. Namely, theoretical basis, interpretability and communicability, data requirements, and useability in social and economic evaluations. This, concluding that the following are key theoretically components in measuring accessibility in the context of land-use/transport studies: *land-use*, *transportation*, *temporal*, and *individual* components.

Ideally, an accessibility measure would take all four components in consideration and should be sensible to changes in the transportation system; sensible to changes in land-use system; consider spatial distribution of demand including competition effects; sensitive to temporal constraints of opportunities; and to the extent possible, and given data availability, it should consider individual needs and abilities. Fully addressing these, however, is a tall order that can serve as a guide for researchers but often practically impossible (Levinson and Wu, 2020). This obligates to recognize any violations of theoretical criteria in studies incorporating accessibility measures.

Location-based accessibility measures come in four distinct variations (Geurs and Van Wee, 2004): the more basic *contour measure*; Hansen's (1959) seminal *potential* model (**Equation (2)**); *adapted potential measures;* and *balancing factors*. The latter two which could be considered extensions of Hansen's model.

As noted by Horner (2004), origin-specific potential models can be developed for metropolitan studies on accessibility, urban form, commute travel, among other relevant topics as an extension of Harris (1954) and Hansen (1959). These measures are resultant of summing for a center i all possible gravity model interactions to m centers j. Specifying that local opportunities at origin (i) are not included in the calculation ($j \neq i$) helps differentiate from the typical cumulative and contour models (**Equation (3)**). As such, a basic modified Hansen model for an *origin-based* potential accessibility measure may be written as:

$$A_i = \sum\nolimits_{j,\, j \neq i}^{m} O_j {*} f\left(\mathrm{C_{ij}}\right) \tag{3}$$

Where:

A_i = accessibility at location i.

O_j = opportunities at location j ($j \neq i$).

C_{ij} = cost (distance or travel time) between location i and location j.

$f(*)$ = impedance function describing non-linear effect, ideally empirically derived.

Nevertheless, it is important to consider activity schedules and individuals' constraints for a more comprehensive and nuanced definition of accessibility (Karner, 2022). However, these more recent operationalizations require higher levels of disaggregation and data, which are difficult to obtain and opposite to the intent of DDM approaches that seek more manageable levels of resources and time.

2.4. Recent transit /Accessibility DDM studies

In the past 10 years there have been few DDM transit patronage studies that have controlled for a stations' nodal *metropolitan accessibility* (n = 4, **Appendix A**). More often *local accessibility* measures are found in studies that aim to assess the role of the local built environment on ridership (Aston et al., 2020). *Metropolitan accessibility* is sometimes operationalized as a more basic distance-to-CBD parameter or topological *centrality* measure. Yet, the potential interaction between *metropolitan-* and *local-accessibility* measures nor the policy implications that may flow from it, remained unexplored.

Recent studies reflect and assess a diversity of accessibility measures at both *local* and *metropolitan* (sometimes referred to as 'regional') scales. A variety of units of analysis, geographical scales, locations, and methods are implemented for example by Moniruzzaman & Páez (2012), Li et al. (2017), and others.

Moniruzzaman & Páez (2012) use the terms 'access *to* transit' and 'access *by* transit' when referring to local and metropolitan accessibility, respectively. Similarly, Li et al. (2017) refers to stations' '*attraction*' and '*radiation*' when referring to the similar concepts of local- and metropolitan-accessibility, respectively. The first study relies on mode-choice logit modeling and the second on a comparison of accessibility indicators. Yet, their operationalizations of accessibility differ between *local* and *metropolitan* scales. Furthermore, each study focuses on different units of analyses, which limits development of a cohesive theory and generalization.

Irrespective of unit of analyses or operalization, *metropolitan accessibility* emerges as a highly significant and positive factor in both discrete-choice as well as DDM transit ridership models. The more nuanced origin-based cumulative opportunities (potential) model for *metropolitan accessibility* is found in Gutiérrez et al. (2011); Moniruzzaman & Páez (2012); Chen and Zegras (2016); Cui et al. (2020); and Wu

& Levinson (2021). Other researchers rely on more basic measures like distance-to-CBD (Sung et al., 2014); Likert-scale measures based on users' perceptions of ease of access (Sung et al., 2014); or topological attributes of the transit network assessed via space-syntax or functional classifications (Li et al., 2017).

In regard to *local accessibility*, it is evident that a variety of measures, ranging from simple to complex, have been evaluated with significant results. Whether using a set of predictors along the 5Ds built-environment typology (Chen and Zegras, 2016); composite indicators (Lin et al., 2014); distance-decay weighted predictors (Gutiérrez et al., 2011); multimodal cost-based utility measures (Li et al., 2017); or more basic distance or time measures to nearest stop or station (Moniruzzaman and Páez, 2012; Chowdhury et al., 2016) *local accessibility* has also been found to be positive and significant for patronage in a variety of transit modes.

Few DDM ridership studies, however, incorporate measures for both *metropolitan* and *local* accessibility that feature similar underpinning constructs, such as origin-based cumulative opportunities models with distance-decay functions. This could facilitate comparison and explorations and strengthen internal validity. And none seem to have explored potential interactions between the two phenomena. In addition, recent urban sustainability studies highlight the need to pay more attention to accessibility to opportunities at various geographical scales (e.g., local, and regional) in studying and improving our understanding of associations between travel behavior and build-environment beyond the prevalent 3D's framework (Elldér, 2020); and emphasizing destination accessibility (Næss et al., 2019) as compared to other accessibility measurement approaches.

3. Main propositions

Proposition 1. *The regional accessibility to jobs, here considered as a proxy for a wider set of socioeconomic opportunities, is also a key station-level trip production factor in Los Angeles (LA) rapid-transit network. That is, a rapid-transit station's metropolitan accessibility, operationalized as an origin-station's nodal cumulative opportunity measure is posited to be a significant and positive factor in station-based DDM ridership models.*

Proposition 2. *A station's metropolitan accessibility and local accessibility (e.g., pedestrian-friendliness) synergistically interact in producing more boardings and improve DDM model fit and explanatory power. I frame this hypothesis in a utilitarian theoretical model, expecting that their combined influence would increase overall transit patrons' utility and decrease their generalized travel cost along two potential mechanisms: 1- access to more socio-economic opportunities at both local and metropolitan scales; and/or 2- their combined effect on reducing total multimodal travel time (pedestrian + transit). In other words, greater pedestrian accessibility to a rapid-transit station and within a rapid-transit station's Pedshed (local-accessibility); as well as greater accessibility from a rapid-transit station to socio-economic opportunities along the rapid-transit network (metropolitan-accessibility) would produce greater patronage when compared to the individual influence of each factor alone, as is usually evaluated in precedent station-level DDM ridership models.*

Moreover, supplementing station-level DDMs with nodal (origin-based) metropolitan accessibility measures increases the level of information in the model. It brings into evaluation the potential spatial interaction of an origin-station (i) with others (j) as well as a users' perspective rather than operators' supply-side attributes, local built-environment, and network topological factors.

The inclusion of stations' multi-scalar accessibility measures (*local* + *metropolitan*) is particularly useful in a situation where station-to-station person trip data is not available or non-existent. This ameliorates station-level DDM limitations. As noted by Duncan (2010), origin–destination (OD) pairs allow to better account three main factors linked to transit patronage: trip-production attributes at origin-station; trip attraction attributes at destination-station; and service quality relative to other modal options. Also, for DDM models based on OD station pairs the number of observations greatly increases, which allows for more robust statistical inference.

On the other hand, OD data is often difficult to obtain, whilst station-based DDM models have recently increased in their predictive power and applications (Ramos-Santiago, 2021). Nevertheless, as previously noted by Taylor et al. (2009), multicollinearity appears to be a recurring issue in DDM modeling and careful assessment is needed when introducing new predictors, such as metropolitan- and/or local-accessibility measures.

4. Research design and methods

Given that transit users inevitably experience and potentially value both phenomena (*local*- and *metropolitan*-accessibility) as part of their one-way trip, their independent and/or combined effects may reflect on aggregate travel behavior.

In order to explore this and overcome precedent research limitations, this study relies on *metropolitan* and *local* accessibility indicators (MA_i and LA_i, respectively) that are underpinned by the same spatial interaction construct: an origin-based cumulative opportunities model with distance decay treatment. These accessibility indicators were estimated for a consistent set of observations and for a unique unit of analysis, rapid-transit stations. This provides a more robust research design framework to examine multi-scalar phenomena and their potential interaction in a consistent node; contributes to theory and practice for transit planning; and possibly improves station-level DDM's explanatory power (**Equation (4)**.

$$Boardings^{station} = f(\ SE, LU.BE, TS, NT, MA_i, LA_i, [MA_i x LA_i]\ ) \tag{4}$$

Since a majority of rapid-transit trips in LA begin with walking (Ramos-Santiago, 2021), the proposed interaction term would capture the combined effects of scale- and mode-specific accessibility experiences: one *pedestrian* in mode and *local* in scale; the other *rapid-transit* in mode and *metropolitan* in scale.

This study was framed as a single case study of LA multimodal rapid-transit network, which features heavy rail (HRT), light rail (LRT), and bus rapid transit (BRT). The rapid-transit system is complemented by a large and variegated network of feeder bus lines that operate across multiple jurisdictions within LA's metropolitan statistical area (Ramos-Santiago, 2021). Considered a polycentric agglomeration, some urban scholars argue that LA is better characterized as a *dispersed* mega-city (Gordon and Richardson, 1996). This generates a diverse and expansive landscape of socio-economic opportunities across the metropolitan area, and within rapid-transit station Pedsheds, ideal for exploring the core issues in this paper, at both local and metropolitan scales.

The unit of analysis is rapid-transit stations (n = 100; excluding extreme outliers) and the study relied on multi-level generalized linear regressions (ML-NBGREG) and Exploratory Factor Analysis (EFA). Because variability is expected across the diverse transit technologies in LA's rapid-transit network, and the outcome data is sourced from a line-based person-trip survey, a multi-level modeling framework is applied to the data. The assessment of propositions and test of hypotheses was based on comparison of a set of restricted and un-restricted models where key variables of interest and controls were regressed onto a common ridership outcome, average weekday boardings.

In addition, the outcome variable reflects a highly skewed distribution, typical of count measures such as station boardings. Applied statistics literature (Hilbe, 2011) and more recent transit DDM studies (Ramos-Santiago, 2021; Ramos-Santiago et al., 2022; Ramos-Santiago and Brown, 2016; Aston et al., 2020) point to generalized linear models geared for negative-binomial distributions as a more robust approach to work with this type of data. Hence, a multi-level *random-intercept* model fits a negative-binomial distribution (**Equation (5)**:

$$ln\mu_{il} = (n_{il} + e_{il}) = \gamma_0 + \sum_{h=1}^{r} \gamma_h x_{hil} + R_{il} + U_{0l} \tag{5}$$

Where:

i = indicates level-one unit (e.g., rapid-transit station).

l = indicates level-two unit (e.g., grouping: *Rapid-Transit Line*).

μ_{il} = expected number of average daily boardings at station i of rapid-transit line l.

$(n_{il} + e_{il})$ = allows for random variation of the expected number of boardings (nbreg).

γ_0 = average intercept.

γ_h = coefficient vector.

x_{hil} = explanatory variable (including *local-accessibility* LA_i, *metropolitan-accessibility* MA_i, and their interaction term).

R_{il} = level-one residuals.

U_{0l} = level-two residuals (group effects).

The key variables of interest are *metropolitan*- and *local*-accessibility indicators, hereafter referred to as a station's *metropolitan accessibility* and *local accessibility*; their interaction; and latent exploratory variables *panoptic accessibility* and *conditioned walkability* that were developed and implemented in the last model and further discussed below.

This study also addresses Elldér (2020) and Levinson & Wu (2020) calls for more research exploring and extending concepts of accessibility in travel studies. Specifically, this author developed a final exploratory model in which an extended definition for a rapid-transit station's general accessibility was developed via Exploratory Factor Analysis (EFA). Given the relatively small number of observations in the dataset and high collinearity between several candidate variables EFA became a useful and reliable approach to identify robust latent variables. These capture the underlying correlational structure of a more complete station accessibility phenomena (see (Levinson and Wu, 2020) for a detailed discussion of general accessibility). Finally, test of parameters' non-stationarity was assessed via geographically weighted regressions (GWR) in pursuit of a more comprehensive understanding of underlying forces and more reliable DDM models.

4.1. Metropolitan, Local, and latent construct measures

4.1.1. Metropolitan accessibility measure [MA_i]

The main function of a rapid-transit system is to facilitate metropolitan-scale access. In this study a modified Hansen origin-based cumulative opportunity model was implemented to operationalize the concept of a station's *metropolitan accessibility* (**Equation** (6). The number of jobs within destination-station j pedestrian service areas serves as proxy for potential socio-economic opportunities. Schedule-based travel time between origin-station i and destination-station j (tt_{ij}), including transfers, is based on LA Metro transit schedules as registered in GTFS files for year 2012. Because the response variable is average weekday boardings, and rapid-transit lines service levels vary in a typical weekday, the cumulative opportunities measure was calculated as an average of 1hr periods (n = 20):

$$MA_i = \left(\frac{1}{n}\right) \sum_{h=1}^{n} \sum_{j,j \neq i}^{m} O_j * f\left(tt_{ij}, h\right) \tag{6}$$

Where:

MA_i = weekday *metropolitan-accessibility* at rapid transit origin-station i.

h = weekday rapid-transit schedule 1 h period (1–20) based on GTFS files, yr2011.

i = rapid-transit origin station.

O_j = number of opportunities (jobs) at rapid-transit destination station j.

j = rapid-transit destination-station.

tt_{ij} = weekday travel time (including transfer) between origin-station i and destination-station j.

$f(*)$ = distance-decay function *gamma*: ($tt_{ij}^{-b} * e^{-c*tt_{ij}}$), where b = -0.503, e = base of natural logarithm, and c = −0.078 are parameters for large MPO (pop > 3million, NCHRP Report 716).

Three distance-decay functions were tested (inverse, exponential, and *gamma*). Travel time was modulated with best-fit distance-decay function *gamma* to capture a more realistic travel behavior. Recent travel forecasting and transportation planning literature has noted better performance of this decay function in the trip distribution phase of the 4-step model (Systematics, 2012).

4.1.2. Local accessibility measure [LA_i]

In this study *local accessibility* refers to the ease of access to a rapid-transit station on foot, as well as access to socio-economic opportunities within a station's Pedshed. This dual interpretation is operationalized through the concept of *walkability*. *Local accessibility* takes into consideration multiple factors such as built environment; number and proximity of amenities (opportunities); pedestrian travel behavior; pedestrian infrastructure; and urban design characteristics. As noted by Weinberger and Sweet (2012) *walkability* measures the opportunity to walk, instead of actual walking behavior. Multiple definitions have been posited in the literature that vary based on research contexts and purpose, unit of analysis, or on data availability and methods, see for example Schlossberg and Brown (2004) or Southworth (2005).

Following Weinberger and Sweet (2012), and Cui et al. (2022) this study relies on a proprietary, yet publicly available metric for operationalizing *local accessibility* at rapid-transit stations. WalkScore® has often been used as a proxy for walkability in several studies and outperformed population density, a measure often used to operationalize pedestrian-friendliness, in its ability to predict walk mode share for a variety of purposes (Weinberger and Sweet, 2012). In essence, WalkScore® is a composite indicator of accessibility by foot based on a spatial interaction model for calculating cumulative opportunities to nearby amenities from a specific address (e.g., rapid-transit station). The algorithm weights more the opportunities located within a 1/4mile (~400mts) network-based buffer than those located beyond, and up to a maximum 30 min walking threshold. It also considers aspects of the built environment such as block length and intersection density. For details see description below, and Weinberger and Sweet (2012), Duncan and Aldstadt (2013) and Walk Score (2019):

> "… WalkScore® can be described as measuring the walkability of any address using a patented system. For each address, Walk Score analyzes hundreds of walking routes to nearby amenities. Points are awarded based on the distance to amenities in each category. Amenities within a 5 min walk (0.25 miles) are given maximum points. A decay function is used to give points to more distant amenities, with no points given after a 30 min walk. Walk Score also measures pedestrian friendliness by analyzing population density and road metrics such as block length and intersection density. Data sources include Google, Education.com, Open Street Map, the U.S. Census, Localeze, and places added by the Walk Score user community."

(https://www.walkscore.com/transit-score-methodology.shtml; accessed 11/19/2019).

WalkScore has been used in several build-environment and transportation studies across disciplines. Hirsch et al. (2013) found significant correlations between WalkScore® and higher odds of walking to transport and more time spent in transport-related walking; and Duncan and Aldstadt (2013) confirmed a better performance of WalkScore® when compared to several objective GIS measures of walkability at neighborhood level. In addition, WalkScore® has also been used in real estate economics; land-use; health science; and transit patronage studies (; Cui et al., 2020; Hall and Ram, 2018).

Yet WalkScore® does not consider other environmental and social factors found to influence pedestrian behavior. For instance, crime; existence of sidewalks and/or pedestrian crossings; and aesthetic factors like the presence of trees and/or quality of the streetscape (Weinstein Agrawal et al., 2008; Park et al., 2014; Tilahun and Li 2015; Cao and

Duncan, 2019). For example, Cao & Duncan found that the presence of sidewalks and pedestrian crossings were very important influences in the propensity of 'Park & Riders' to walk to a transit station (Cao and Duncan, 2019). Likewise, based on input from 355 respondents from Chicago Tilahun & Li (2015) identified access time, safety from crime, and sidewalk availability as statistically significant factors in explaining the propensity of transit users to walk to transit (Tilahun and Li, 2015). Nevertheless, Duncan et al. found a statistically significant medium correlation between WalkScore® ranks, sidewalk width, and sidewalk completeness within an 800mts (1/2mile) buffer radius (Duncan et al., 2013). As consequence, it could be argued that WalkScore® captures part of the information related to sidewalk infrastructure.

Of particular value to this paper, WalkScore® combines land-use and design attributes (e.g., street network characteristics, proximity to amenities), pedestrian access from/to a location (e.g., access to a rapid-transit station on foot), as well as access to opportunities within the station's pedshed. It can be considered an origin-based cumulative opportunity measure that is conceptually similar to the stations' *metropolitan accessibility* calculated in this study.

4.1.3. Latent constructs

Latent constructs are variables not directly observed that capture

Table 1
Summary statistics of candidate variables.

Variable	*source (year)*	*transformation(s)*	**Mean**	**Std. Dev.**	**Min**	**Max**	Raw (untransformed) Data			
							Mean	**Std. Dev.**	**Min**	**Max**
Dependent:										
Avg. Weekday Station Boardings	*LA Metro (2011)*	*n.a.*	3405	5545	50	38,665	n.a	n.a	n.a	n.a
Independent: demographic/ socioeconomic										
Population	*US Census (2011)*	*scaled (000's), centered*	0.00	10.26	−17.77	37.43	17,957	10,256	183	55,385
Jobs	*US Census LEHD (2011)*	log, *centered*	0.00	1.10	−2.39	2.94	15,743	26,298	742	154,439
Avg. Number of Vehicles / HU	*US Census (2011)*	*n.a.*	1.26	0.43	0.29	2.06	n.a	n.a	n.a	n.a
accessibility										
Metropolitan-Accessibility	*author*	*scaled (0–100), centered*	−0.92	21.97	−16.03	83.97	65,116	89,735	3398	411,896
Local-Accessibility	*WalkScore® (2015)*	*[WalkScore; 0–100], centered*	0.93	18.02	−44.48	21.52	78.41	18.02	33.00	99.00
Panoptic Accessibility	*author*	*n.a.*	9.69E-10	0.98	−0.86	2.95				
network / topological										
One-Way Service	*LA Metro (2011)*	*[0,1] dummy*	0.22	0.42	0.00	1.00	n.a	n.a	n.a	n.a
Terminal	*LA Metro (2011)*	*[0,1] dummy*	0.08	0.27	0.00	1.00	n.a	n.a	n.a	n.a
Transfer Hub	*LA Metro (2011)*	*[0,1] dummy*	0.07	0.26	0.00	1.00	n.a	n.a	n.a	n.a
Nodal Split (fork)	*LA Metro (2011)*	*[0,1] dummy*	0.01	0.09	0.00	1.00	n.a	n.a	n.a	n.a
Union Station	*LA Metro (2011)*	*[0,1] dummy*	0.01	0.10	0.00	1.00	n.a	n.a	n.a	n.a
service level										
Rapid-Transit Vehicles per Weekday	*LA Metro (2011)*	*n.a.*	227	109	80	667	n.a	n.a	n.a	n.a
Number of Parking Spaces	*LA Metro (2011)*	*scaled (00's), centered*	0.00	4.20	−2.46	21.94	246	420	0	2440
bus network connectivity										
Count of Bus Vehicle Trips	*LA Metro (2011) / SCAG (2011)*	*n.a.*	1509.39	1841.83	44.00	7506.00	n.a	n.a	n.a	n.a
Number of Bus Lines	*LA Metro (2011) / SCAG (2011)*	*n.a.*	16.79	18.61	1.00	70.00	n.a	n.a	n.a	n.a
Number of Bus Local Lines	*LA Metro (2011) / SCAG (2011)*	*n.a.*	10.61	10.73	1.00	48.00	n.a	n.a	n.a	n.a
Number of Bus Rapid Lines	*LA Metro (2011) / SCAG (2011)*	*n.a.*	1.48	2.14	0.00	8.00	n.a	n.a	n.a	n.a
Number of Bus Commute Lines	*LA Metro (2011) / SCAG (2011)*	*n.a.*	1.75	3.02	0.00	9.00	n.a	n.a	n.a	n.a
build-environment										
Land-Use Mix Score	*WalkScore® (2015)*	*centered*	−0.01	1.47	−2.50	5.69	78.83	17.65	31.12	99.89
Avg. Block Length Score	*WalkScore® (2015)*	*centered*	1.13	26.27	−49.61	43.39	156.03	47.60	107.00	439.00
Intersection Density Score	*WalkScore® (2015)*	*centered*	0.57	13.42	–33.26	26.41	107.78	39.92	6.00	185.50
land-use										
Special Generators	*SCAG Land Use Shapefile (2011), Google Maps, OSM Data*	*n.a.*	4.32	4.42	0.00	32.00	n.a	n.a	n.a	n.a

underlying relationships between a larger set of observed, and possibly correlated variables. Exploratory Factor Analysis (EFA) allows to uncover this underlying structure and is often used in multiple disciplines to develop more parsimonious models, manage multicollinearity issues, and/or uncover relationships for theoretical and/or practical modeling purposes. As noted by Fabrigar et al. (1999) and Watkins (2018), EFA allows for construction of more parsimonious models via latent variables whilst maintaining relatively high levels of information from measured correlated variables.

Six observed variables that account for a variety of access modes that overlap at origin-stations informed EFA analysis. Specifically, pedestrian; bus; rapid transit; and automobile availability (Table 2). This set of modal dimensions better captures rapid-transit stations' general accessibility. This is particularly relevant when studying polycentric agglomerations that rely to a larger extent on feeder bus networks for a large share of their patronage (>25%; Ramos-Santiago, 2021). Also, Levinson and Wu's (2020) proposition that a more general measure of accessibility is better than a less general measure further motivates this exploration. EFA analysis and constructs are informed in part by the previously calculated *metropolitan-accessibility* and *local-accessibility* indicators. Latent constructs derived from EFA inform Model.4ab specification, which in turn complements analyses, findings, and results from the preceding ML-NBREG models.

4.2. Candidate variables and models

Table 1 registers a summary of statistics of the dependent and candidate independent variables including those used in EFA. The set of independent variables is drawn from the literature review of transit ridership factors; precedent DDM studies; and this author's hypotheses. Rapid transit station-level boarding data is drawn from Los Angeles' LA Metro Bus and Rail On-board Survey (LACMTA 2012a,b,c). Other sources of data are LA Metro stations' WalkScore® ranks and their subcomponents; US Census ACS and LEHD databases for population and employment levels; and LA's multi-agency GTFS files (LA Metro and SCAG agencies; Southern California Association of Governments). These latter files were used for development of the *metropolitan accessibility* travel time matrices ($tt_{ij,h}$), which were developed with Transit Network tools in ArcGIS Pro (v.2.X).

Socioeconomic values for stations' pedestrian service areas (SE; Equation (4), Table 1) were captured in GIS using a 1/2mile (~800mts) network-based distance parameter with ArcGIS 'Network Analyst' service area tool. Underlying georeferenced data was sourced from US Census Bureau survey (Table 1).

Linear mixed models (a.k.a. Multilevel, Hierarchical) provide the basis for analyses and comparison in this study. This technique addresses the situation where observations are not independent (e.g., stations cluster along lines with distinct rapid-transit modes and service characteristics), and correctly model correlated error (Garson, 2013). Specifically, a *random-intercept* model is specified for this situation and reflects as variance in the random effects model output. Similar to conventional regression analyses, the model predictor variables on the right-hand side of the estimation equation are linearly related to the outcome variable on the left-hand side. Models that treat the relationship between predictors and outcome in a non-linear fashion are classified as *generalized* linear models (GLM) and a 'link' function is specified. The link function specification is based on outcome distribution characteristics, among other factors (Hilbe, 2011; Cameron and Trivedi, 2010).

As per applied statistics best practice for count models (Ramos-Santiago, 2021; Ramos-Santiago et al., 2022; Hilbe, 2011; Cameron and Trivedi, 2010), this study uses a multi-level generalized linear model with a negative binomial link function (ML-NBREG). The model output reports estimated parameters as incidence rate ratios (IRR), which can be interpreted as semi-elasticities. Other reported fit statistics, and parameter and model attributes, such a *p*-value, directionality (sign), likelihood-ratio test, and pseudo-R^2 are similarly interpreted as in conventional regression analysis.

4.3. Analytical strategy

The overall analytical strategy is based on comparison of sequential model fit statistics; explanatory power; and key variables' significance, magnitude, and directionality. It begins by contrasting a restricted base model of fixed-effects only (**Model.00**) that excludes accessibility measures, to a restricted model with full-effects (fixed + random; **Model.01**). This is done to assess the appropriateness of a multi-level approach.

Post-regression statistics of model-fit, such as likelihood-ratio test, AIC and BIC penalized-likelihood statistics, pseudo-R^2 for fixed- and full-effects, as well as visual inspection of predicted vs. observed boardings and residuals were used in confirming model's adequacy and improvements. Also, violation of regression assumptions related to multicollinearity, such as high Variance Inflation Factor (VIF) and/or nonsensical covariates sign and/or unexpected parameter magnitude changes were also evaluated.

The restriction in the first two models (**Model.00** and **Model.01**) is based on excluding stations' *local-* and *metropolitan-accessibility* indicators, which are then specified independently and in tandem in subsequent unrestricted models **Model.02a**, **Model.02b**, and **Model.02ab.** This is done to test the first hypothesis in this study, whether a station's *local-* and *metropolitan accessibility* are significant and positively related to boardings in LAs DDM rapid-transit ridership model.

The subsequent **Model.03** includes an interaction term to test the second hypothesis of this study (that a synergistic effect exists between *local-* and *metropolitan accessibility*). The interaction was then assessed for significance; directionality; evidence of model improvement in explanatory power; and interpretation of main and interaction effects.

These results were then contrasted with those from the preceding **Model.02ab** where no interaction is specified. Continuous variables were centered to better handle structural collinearity when specifying interaction factors in **Model.03**, and some variables were scaled and/or transformed to address residuals non-linearity and to facilitate interpretation of estimated parameters (see Table 1). An exploratory **Model.04** was also developed and fitted with latent constructs that were estimated via Exploratory Factor Analysis (EFA). In addition to the *local-* and *metropolitan-accessibility* indicators, variables representing bus-rail connectivity (classified across service levels [local, rapid, express]), and a variable capturing automobile availability per household were processed and tested for reliability and feasibility for use in EFA (Table 2**;** Table 3**, Model.04**).

Resulting latent variables scores were then calculated for two relevant factors and results from this model were then assessed and compared to the analogous **Model.02ab** results. Finally, the possibility of spatial non-stationarity in estimated parameters (see Cardozo et al., 2012) is explored for LAs system in two Geographically Weighted Regressions: **GWR.02ab** and **GWR.04ab**. The first relies on the estimated *local-* and *metropolitan-accessibility* indicators; and the second relies on the latent constructs from the EFA analysis. The intent is to verify which parameters, if any, is/are more spatially stable and reliable, and perhaps gain deeper insights into the phenomena of interest, theoretical and practical implications.

5. Results and discussion

5.1. Metropolitan and local accessibility measures

Three distance decay functions were calculated for the *metropolitan-accessibility* measure and tested for best model fit. The empirically derived *Gamma* function performed best, registering a highly significant parameter; better model fit (lower AIC and BIC); and was used in the final models. Fig. 1 registers variability in boarding stations'

Table 2
Exploratory factor analysis: factor-1 and factor-2 eigenvalues, loadings, and number-of-factors fit statistics.

	[a] Variance	Difference	Proportion
Factor 1	3.04709	1.30532	0.6363
Factor 2	1.74177	.	0.3637
LR test: independent vs. saturated: chi2(15) = 514.07 Prob > chi2 = 0.0000			
LR test: 2 factors vs. saturated: chi2(4) = 3.61 Prob > chi2 = 0.4614			

	[b] Loadings		
	Panoptic Accessibility	*Conditioned Walkability*	
Variables	Factor 1	Factor 2	Uniqueness
Num. Rapid Bus Lines	**0.8747**	0.3162	0.1350
Num. Commuter Bus Lines	**0.8587**	0.2066	0.2199
Metropolitan-Accessibility	**0.8402**	0.3942	0.1387
Num. Local Bus Lines	**0.7565**	0.3646	0.2947
Local-Accessibility	0.2131	**0.8338**	0.2595
Num. Vehicles / Household	−0.4701	**−0.7846**	0.1633

Notes: a. maximum-likelihood method.
b. orthogonal rotation.

metropolitan accessibility (top) and *local accessibility* (bottom). Higher values of *metropolitan accessibility* tend to cluster at and around LA's central business district (CBD) where historically high levels of employment and more dense transit network locates. On the other hand, *local accessibility* levels show a more diverse and dispersed geographical distribution, with clusters at the CBD; along the HRT red/purple lines North-West from the CBD; towards the South in the Long Beach area; and North-East from the CBD in Pasadena. This pattern in part reflects the historical development of towns and cities that coalesced in time and now form part of the larger LA metropolitan area.

5.2. Multilevel and Negative-Binomial generalized regression modeling approach

A comparison of post-regression statistics (Table 3) and visual assessment between **Model.00** (fixed-effects only) and **Model.01** (mixed-effects) indicates that a multi-level approach better fits the data and confirms variability at line-level (Fig. 2). This approach and generalized linear regressions for data fit with a negative-binomial distribution link (NBREG) is also suggested as best-practice for transit modeling and research (Aston et al., 2020).

5.3. Overall model results and hypotheses testing

In this study all ML-NBREG models reached convergence and all likelihood-ratio tests of alpha = 0, comparing to a Poisson model, suggest that alpha is non-zero (over-dispersion; $p \leq 0.05$). This supports a negative binomial modeling approach as more appropriate to fit the data (Table 3). Also, all estimated parameters yield expected directionality in alignment with precedent studies and theory; and stations' topological controls effectively managed outliers and produce better models fit. They likely capture specific advantages and/or disadvantages related to nodal functions and a station's relative position in the rapid-transit network (Fig. 2).

Proposition 1: *That Both Local- and Metropolitan-Accessibility are Significant Predictors of Ridership at Station-Level.*

The introduction of the *local accessibility* measure in **Model.02a** yields a positive yet non-significant parameter (Table 3). The non-significance could be result of the relative low number of observations in the dataset, and/or its partial correlation with the variable *Jobs*. Although the VIF statistic is non-problematic (VIF < 3) removing *Jobs* from the model results in a significant parameter for *local accessibility* at a 99% confidence level. However, this is not the first study that reports non-significant results for a measure of local land-use and build-environment characteristics (e.g., TOD) on transit patronage. For example, Brown and Thompson (2012) found that TOD characteristics were not relevant for bus and most rail riders in Atlanta's multi-modal system. Another potential explanation of the non-significant result is that a potential cross-over interaction effect exists where *local accessibility* could act as mediator of another factor. This possibility is explored in **Model.03** and discussed later in this section.

The introduction of the *metropolitan accessibility* indicator in **Model.02b** (independently) and in **Model.02ab** (in tandem with *local accessibility*) results in a positive and significant parameter at a high 99% confidence level; higher pseudo-R^2 for full effects; and improved model fit (e.g., lower AIC and BIC statistics and highly significant likelihood-ratio test) when compared to the preceding **Model.02a** (Table 3). This confirms part of the first proposition that states *metropolitan accessibility* and *local-accessibility* are significant trip-production factors in LA's rapid-transit network.

Due to partial correlation of *Jobs* with both *local- and metropolitan-accessibility* the variable *Jobs* was removed from the subsequent interaction **Model.03** and from **Model.04**, with no significant penalty on models fit nor explanatory power. This is likely result of spatial correlation, where areas closer together tend to share similar values across a variety of attributes. That is, stations with high number of jobs tend to locate close to each other at or near historical urban centers (e.g., CBD's, downtowns) and/or in more recent metropolitan sub-centers. These areas also tend to register higher density of transit service. Hence, higher levels of both *local-* and *metropolitan-accessibility* are usually found together in stations with higher number of jobs. This possibility was further assessed with Geographically Weighted (GWR) regressions and results are discussed further below.

The effect of *metropolitan accessibility* is similar in effect size as compared to *population*, but lower than that of *Jobs*. This likely reflects the predominance of commute travel in LA's rapid-transit system (Ramos-Santiago, 2021) contrasting with previous DDM studies where *metropolitan accessibility* played a stronger role in transit ridership, especially for peak-time ridership (see Chen & Zegras Boston's study; 2016). The difference in effect size may result from this study's more aggregate outcome data (average weekday boardings) that conflates peak and non-peak boardings.

Proposition 2: *Local-Accessibility and Metropolitan-Accessibility Synergistically Interact in Producing More Boardings.*

Model.03 specifies the interaction term between *local accessibility* and *metropolitan accessibility* and yields a highly significant and positive interaction parameter ($p < 0.000$), as well as significant main effects (Table 3), with a relatively small effect size. The overall model fit notably improved with a significant likelihood-ratio test (compared to the less saturated preceding model) and reported the lowest AIC and BIC scores of the entire model set. This implies that the interaction model allows for better data fit than the preceding models.

Notable improvement in explanatory power as registered in the highest pseudo-R^2 statistics for both fixed- and full-effects is also reported (fixed effects = 0.76; full effects (fixed + random) = 0.91) (Table 3). Directionality of main effects indicate a cross-over interaction where *local-accessibility* appears to act as mediator. These results allow to reject the null hypothesis that there is no significant interaction effect between *local-* and *metropolitan-accessibility* with a high-confidence level.

Table 3
ML-NBREG Fit Statistics, Regression Model Results, and Exploratory Factor Analysis Statistics.

	interaction term (i)	**MODEL.00** Restricted FIXED-EFFECTS ONLY			**MODEL.01** - Restricted - MIXED-EFFECTS			**MODEL.02a** - Un-restricted - MIXED-EFFECTS inc./ Local-Accessibility (Walkcore®)			**MODEL.02b** Un-restricted MIXED-EFECTS inc/ Metropolitan-Accessibility			**MODEL.02ab** Un-restricted MIXED-EFECTS inc/ Local-Accessibility and Metropolitan-Accessibility			**MODEL.03** Un-restricted MIXED-EFECTS Interaction Term: Local-Accessibility (WalkScore®) X Metro-Accessibility			**MODEL.04ab** Un-restricted MIXED-EFECTS Latent Factors: *Panoptic Accessibilty and Conditioned Walkability*		
Model-fit statistics:																						
N:		100			100			100			100			100			100			100		
LR test vs. nbinomial model:		chi2(01) = 1.2e + 05, Prob >= chi2 = 0.000			chi2(2) = 95.31, Prob > chi2 = 0.0000			chi2(2) = 95.96, Prob > chi2 = 0.0000			chi2(2) = 114.12, Prob > chi2 = 0.0000			chi2(2) = 116.30, Prob > chi2 = 0.0000			chi2(2) = 122.20, Prob > chi2 = 0.0000			chi2(2) = 124.65, Prob > chi2 = 0.0000		
Likelihood-ratio test:		n.a.			[m.00 nested in m.01]: LR chi2(2) = 95.31 Prob > chi2 = 0.0000			[m01 nested in m02a] LR chi2(1) = 1.37 Prob > chi2 = 0.2425			[m01 nested in m02b] LR chi2(1) = 21.00 Prob > chi2 = 0.0000			[m01 nested in m02ab] LR chi2(1) = 19.78 Prob > chi2 = 0.0000			[m02_1b nested in m03] LR chi2(1) = 19.44 Prob > chi2 = 0.0000			n.a.		
AIC:		1733.742			1642.435			1643.070			1623.44			1624.657			1605.998			1613.281		
BIC:		1759.793			1673.697			1676.940			1657.31			1658.524			1642.470			1647.148		
[a] Fixed-Effects only Pseudo-R^2		0.67			0.67			0.67			0.68			0.50			0.71			0.67		
[a] Total-Effects Pseudo-R^2		n.a.			0.87			0.87			0.90			0.89			0.91			0.91		
Model (OLS) VIF		1.31			1.31			1.58			1.68			1.81			[c] 20.53			1.59		
Highest Ind. Var. VIF		1.69			1.69			2.27			2.82			2.55			[c] 6.13			2.34		
Factor Analysis Statistics:																				[e] Exploratory Factor Analysis Statistics:		
Det. Correlation matrix			n.a.			n.a.			n.a.			n.a.			n.a.			n.a.		Det. Correlation matrix	0.005	
Bartlett sphericity			n.a.			n.a.			n.a.			n.a.			n.a.			n.a.		Bartlett sphericity	0.000	*p-value*
KMO			n.a.			n.a.			n.a.			n.a.			n.a.			n.a.		*KMO*	0.86	
Cronbach's alpha			n.a.			n.a.			n.a.			n.a.			n.a.			n.a.		Cronbach's alpha	0.705	
DV: Avg. Weekday Boardings		IRR	*p*	sig.	IRR	*p*	sig.	IRR	*p*	sig.	IRR	*p*	sig.	IRR	*p*	sig.	IRR	*p*	sig.	IRR	*p*	sig.
Fixed-effects:																						
Population		1.039	*0.000*	***	1.022	*0.000*	***	1.018	*0.002*	***	1.021	*0.000*	***	1.018	*0.000*	***	1.016	*0.001*	***	1.021	*0.000*	***
Jobs		1.145	*0.032*	**	1.203	*0.000*	***	1.171	*0.003*	***	1.059	*0.249*		[b] *not used*	[b] *not used*	[b] *not used*	[b] *not used*	[b] *not used*	[b] *not used*	*not used*	*not used*	*not used*
Number of Parking Spaces		1.145	*0.000*	***	1.094	*0.000*	***	1.096	*0.000*	***	1.122	*0.000*	***	1.124	*0.000*	***	1.118	*0.000*	***	1.119	*0.000*	***
[d] (*dummy*) OneWay Service		0.182	*0.000*	***	0.433	*0.000*	***	0.390	*0.000*	***	0.364	*0.000*	***	0.346	*0.000*	***	0.233	*0.000*	***	0.307	*0.000*	***
[d] (*dummy*) Terminal		4.899	*0.000*	***	2.810	*0.000*	***	3.105	*0.000*	***	3.356	*0.000*	***	3.548	*0.000*	***	4.835	*0.000*	***	3.884	*0.000*	***
[d] (*dummy*) Transfer Hub		3.088	*0.000*	**	3.756	*0.000*	***	3.848	*0.000*	***	2.645	*0.000*	***	2.616	*0.000*	***	2.670	*0.000*	***	3.191	*0.000*	***
[d] (*dummy*) Split Node		1.941	*0.377*		1.601	*0.291*		1.707	*0.231*		1.426	*0.382*		1.495	*0.326*		1.397	*0.364*		1.612	*0.211*	
[d] (*dummy*) Union Station		5.539	*0.033*	***	1.670	*0.319*		1.805	*0.253*		1.211	*0.684*		1.187	*0.720*		10.372	*0.000*	***	0.802	*0.628*	
Local-Accessibility		*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	**1.005**	***0.239***		*not used*	*not used*	*not used*	**1.001**	***0.744***		**1.035**	***0.000***	***	*not used*	*not used*	*not used*

(*continued on next page*)

Table 3 (*continued*)

	interaction term (i)	**MODEL.00** Restricted FIXED-EFFECTS ONLY			**MODEL.01** - Restricted - MIXED-EFFECTS			**MODEL.02a** - Un-restricted - MIXED-EFFECTS inc./ Local-Accessibility (Walkcore®)			**MODEL.02b** Un-restricted MIXED-EFECTS inc/ Metropolitan-Accessibility			**MODEL.02ab** Un-restricted MIXED-EFECTS inc/ Local-Accessibility and Metropolitan-Accessibility			**MODEL.03** Un-restricted MIXED-EFECTS Interaction Term: Local-Accessibility (WalkScore®) X Metro-Accessibility			**MODEL.04ab** Un-restricted MIXED-EFECTS Latent Factors: *Panoptic Accessibilty and Conditioned Walkability*		
Metropolitan-Accessibility		*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	**1.018**	***0.000***	***	**1.020**	***0.000***	***	**0.979**	***0.016***	**	*not used*	*not used*	*not used*
Local-Access × Metro-Access	*i*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	**1.002**	***0.000***	***	*not used*	*not used*	*not used*
[e] *Panoptic Accessibility*		*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	[e] **1.614739**	***0.000***	***
[e] *Conditioned Walkability*		*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	*not used*	[e] **1.158472**	***0.013***	**
_cons		*2514.364*			*2834.650*			*2858.237*			*3125.644*			*3168.29*			*1908.922*			*2944.702*		
/lnalpha		−0.693	*0.132*		−1.760	*0.147*		−1.772	*0.147*		−1.957	*0.146*		−1.950	*0.147*		−2.160	*0.147*		−2.066	*0.147*	
Random-effects:																						
Rapid-Transit Line					var.	*std. err*		var.	*std. err*		var.	*std. err*		var.	*std. err*		var.	*std. err*		var.	*std. err*	
var (BRT_Silver)		n.a.	n.a.		3.777	*5.581*		3.772	*5.569*		5.433	*7.864*		5.452	*7.903*		4.324	*6.303*		5.657	*8.193*	
var (_cons)		n.a.	n.a.		0.133	*0.096*		0.130	*0.092*		0.098	*0.064*		0.106	*0.069*		0.103	*0.061*		0.106	*0.068*	

Notes:

[a] Pseudo-R^2 calculations are based on the log-likelihood method; see Kramer (2005): $R^2_{LR} = 1$ - exp(- 2/n * (logL_M - logL_0)).

[b] The variable *Jobs* is highly correlated with *local-* and *metropolitan-accessibility* and registers a high VIF statistic, thus removed from specification to reduce multicollinearity effects.

[c] High VIF statistics (multicollinearity) is expected from interaction terms and their factors, as the first is product of the second. Yet this does not influence predictions, precision of the predictions, and the goodness-of-fit statistics (Neter et al. 1996).

[d] Estimated *dummy* [0,1] variable parameters are interpreted as multiplicative factors.

[e] Latent variables from Exploratory Factor Anaysis (EFA - maximum likelihood method (ML); 2 factor structure, orthogonal rotation).

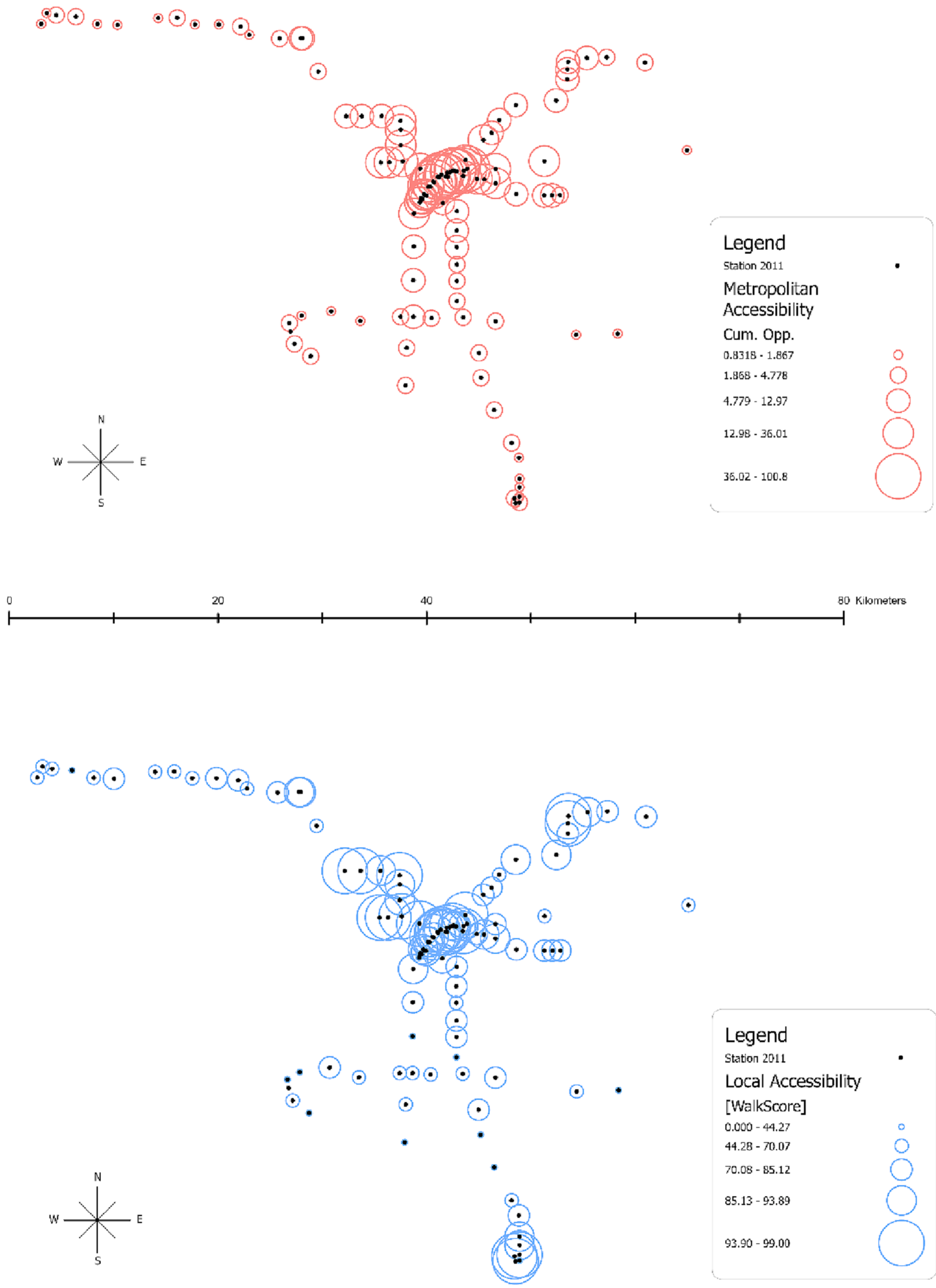

Fig. 1. Rapid-Transit Station Metropolitan Accessibility Levels (top) and Local Accessibility Levels (bottom) in Los Angeles – 2012.

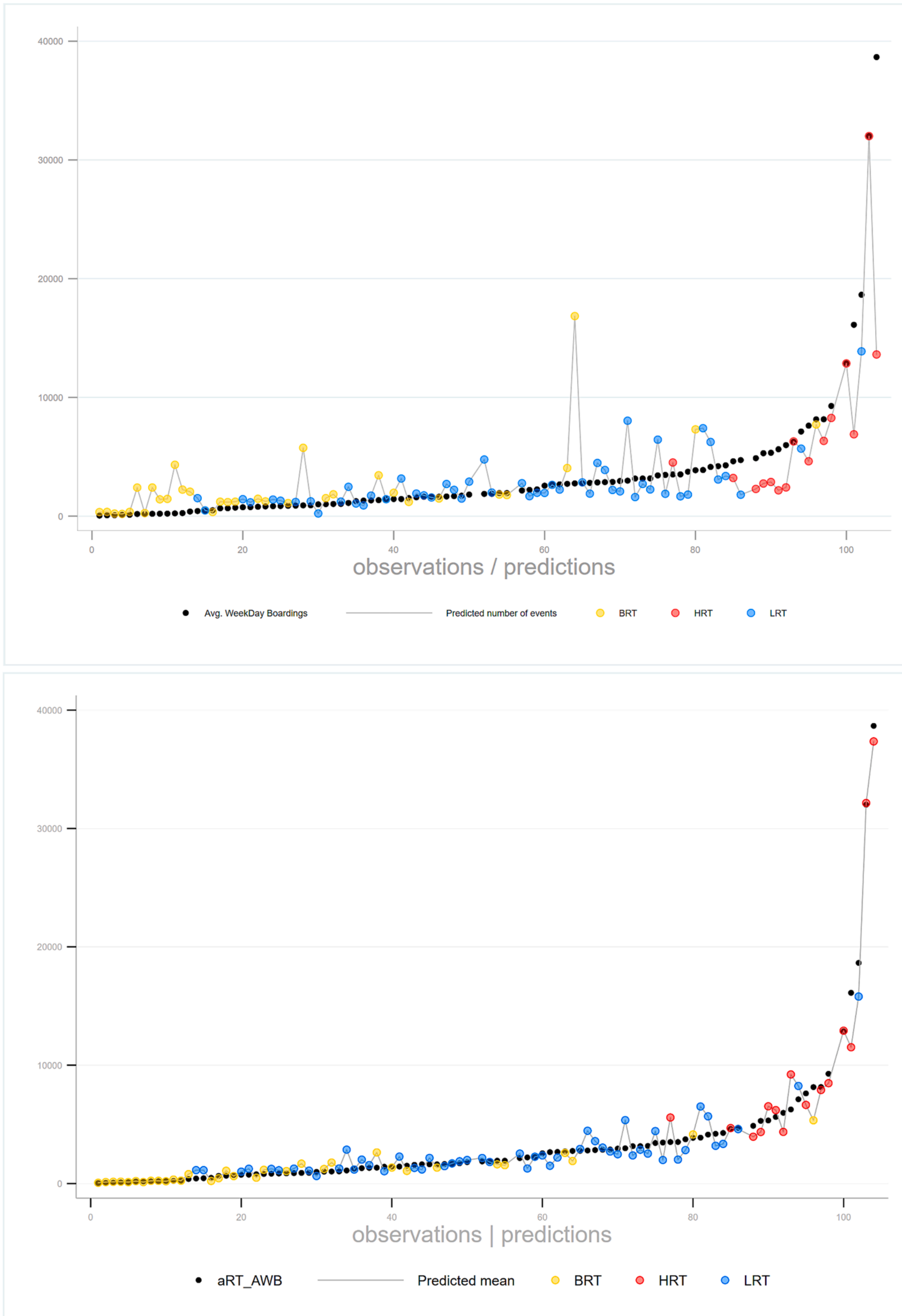

Fig. 2. Predicted vs. Observed Plots for Restricted Model.00 and Model.04. (Above)Model.00 Restricted Model, Fixed-Effects Only. (Below) Model.04 Unrestricted w/ Latent Variables 'Panoptic Accessibility' and 'Conditioned Walkability'.

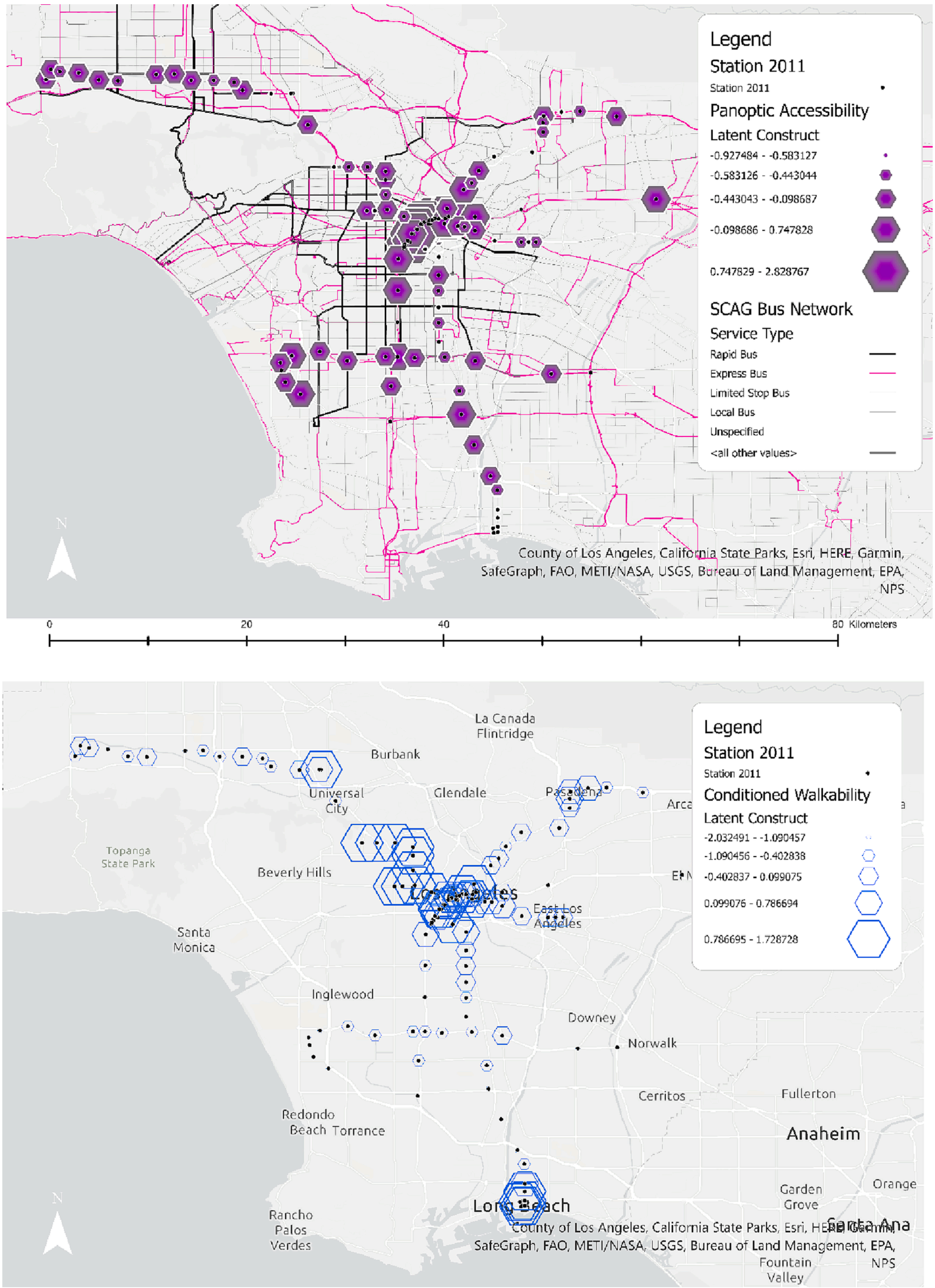

Fig. 3. LA Metro Rapid-Transit Station *Panoptic Accessibility* levels (top), SCAG Bus Network by Service Type (yr2011), and *Conditioned Walkability* levels (bottom).

Table 4
Geographically weighted regression results.

	GWR.02ab		GWR.04ab	
[a] *Global OLS Model-fit statistics:*				
N:	100		100	
df	9		9	
Prob > F	0.0000		0.0000	
R^2	0.8396		0.8497	
Adjusted R^2	0.8216		0.8328	
Global OLS Vs. *Local GWR Model Test:*				
Observed	3.6e + 04		1.10E + 04	
p-value	0.234		**0.000**	
Local GWR Model Results				
Significance Tests for Non-Stationarity:	[b] ***Si***	***p-value***	[b] ***Si***	***p-value***
Constant	0.0434	*0.065*	0.5952	0.000
Local-Accessibility	0.0030	***0.024***	*not used*	*not used*
Metropolitan-Accessibility	0.0004	*0.973*	*not used*	*not used*
Panoptic Accessibility	*not used*	*not used*	0.5693	***0.025***
Conditioned Walkability	*not used*	*not used*	0.2862	***0.018***
Population	0.0024	*0.130*	0.0179	*0.132*
Number of Parking Spaces	0.0001	***0.044***	0.0003	*0.506*
(*dummy*) OneWay Service	0.0387	*0.730*	0.4631	*0.352*
(*dummy*) Terminal	0.0272	*0.931*	0.3921	*0.726*
(*dummy*) Transfer Hub	0.0159	*0.979*	0.9676	***0.034***
(*dummy*) Split Node	0.0143	*0.926*	0.7650	***0.001***
(*dummy*) Union Station	0.1087	*0.179*	0.4692	*0.697*
(*dummy*) BRT-Silver Line	0.0300	*0.607*	1.0200	***0.049***

a. GWR models were regressed onto log-transformed boardings.
b. Standard deviations of the observed parameter estimates.

5.4. Exploratory Model.04: Panoptic accessibility and Conditioned walkability

EFA yielded two strong factors from a 6-item model for inclusion in the exploratory **Model.04**. A 2-factor structure was found more appropriate after conducting 3-, 2-, and 1-factor EFA. This structure was also supported by assessing internal consistency (Cronbach's *alpha*); model fit results from saturated versus independent model tests (ML method; AIC and BIC); and reliability tests.

Factor-1 and Factor-2 explain 63% and 36% of variance, respectively (Table 2). Factor-1 is informed for the most part by a set of regional multimodal accessibility measures, including the *metropolitan-accessibility* indicator (rapid-transit) and three *bus connectivity* measures. Factor-2 combines measured local factors: *local accessibility* and the negative influence of *automobile ownership* on transit ridership (Table 2). Thus, factor-2 operates as a composite measure that balances the positive local influence of a walkable station with the negative local influence of high vehicle ownership, which usually operates as a disincentive of transit patronage in the United States. This author interprets this set of factors as a station's *panoptic accessibility* (factor-1) and a station's *conditioned walkability* (factor-2; Fig. 3).

Factor scores were then calculated for each of the two constructs in Stata/IC 15.1 and values were loaded in models for each latent variable. The resulting regression parameters for *panoptic accessibility* and *conditioned walkability* are highly significant at a 99% and 95% confidence levels, respectively. The two factors present the largest effect sizes of all continuous predictor variables in **Model.04** and for the entire model set (Table 3). Model fit indicators for **Model.04**, such as AIC, BIC, Pseudo-R^2, and VIF statistics support a better fit with specification of latent variables as compared to the analogous **Model.02ab** that relies on the relatively more basic origin-based cumulative measures (Fig. 1,3). In addition, Pseudo-R^2 for full effects is identical to that of the best fit **Model.03** (0.91) allowing for a more parsimonious specification that accounts for a more diverse set of influences and for automobile competition effects (note the negative loading for *vehicle per household* in Table 2). It is also worth noting the greater influence of regional factors when compared to local, especially bus-rail connectivity that has been highlighted in previous transit studies (Ramos-Santiago, 2021). Furthermore, data on all components for the 6-item construct via EFA is readily and publicly accessible in most transit agency and census web portals. For EFA statistics and tests of consistency and reliability please refer to Model.04 results in Table 3 (top).

5.5. Tests of Global Vs. Local GWR models and parameters spatial Non-Stationarity

A generalized GWR modeling approach for fitting a negative-binomial distribution did not reach convergence. Instead, a traditional OLS regression approach onto log-transformed station boardings yielded robust GWR models with high explanatory power (Table 4). Results from model **GWR.02ab**, which specifies *local-* and *metropolitan-accessibility*, did not report significant differences between the 'global' model and the 'local' GWR model (p-value = 0.234); and registered significant spatial heterogeneity in two of the predictors.

Model **GWR.04ab**, which specifies latent variables *panoptic accessibility* and *conditioned walkability*, reports that a 'local' geographically weighted regression approach provides a better fit than a typical 'global' model that assumes parameter spatial stationarity (p-value = 0.000). The spatial non-stationarity for *panoptic accessibility* and *conditioned walkability* is significant at $p = 0.025$ and $p = 0.018$, respectively (Table 4**)**. As to why these two key variables exhibit non-stationarity, this author suggests that they are more sensible to local socio-economic and network connectivity conditions which typically vary across urban landscapes and tend to exhibit clustering patterns (e.g., spatial correlation). These local conditions are not captured by the simpler origin-based cumulative opportunities measure of *metropolitan-accessibility*.

These results suggest that a more general construct of accessibility as derived by EFA offers more nuanced and location-specific parameters (spatially heterogeneous). On the other hand, the origin-based cumulative opportunity measures of *local-* and *metropolitan-accessibility*, although spatially homogeneous and less sensitive to local context are easier to compute. They also yield a significant interaction effect.

Both types of accessibility indicators and modeling techniques could be used by analysts and transit planners. Note that both the interaction **Model.03** and the latent variables **Model.04ab** (Table 3) offer identical explanatory power for full effects (0.91), and the analyst/modeler could use one approach or the other in DDMs or GWRs depending on the focus and purpose of the study, whether for predictive or analytical purposes. Data availability and characteristics, time resources, and statistical software capabilities and availability may impact which model to use.

It is important to note that the more parsimonious **Model.04ab**, which relies on latent constructs of accessibility, requires more data and processing time for EFA development. Also, the GWR models take notably more time to resolve in a standard desktop setting (>4hr) as compared to the generalized ML-NBREG linear regression used for **Model.03** (<0.25hr). Thus, there is a trade-off between modeling speed and the more locally nuanced GWR models that specify latent constructs of accessibility.

Practical applications of station-level DDMs (whether ML-NBREG or GWR) that incorporate multiscale accessibility indicators could facilitate, for example, TOD scenario planning for new or existing rapid-transit stations. This could be applied in single- or multi-line rapid-transit systems. Land use planners, transit planners, and/or developers (and their design staff) could specify different levels of population, employment, transit service, bus feeder connectivity levels, among other topological and contextual characteristics for both ridership and real-estate feasibility analyses. As such, it could support interdisciplinary explorations and discussions, and inform negotiations among key TOD stakeholders. Station-level DDMs could also be used in alternative route analyses and Benefit-Costs analyses. Candidate locations for stations, and hypothetical WalkScores together with hypothetical service levels and land-use characteristics could be specified in equations fitted with estimated parameters and appropriate link function.

Data required for building accessibility indicators and for fitting DDM models is readily and publicly available via US Census, WalkScore and transit agencies GTFS files. Also, the required GIS and statistics software packages are relatively accessible and affordable (at least in institutional contexts) and would only require staff training to intermediate level proficiency. This is an advantage as compared to the typical 4-step Urban Transportation Model that relies on larger and more expensive survey-based databases, and on more costly licenses associated with current industry-level transport/land-use modeling applications.

6. Conclusions, study limitations, and future lines of research

Quantitative analyses of LA's transit system at station-level indicate that independently a station *metropolitan accessibility* is a significant predictor at a 99% confidence level with a relatively moderate effect, and notably improves DDM model fit and explanatory power when interacted with a stations' *local accessibility*. The interaction between *metropolitan accessibility* and *local accessibility* at station-level is also highly significant and positive with a 99% confidence level; produces an increment in explanatory power for fixed-effects of 21%; and results in a model with high explanatory power for full-effects. Therefore, the null hypothesis in this study (that there is no synergistic interaction between *local-* and *metropolitan-accessibility*) is rejected, and modelers, land-use and transit planners would likely benefit from integrating multi-scalar accessibility measures in direct-demand models, and their interactions.

Furthermore, origin-based cumulative opportunity measures proved to be effective operationalizations of accessibility and can be feasibly calculated with publicly available transit and land-use data. This study also illustrates how Exploratory Factor Analysis (EFA) can be an effective tool to handle multicollinearity and develop more general constructs of accessibility for local and metropolitan scales. The use of latent constructs of accessibility can result in a more parsimonious model with higher information and with similar predictive power as the interaction model. EFA can also help develop DDM models where a relatively low number of observations and high collinearity among candidate predictors exist.

In conjunction, the first part of this study that is hypothesis-driven, and the second part that is exploratory, illustrate the utility and desirability of integrating an accessibility lens in transit DDMs and in transit research in general. This allows for a better understanding of multimodal and multi-scalar interactions in complex networks. This is valuable for multimodal transit systems that operate in large polycentric agglomerations. Origin-based cumulative opportunities models with distance-decay functions can capture complex land-use/transportation influences on aggregate travel behavior; and could provide theoretical insights for future research efforts.

Practical applications of this modeling approach could benefit scenario planning for rapid-transit TOD's, especially when origin–destination travel information is absent or inaccessible. DDM models like the ones developed in this study could also be used for didactive purposes in teaching land-use/transportation interactions, predictive exercises, and/or preliminary benefit/cost evaluations of route alignment alternatives or line extensions.

Likewise, results from this study illustrate the importance of a multi-level modeling approach when dealing with a multi-modal rapid-transit networks. It provides further evidence of the benefits of a generalized linear regression approach with negative-binomial link function when working direct-demand models of transit ridership.

Nevertheless, there are limitations to this study. The first relates to the nature of single case-studies and their weaker generalizations. Still, it is plausible to expect similar results in similar contexts related to polycentric agglomerations and multi-modal networks. The second limitation relates to potential instrument validity issues with the local-accessibility indicator, WalkScore®, which lacks attention to crime, sidewalk completeness and crosswalks. Likewise, emerging micro-mobility and transportation network companies (TNC; e.g., Uber, Lyft) have potential impact as feeder modes. Nevertheless, when we consider that bus and walk access accounts for ~90% of all boardings in LAs rapid transit network their impact appears minor. Also, this study relied on aggregate data that does not allow for more nuanced calculations of accessibility measures along individual, time of day, demographic, and/or trip purpose characteristics (among many others).

These limitations open doors for future lines of research. Studies with more disaggregate data and higher number of observations based on other cities could further explore the issues and perhaps improve generalization. Improvement of WalkScore® ranking to incorporate pedestrian infrastructure attributes (such as sidewalk availability, completeness, and crosswalk presence); or development of new multi-dimensional *local-accessibility* indicators that emphasize pedestrian infrastructure elements could increase instrument validity. Similarly, incorporating new feeder modes into *local-* and *metropolitan-accessibility* indicators could further enhance DDM models. Meanwhile, transit analysts, TOD planners and policy analysts would benefit from including both independent and combined effects of *metropolitan-* and *local-accessibility* measures, and relevant interactions in station-level DDM models.

CRediT authorship contribution statement

Luis Enrique Ramos-Santiago: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

I would like to thank the 'Walk Score® Professional' Data Services (www.walkscore.com) who contributed at no cost to this investigation WalkScore rankings and associated measurement sub-components for LA's rapid-transit station portals (data provided by Redfin Real Estate; https://www.redfin.com).

Appendix A

Part A. Summary of recent transit studies where local and/or metropolitan accessibility indicators were incorporated in Models.

Author(s)	Year	Location	Number of Cases	Unit of Analysis	Number of Observations	Transit Mode	Outcome of Interest	Modeling	Methods	Explanatory Power	*Local Accessibility*	*Metropolitan / Regional Accessibility*	Accessibility Results
Gutiérrez et al.	2011	Madrid, Spain	1 city	Metro station	158	Metro transit	monthly boardings at stations	DDM	Distance-decay weighted regressions	Pseudo-R2:0.736	Distance-decay weighted independent variables	Nodal accessibility'; Cumulative opportunities measure (jobs)	Metropolitan accessibility (nodal) registers a highly significant positive relationship with boardings, and larger standardized *Beta* coefficient as compared to all other predictors in model.
Moniruzzaman & Páez	2012	Hamilton, Canada	1 city	DA (Dissemination Areas; ~ smaller than US Census *block-group* [pop])	761	Bus	% Share of commute by transit(vis-à-vis non-transit)	Discrete-Choice	Logistic regression w/ spatial lag component and control of over-dispersion	*Pseudo-* R^2:0.450	Distance to nearest bus-stop	Cumulative opportunities measure (jobs) in transit-served DA_i	Accessibility ***by*** transit (metropolitan) is highly significant and positively associated with higher transit commute share, with a modest effect. Accessibility ***to*** transit (local) is also significant and negatively associated with transit commute share.
Sung et al.	2014	Seoul metropolitan region,Korea	1 city	Rail stations	473	Intra- and Inter-Urban Railways	Rail transit ridership; average daily ridership (2010)	DDM	Spatial regressions (best fit: spatial error)	[0.258–0.522]	Number of station entrances and exist, and number of connecting bus lines	Distance to City Hall; distance to Gangnam station	Station-level accessibility (local accessibility) is as important as local land use for transit ridership, including well-connected transit networks (e.g. bus feeders)
Lin et al.	2014	Perth metropolitan area,Australia	1 city	Elderly train users at Train stations	7 train stations; 940 usable surveys	Rail transit: heavy-rail and light-rail	Rate of the elderly's patronage at train station	Other (Analytical Hierarchy Process)	Intercept survey; AHP	n/a	Place-based composite indicator that includes: mode (multiple), spatial separation, and activity opportunities; together with Analytic Hierarchy Process for weighting	n/a	The developed multidimensional and multi-modal accessibility composite-index (local accessibility) for elderly population helps in ranking of stations and identifying key factors associated with greater ease of access for the elderly. Shopping opportunities, seating availability at station, intermodal and network (topological) connectivity (route directness) are some of the highlighted factors.
Chen & Zegras	2016	Boston, Massachusetts (USA)	1 city;with 3 heavy-rail lines and 2 light-rail lines	Rapid-transit Stations	120	Rail transit: heavy-rail and light-rail	Station Weekdays Daily Boardings, and at peak, non-peak, and weekend periods	DDM	Bootstrapped OLS Regressions	adjusted-R^2: [0.764–0.795]	Urban design: % four-way intersections, walk-index, avg. road width, sidewalk density, land-use mix, retail employment density	Cumulative opportunities measure (jobs) in transit-served areas	Together with attributes associated with TOD developments (e.g. population and employment densities and street network connectivity), and intermodality (e.g. bus connections), 'Accessibility'

(*continued on next page*)

(*continued*)

Author(s)	Year	Location	Number of Cases	Unit of Analysis	Number of Observations	Transit Mode	Outcome of Interest	Modeling	Methods	Explanatory Power	*Local Accessibility*	*Metropolitan / Regional Accessibility*	Accessibility Results
													(metropolitan accessibility) is not only highly significant but also registers the largest elasticity of all predictive variables.
Chowdhury et al.	2016	Auckland,New Zealand	1 city	Users of two (2) PT terminal stations	300	Bus rapid transit	M1: attitudes towards public transport (satisfaction with existing PT services); M2: intention to use PT in the future	Discrete-Choice	User-preference survey; ordinal and logistic regressions	[0.360–0.420]	Access time or distance to PT terminal/stop	Likert scale on question assessing ease of access to suburbs from terminal station.	Ease of access to terminals (local accessibility) and connectivity to various destinations (metropolitan accessibility) are statistically significant in affecting attitudes towards public transport, and that both types of accessibility need to be of high standard.

Part B. Summary of recent transit studies where local and/or metropolitan accessibility indicators were incorporated in models (Continuation)

Author(s)	Year	Location	Number of Cases	Unit of Analysis	Number of Observations	Transit Mode	Outcome of Interest	Modeling	Methods	Explanatory Power	*Local Accessibility*	*Metropolitan/ Regional Accessibility*	Accessibility Results
Li et al.	2017	Xi'an,China	1 city;1 line is used for attraction accessibility analysis; 4 lines are used for radiation accessibility analysis	Metro Stations	2553 metro riders' surveys for M1; 225 questionnaires for M2	Metro transit	Comparison of attraction accessibility and radiation accessibility for policy analysis	Other (Development and Comparison of Quantitative Accessibility Indicators)	Utility Model(for multi-modal 'attraction accessibility' indicator) andSpace-Syntax (for 'radiation accessibility' indicator)	n/a	"Attraction accessibility"; ease of reaching a station considering time, fare, and fatigue costs (questionnaire survey): walking connection + bus connection; utility-based method	"Radiation accessibility"; measured via an axis model of space-syntax (integration); and transfer stations function (network topological attribute)	1- "attraction accessibility" (local) by bus was lower than that by foot, which reflects actual access mode share in Xi'an; attraction depends mostly on walking connection;2- "radiation accessibility" (metropolitan) reveals the importance of transfer stations and variation based on topological attributes;3- bus crowdedness and

(*continued on next page*)

(*continued*)

Author(s)	Year	Location	Number of Cases	Unit of Analysis	Number of Observations	Transit Mode	Outcome of Interest	Modeling	Methods	Explanatory Power	*Local Accessibility*	*Metropolitan/ Regional Accessibility*	Accessibility Results
													wait time were key detractors; authors recommend improved bus frequency, streamlined bus-rail interchange, among other service improvements geared for a more comfortable and convenient link to rail station.
Aston et al.	2020	[various], International	146 studies	Transit travel and built-environment studies	467 models;1662 data points	[various]	Influence of research design on built-environment/ transit studies results along four dimensions: Density, Diversity, Design, Accessibility (regional)	Other(Meta-Regression)	Meta-regression	[various]; correlation coefficients and significance levels	[various]	[various]	Accessibility (regional) is significant and correlated in reducing bias in built-environment/ transit ridership studies. It is recommended as best-practice to control for regional accessibility in model specifications.
Cui et al.	2020	Vancouver, Calgary, Edmonton, Winnipeg, London, Kitchener-Cambridge-Waterloo, Toronto-Hamilton, Ottawa-Gatineau, Montreal, Quebec City, Halifax; Canada	11 cities	Canadian metropolitan regions; census tracts.	106–109	Public Transit(bus, subway, elevated and light-rail, streetcar, commuter train, passenger ferry)	% Share of commuters using public transport, living census-tractacross low- and high-income groups	DDM	Linear regressions	[0.432–0.781]	Network distance to nearest rapid transit station	Cumulative accessibility measures (jobs) by income-group (low vs. high)	Lower-order quadratic term of percentage of jobs accessible by public transport (metropolitan) is positively associated with mode share, except for higher-income groups in a few metropolitan regions. The higher-order term of Access registers a negative directionality (concave parabola) that suggests a threshold where

(*continued on next page*)

(*continued*)

Author(s)	Year	Location	Number of Cases	Unit of Analysis	Number of Observations	Transit Mode	Outcome of Interest	Modeling	Methods	Explanatory Power	*Local Accessibility*	*Metropolitan/ Regional Accessibility*	Accessibility Results
													more accessibility will not produce more public transit share.
Wu & Levinson	2021	Analytical model (theoretical)	n/a	Station-Spacing		Urban train	Person-weighted accessibility	Other (analytical model)	Mathematical analytical model	n/a	n/a	Cumulative opportunities measure (jobs) in transit-station served areas	Through an analytical accessibility analysis, the authors find that for each transit service type, there is an optimal station spacing that maximizes accessibility. Operational speed and vehicle gate configurations are key parameters to consider.

Appendix B. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.trip.2023.100834.

References

Aston, L., Currie, G., Kamruzzaman, M., Delbosc, A., Teller, D., 2020. Study design impacts on built environment and transit use research. J. Transp. Geogr. 82, 102625.
Balcombe, R., Mackett, R., Paulley, N., Preston, J., Shires, J., Titheridge, H., et al. 2004. The demand for public transport: a practical guide.
Brown, J.R., Thompson, G.L., 2008. Service orientation, bus–rail service integration, and transit performance: examination of 45 US metropolitan areas. Transp. Res. Rec. 2042 (1), 82–89.
Brown, J.R., Thompson, G.L., 2012. Should transit serve the CBD or a diverse array of destinations? A case study comparison of two transit systems. J. Public Transp. 15 (1), 1.
Cameron, A.C., Trivedi, P.K., 2010. Microeconometrics Using Stata, Vol. 2. Stata Press, College Station, TX.
Cao, J., Duncan, M., 2019. Associations among distance, quality, and safety when walking from a park-and-ride facility to the transit station in the Twin Cities. J. Plan. Educ. Res. 39 (4), 496–507.
Cardozo, O.D., García-Palomares, J.C., Gutiérrez, J., 2012. Application of geographically weighted regression to the direct forecasting of transit ridership at station-level. Appl. Geogr. 34, 548–558.
Cervero, R., 2001. Walk-and-ride: factors influencing pedestrian access to transit. J. Public Transp. 3 (4), 1–23.
Cervero, R., 2006. Alternative approaches to modeling the travel-demand impacts of smart growth. J. Am. Plann. Assoc. 72 (3), 285–295.
Cervero, R., Duncan, M., 2002. Land value impacts of rail transit services in Los Angeles County. Rep. Prep. Natl. Assoc. Realtor Urban Land Inst.
Cervero, R., Round, A., Goldman, T., Wu, K.L. 1995. Rail access modes and catchment areas for the BART system. 1995.
Cervero, R., Murakami, J., Miller, M., 2010. Direct ridership model of bus rapid transit in Los Angeles County, California. Transp. Res. Rec. 2145 (1), 1–7.
Chakraborty, A., Mishra, S., 2013. Land use and transit ridership connections: Implications for state-level planning agencies. Land Use Policy 30 (1), 458–469.
Chen, S., Zegras, C., 2016. Rail transit ridership: station-area analysis of Boston's Massachusetts Bay transportation authority. Transp. Res. Rec. 2544 (1), 110–122.
Chowdhury, S., Zhai, K., Khan, A., 2016. The effects of access and accessibility on public transport users' attitudes. J. Public Transp. 19 (1), 97–113.
Cui, B., Boisjoly, G., Miranda-Moreno, L., El-Geneidy, A., 2020. Accessibility matters: exploring the determinants of public transport mode share across income groups in Canadian cities. Transp. Res. Part Transp. Environ. 80, 102276.
Currie, G., Ahern, A., Delbosc, A., 2011. Exploring the drivers of light rail ridership: an empirical route level analysis of selected Australian, North American and European systems. Transportation 38 (3), 545–560.
Duncan, M., 2010. To park or to develop: trade-off in rail transit passenger demand. J. Plan. Educ. Res. 30 (2), 162–181.
Duncan, D.T., Aldstadt, J., Whalen, J., Melly, S.J., 2013. Validation of Walk Scores and Transit Scores for estimating neighborhood walkability and transit availability: a small-area analysis. GeoJournal 78 (2), 407–416.
Durning, M., Townsend, C., 2015. Direct ridership model of rail rapid transit systems in Canada. Transp. Res. Rec. 2537 (1), 96–102.
Elldér, E., 2020. What kind of compact development makes people drive less? The "ds of the built environment" versus neighborhood amenities. J. Plan. Educ. Res. 40 (4), 432–446.
Fabrigar, L.R., Wegener, D.T., MacCallum, R.C., Strahan, E.J., 1999. Evaluating the use of exploratory factor analysis in psychological research. Psychol. Methods 4 (3), 272.
Foletta, N., Vanderkwaak, N., Grandy, B., 2013. Factors that influence urban streetcar ridership in the United States. Transp. Res. Rec. 2353 (1), 92–99.
Garson, G. David. 2013. "Fundamentals of hierarchical linear and multilevel modeling." *Hierarchical linear modeling: Guide and applications* 3-25.
Geurs, K.T., Van Wee, B., 2004. Accessibility evaluation of land-use and transport strategies: review and research directions. J. Transp. Geogr. 12 (2), 127–140.
Gordon, P., Richardson, H.W., 1996. Beyond polycentricity: the dispersed metropolis, Los Angeles, 1970–1990. J. Am. Plann. Assoc. 62 (3), 289–295.
Guerra, E., Cervero, R., 2011. Cost of a ride: the effects of densities on fixed-guideway transit ridership and costs. J. Am. Plann. Assoc. 77 (3), 267–290.
Gutiérrez, J., Cardozo, O.D., 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.
Hall, C.M., Ram, Y., 2018. Walk score® and its potential contribution to the study of active transport and walkability: a critical and systematic review. Transp. Res. Part Transp. Environ. 61, 310–324.
Hansen, W.G., 1959. How accessibility shapes land use. J. Am. Inst. Plann. 25 (2), 73–76.
Harris, C.D., 1954. The market as a factor in the localization of industry in the United States. Ann. Assoc. Am. Geogr. 44 (4), 315–348.
Hilbe, J.M. (Ed.), 2011. Negative Binomial Regression. Cambridge University Press.
Hirsch, J.A., Moore, K.A., Evenson, K.R., Rodriguez, D.A., Roux, A.V.D., 2013. Walk Score® and Transit Score® and walking in the multi-ethnic study of atherosclerosis. Am. J. Prev. Med. 45 (2), 158–166.
Horner, M.W., 2004. Exploring metropolitan accessibility and urban structure. Urban Geogr. 25 (3), 264–284.

Karner, A., 2022. People-focused and near-term public transit performance analysis. J. Public Transp. 24, 100019.
Kim, S., Ulfarsson, G.F., Hennessy, J.T., 2007. Analysis of light rail rider travel behavior: impacts of individual, built environment, and crime characteristics on transit access. Transp. Res. Part Policy Pract. 41 (6), 511–522.
Kuby, M., Barranda, A., Upchurch, C., 2004. Factors influencing light-rail station boardings in the United States. Transp. Res. Part Policy Pract. 38 (3), 223–247.
LACMTA (Los Angeles County Metropolitan Transportation Authority) - PTV NuStats. *Los Angeles County Metropolitan Transportation Authority System-Wide On-Board Origin-Destination Study, Draft Final Report*. 2012.
Levinson, David M., Wesley Marshall, Kay Axhausen. 2017. *Elements of Access: Transport Planning for Engineers, Transport Engineering for Planners*. Network Design Lab.
Levinson, D., Wu, H., 2020. Towards a general theory of access. J. Transp. Land Use 13 (1), 129–158.
Li, L., Ren, H., Zhao, S., Duan, Z., Zhang, Y., Zhang, A., 2017. Two dimensional accessibility analysis of metro stations in Xi'an, China. Transp. Res. Part Policy Pract. 106, 414–426.
Lin, T.G., Xia, J.C., Robinson, T.P., Goulias, K.G., Church, R.L., Olaru, D., Tapin, J., Han, R., 2014. Spatial analysis of access to and accessibility surrounding train stations: a case study of accessibility for the elderly in Perth, Western Australia. J. Transp. Geogr. 39, 111–120.
Litman, T.A. 2021. Evaluating Accessibility for Transport Planning-Measuring People's Ability to Reach Desired Services and Activities-22 April 2021.
McLeod, Jr M.S., Flannelly, K.J., Flannelly, L., Behnke, R.W. 1991. Multivariate Time-Series Model of Transit Ridership Based on Historical, Aggregate Data: the Past, Present and Future of Honolulu.
Mees, P., 2009. Transport for Suburbia: Beyond the Automobile Age. Routledge.
Mees, P., Stone, J., Imran, M., Nielson, G. 2010. Public transport network planning: a guide to best practice in NZ cities. Report No.: 0478352905.
Moniruzzaman, M., Páez, A., 2012. Accessibility to transit, by transit, and mode share: application of a logistic model with spatial filters. J. Transp. Geogr. 24, 198–205.
Næss, P., Strand, A., Wolday, F., Stefansdottir, H., 2019. Residential location, commuting and non-work travel in two urban areas of different size and with different center structures. Prog. Plan. 128, 1–36.
Park, S., Kang, J., Choi, K., 2014. Finding determinants of transit users' walking and biking access trips to the station: a pilot case study. KSCE J. Civ. Eng. 18 (2), 651–658.
Parsons, Brinckerhoff, Quade & Douglas, United States. Federal Transit Administration and Transit Development Corporation, 1996. Transit and Urban Form: pt. III, A guidebook for practioners; pt. IV, Public policy and transit-oriented development (Vol. 16). National Academy Press.
Pushkarev, B.S., Zupan, J.M., 1977. Public Transportation and Land Use Policy. Indiana University Press, Bloomington.
Ramos-Santiago, L.E., 2021. Towards a better account and understanding of bus/rapid-transit interactions: the case of Los Angeles. Case Stud. Transp. Policy 9 (3), 1167–1179.
Ramos-Santiago, L.E., Brown, J., 2016. A comparative assessment of the factors associated with station-level streetcar versus light rail transit ridership in the United States. Urban Stud. 53 (5), 915–935.
Ramos-Santiago, L.E., Brown, J.R., Nixon, H., 2015. Transit performance of modern-era streetcars: consideration of five US cities. Transp. Res. Rec. 2534 (1), 57–67.
Ramos-Santiago, L.E., Novales, M., Varela-García, F.A., 2022. Identifying and understanding determinants of regional differences in light-rail patronage and performance. Case Stud. Transp. Policy 10 (2), 1188–1206.
Schlossberg, M., Brown, N., 2004. Comparing transit-oriented development sites by walkability indicators. Transp. Res. Rec. 1887 (1), 34–42.
Sohn, K., Shim, H., 2010. Factors generating boardings at metro stations in the Seoul metropolitan area. Cities 27 (5), 358–368.
Southworth, M., 2005. Designing the walkable city. J. Urban Plan. Dev. 131 (4), 246–257.
Sung, H., Choi, K., Lee, S., Cheon, S., 2014. Exploring the impacts of land use by service coverage and station-level accessibility on rail transit ridership. J. Transp. Geogr. 36, 134–140.
Taylor, B.D., Fink, C.N. 2003. The factors influencing transit ridership: a review and analysis of the ridership literature.
Systematics, C., 2012. Inc. NCHRP Report 716: Travel Demand Forecasting: Parameters and Techniques. Transportation Research Board of the National Academies. Washington, DC.
Taylor, B.D., Miller, D., Iseki, H., Fink, C., 2009. Nature and/or nurture? Analyzing the determinants of transit ridership across US urbanized areas. Transp. Res. Part Policy Pract. 43 (1), 60–77.
Tilahun, N., Li, M., 2015. Walking access to transit stations: evaluating barriers with stated preference. Transp. Res. Rec. 2534 (1), 16–23.
Walk Score. Walk Score Methodology [Internet]. 2019. Available from: https://www.walkscore.com/methodology.shtml.
Tilahun, N., Li, M., 2015. Walking access to transit stations: evaluating barriers with stated preference. Transportation research record 2534 (1), 16–23.
Walker, J., 2012. Human Transit: How Clearer Thinking About Public Transit Can Enrich Our Communities and Our Lives. Island Press.
Wardman, M., Whelan, G., Toner, J.P. 1994. Direct Demand Models of Air Travel: A Novel Approach to the Analysis of Stated Preference Data.
Watkins, M.W., 2018. Exploratory factor analysis: a guide to best practice. J. Black Psychol. 44 (3), 219–246.
Weinberger, R., Sweet, M.N., 2012. Integrating walkability into planning practice. Transp. Res. Rec. 2322 (1), 20–30.
Weinstein Agrawal, A., Schlossberg, M., Irvin, K., 2008. How far, by which route and why? A spatial analysis of pedestrian preference. J. Urban Des. 13 (1), 81–98.
Wu, H., Levinson, D., 2021. Optimum stop spacing for accessibility. Eur. J. Transp. Infrastruct. Res. 21 (2), 1–18.
Zhao, J., Deng, W., Song, Y., Zhu, Y., 2014. Analysis of Metro ridership at station level and station-to-station level in Nanjing: an approach based on direct demand models. Transportation 41 (1), 133–155.