



People-focused and near-term public transit performance analysis

Alex Karner

The University of Texas at Austin, United States



ARTICLE INFO

Keywords:

Performance management
Accessibility
Revealed behavior
Cumulative opportunities
Near-term forecasting

ABSTRACT

Public transit ridership forecasts have long played a role in understanding the potential success of a policy or investment, but their limitations have led researchers and practitioners to identify other performance analysis approaches. Accessibility, or the ease of reaching opportunities, has become very popular and widely used for this purpose. But commonly used accessibility measures also embody weaknesses that are seldom acknowledged; these limit their utility for truly understanding the benefits of transit investments. In this paper, we identify the pros and cons of these competing approaches and offer a third strategy. Specifically, we describe how revealed travel behavior data, potentially combined with near-term forecasts, can provide information about how current public transit users will be affected by a new project. While acknowledging the limitations of this approach, we demonstrate how accessibility can be misleading when applied without an understanding of ridership patterns.

1. Introduction

How should practitioners determine whether a particular public transit capital investment or service change is *good*? This question is omnipresent in public transit, since budgets are usually highly constrained, needs often outstrip what can feasibly be done, and the public holds transit projects to a higher standard than highway investments. Because public transit's subsidies are much easier to track, the public demands a higher level of accountability in their expenditure. A cynic would argue that political considerations dominate transit investment decisions, and they would not be wrong (e.g., Hamer, 1976; Kain, 1990; Pickrell, 1992). But demonstrating that a rational basis exists for determining how public transit investments can be prioritized and making those methods more widely available to the public can diminish the influence of rank politics over time.

Two approaches to public transit performance analysis are prevalent today: (1) ridership forecasting and (2) access to opportunities (commonly referred to as *accessibility* or *freedom*) analyses. Ridership forecasting involves combining estimates of demographics, transit levels of service, land use, and ridership to calibrate statistical models that match observations in a base year. These models are subsequently applied to a future year to generate ridership estimates. Decades of practice have emphasized the importance of forecasting future travel demand and guiding near-term decisions using estimates of long-term performance, even though forecasts often turn out to have been inaccurate (Voulgaris, 2020a; Flyvbjerg et al., 2005). Nevertheless, ridership estimates facilitate the creation of performance estimates, like

traditional cost-benefit measures, that can be used to compare between possible investment strategies and/or service changes.

Accessibility-based approaches, sometimes referred to as *freedom* analyses (Walker, 2018), are also very common in practice. Drawing from the vast academic literature on accessibility (e.g., Geurs and van Wee, 2004; Handy and Niemeier, 1997; Páez et al., 2012), this approach identifies the ease of reaching opportunities as the most important benefit of a transportation system (see also Martens, 2012, 2016). Typical implementations count the number of available opportunities (e.g., jobs, square feet of retail space, number of medical facilities) within a travel time threshold (e.g., 45 min) of every origin location where walking to public transit service is viable. With advances in data availability through the General Transit Feed Specification (GTFS) standard, as well as in open-source and/or widely available network routing software such as OpenTripPlanner, travel times throughout a network can be easily generated (Farber and Fu, 2017; Karner, 2018). Accessibility can be assessed for existing conditions (Levinson, 2013; Allen and Farber, 2019) and for proposed system changes (Pereira, 2019). In each case, results can be summarized for population groups to determine their average accessibility levels (Rowangould et al., 2016). In contrast to ridership forecasting, accessibility analyses do not require information about future demographics and land uses, although such data could be employed if available.

When evaluating public transit system performance, these two approaches embody multiple shortcomings that will be discussed further below, but it is important to note that a third approach is available that has many inherent advantages: current-year analysis. Assessing

E-mail address: alex.karner@utexas.edu

<https://doi.org/10.5038/2375-0901.23.2.1>

proposed performance in the current year requires a good understanding of who is using public transit in a region for what purpose, usually gleaned from a high-quality (i.e., statistically controlled and properly expanded) transit rider survey. The rider survey can be used to calibrate a ridership model or to examine how a network change will affect existing travelers by looking at differences in their before/after trip characteristics including travel time, waiting time, number of transfers, overall utility, and so on. Current-year model estimates can also be compared to observed behavior data to understand uncertainty and inform the interpretation of long-range projections. In analyzing applications to its Capital Investment Grant (CIG) program, the US Federal Transit Administration requires sponsors to assess project performance based on current conditions; future-oriented analyses are conducted only at the option of the sponsor (FTA, 2016).

In the interests of contributing to the debate and improving public transit performance analysis, this paper argues that the prevailing approaches embody underappreciated shortcomings. First, those shortcomings are documented with reference to ridership forecasting and typical access to opportunities approaches. An empirical example follows. The final section makes the case for near-term analyses focused on the decisions that current travelers actually make, while engaging with some of the criticisms this approach has faced from elsewhere in the field. In the conclusions, some reflections are offered on the future of public transit performance analysis.

The trouble with forecasts

In the language of demography, *projections* are conditional statements about the future dependent on the veracity of the input data and the posited relationships between inputs and outputs, whereas *forecasts* are a demographer's best guess as to the future that will come to pass (e.g., Smith et al., 2013). One key difference in whether transportation professionals are presenting projections or forecasts involves how they perceive their role or how they are expected to comport themselves based on the needs and requirements of their organization and/or their superiors (Howe and Kaufman, 1979; Wachs, 2001). While discussions of ethics have shifted to favor presenting multiple scenarios/projections and interpreting their tradeoffs for the public, much of practice still involves presenting a single forecast and defending its accuracy (Klosterman, 2013; Wachs, 2001; Voulgaris, 2019).

There are multiple reasons why a single forecast is unlikely to be correct. And ample empirical evidence exists documenting that public transit ridership forecasts are almost always overestimated (Flyvbjerg et al., 2005). Travel demand models—the workhorse tool within transportation planning and the source of most public transit ridership estimates—embody many well-known limitations. But these limitations are seldom acknowledged by practicing transportation planners and travel demand modelers when communicating results. These include the following:

- **Importance of input data.** Projecting travel demand for any mode requires information about future land uses, demographics, and infrastructure. It is these inputs that often drive model outputs regardless of paradigm (four-step, activity based, sketch planning, etc.). Because of the importance of input data, models can be easily made to produce estimates that make projects appear maximally (un)attractive and systematic mischaracterizations are common (Flyvbjerg et al., 2005; Wachs, 1989), although the situation may be improving in the United States (Voulgaris, 2020b).
- **Inability to anticipate unexpected events.** Related to issues with input data, broader socioeconomic shifts (e.g., recession/depression, employment by sector, pandemic) and demographic changes, such as the suburbanization of poverty, are difficult to predict with any degree of accuracy, yet they fundamentally drive future outcomes.

- **Reliance on complex “black box” models.** Travel demand models are inherently and increasingly complex. They are difficult for the lay public to understand, creating a barrier to public involvement (Handy, 2008; Marcantonio and Karner, 2014). Consultants are often relied upon to complete and analyze model runs, taking decision-making power away from the public agencies tasked with transportation decision-making.

- **Expensive software.** The industry has coalesced around several vendors who provide the base software upon which travel demand models are run (including Bentley Systems, Caliper Corporation, PTV Group, and INRO). The software is objectively expensive to operate and maintain, with annual licenses typically costing thousands of dollars.

If these limitations are acknowledged and understood, model-based projections can be exceptionally worthwhile, helping stakeholders understand how decisions made today are likely to affect future outcomes, at least in terms of their magnitude and direction. It is also possible to assess a projection's quality; some are better than others, especially when certain criteria are met and addressed (Voulgaris, 2019). The Federal Transit Administration (FTA) has led an effort to standardize public transit ridership forecasting by developing a tool called Simplified Trips-on-Project Software (STOPS). Through standardization, one of the major limitations of model-based projections can be mitigated, specifically the tendency to manipulate input data to obtain desired outcomes.

STOPS also serves a regulatory purpose. It represents one approach that CIG applicants can use to prepare FTA's required current-year ridership forecast (FTA, 2016). Making near-term instead of longer-term projections generates multiple benefits. In describing its use of current-year forecasts, FTA states that

project evaluation based on existing conditions provides the most easily understood, most reliable, and most readily available information for decision-making...Use of current year data increases the reliability of the projected future performance of the proposed project by avoiding reliance on future population, employment, and transit service levels that are themselves forecasts (FTA, 2016, 10).

Access to opportunities and “freedom” as an alternative

The well-known issues with forecasting have generated demand for different performance assessment methods. If we cannot reliably predict the number of riders on public transit in the future, what should planners and concerned members of the public do? Many have begun to turn to accessibility measures that give a sense of how easily destinations can be reached before and after a project or system change. These types of analyses need not rely on uncertain future projections and require only information about the proposed changes (and optionally demographic data).

Accessibility indicators have substantial theoretical appeal as they are rooted in the notion of travel as a derived demand and reflect transportation's fundamental purpose (Levine et al., 2019). People travel not for its own sake, but to reach opportunities—accessibility indicators capture the extent to which these connections are possible. Well-designed accessibility measures are powerful planning tools; they recognize that both transportation (e.g., increasing travel time, reliability, affordability, or comfort) and land use (e.g., increasing density or diversity) interventions affect transportation system performance and travel demand (Handy and Niemeier, 1997). New methods for evaluating accessibility and a proliferation of data sources are rapidly expanding the frontier of accessibility analysis.

Cumulative opportunities measures are by far the most widely applied accessibility measures in practice and in academic research (e.g., Golub et al., 2013; Golub and Martens, 2014; Owen and Levinson, 2014).

These quantify the opportunities that can be reached from a given origin location, are calculated for a geographic unit (e.g., census tracts or block groups), and most often use constant travel time thresholds (e.g., number of destinations reachable within a 45-minute door-to-door public transit trip) rather than information about how people travel.

A representative example of the cumulative opportunities-based approach is the analysis undertaken by Jarrett Walker and Associates for the Dublin bus network redesign (Jarrett Walker + Associates, 2018). Bus network redesigns are becoming increasingly common. They typically seek to increase the number of high-frequency routes while keeping an agency's operating budget constant. This change usually requires cutting peripheral service and redirecting resources saved into routes likely to generate additional riders. The Dublin example shows how changes in cumulative opportunities can be assessed before and after a proposed system change.

Chapter 8 of the Dublin redesign's 2018 Public Consultation Report is entitled "Network Outcomes."¹ It presents two types of analyses. The first tallies changes in the number of people and jobs located within walking distance of frequent public transit. The second calculates multiple different cumulative opportunities measures, providing estimates of changes in the number of jobs accessible within specified travel time thresholds. Fig. 1 shows a sample of the results, which summarizes the change in the number of jobs accessible within 30 min from origins located near transit stops in Dublin.

Fig. 1 also summarizes the number of jobs accessible from every origin location, but it does not demonstrate precisely *which jobs* are accessible. This information is contained in several appendices to the report that summarize the spatial reach of public transit from one origin using isochrones. An example isochrone map is shown in Fig. 2. The utility of such a map is clear for area residents who have cause to travel to or from the origin location shown; they can clearly see how a proposed service change will affect their ability to reach specific places. All possible isochrones (each appendix contains ~70 pages with one isochrone each) can be summarized using a map like Fig. 1. This approach loses directional information but enables one to visualize broader regional trends.

Fig. 1's results can also be further summarized to generate regional performance measures. For example, one of the key conclusions for Dublin is that, "On average, Dubliners would be able to reach 18% more jobs in 30 min of travel time than today" (Jarrett Walker + Associates, 2018, 155). Although equations are not presented, this result is described as a change in a population-weighted accessibility score. This approach is commonly applied in the literature and is a reasonable way to summarize the average accessibility change experienced by a randomly selected resident. It can be calculated for different groups of interest (e.g., Golub and Martens, 2014; Rowangould et al., 2016).

Despite its ubiquity, population-weighted accessibility is an imprecise measure of the benefits of a public transit improvement because it is not linked to the trips that people want or need to make (see also Karner and Golub, 2015; Farber et al., 2016). In traditional travel behavior analysis, differences in the relative attractiveness of trip-mode combinations before and after a change are assessed and converted to equivalent time or money units, then interpreted as a measure of consumer surplus (e.g., Niemeier, 1997; de Jong et al. 2007). These differences can be interpreted as a measure of the "user benefits" of a proposed system change (e.g., Chorus and Timmermans, 2009). Conveniently, they also represent changes in the ease with which opportunities can be reached—an alternative measure of accessibility (Handy and Niemeier, 1997; Geurs and van Wee, 2004). Simpler measures such

¹ This document is also referred to as the *Revised Network Report June 2018*. It was revised again in October 2020 to reflect responses to public feedback and changes in network design. The performance measures presented in both the 2018 and 2020 versions are very similar.

as travel time changes can also be calculated and similarly interpreted (e.g., Manaugh and El-Geneidy, 2012).

Beyond the Dublin example, the popularity of cumulative opportunities accessibility measures can be seen in the sheer number of software vendors and open-source solutions available for measuring and assessing accessibility under different circumstances, as well as their widespread use. CUBE Access and Remix are both examples of commercially available software for calculating cumulative opportunities accessibility, isochrones, and similar measures. Open-source solutions include OpenTripPlanner Analyst and the Python Pandana library. Examples in practice abound and include the Puget Sound Regional Council's (2018) regional transportation plan (Puget Sound Regional Council, 2018), Virginia Department of Transportation's SMART SCALE project prioritization process (Virginia Department of Transportation, 2020), and the University of Minnesota Accessibility Observatory (Owen and Levinson, 2014).

Despite the widespread popularity and mass appeal of cumulative opportunities measures and isochrones, they embody two major limitations for understanding infrastructure and policy impacts that are scarcely acknowledged in the literature or in practice. These include the following:

- **Limited utility for assessing regional benefits.** As demonstrated in the Dublin example, isochrone maps can be prepared that show the number of destinations that can be reached from any origin location in a region. These types of maps can be profoundly useful for individuals and businesses seeking to understand how a change in transit service will affect their ability to reach opportunities or customers (e.g., Stewart, 2017). Population-weighted average changes in accessibility can also be calculated, but these do not represent user benefits in the traditional sense since changes may or may not be well aligned with the trips people want or need to make.
- **No integration of public transit market information.** The limited utility of a regional benefit analysis based on access to opportunities arises in part because of the absence of information about the public transit market. It is entirely possible to imagine a scenario where apparent accessibility increases substantially for large numbers of people because of a proposed capital improvement or service change but where the actual/future transit market is quite weak. If such an investment or change is pursued on the grounds that it will increase accessibility, it could ultimately perform worse in terms of helping people reach their destinations by public transit than a comparable change whose apparent access to opportunities benefit is smaller.

One immediate way to address these weaknesses is to refocus the analysis to consider who is using public transit today, for what purpose, and for what destination. These types of analyses are addressed in the next section.

People-focused near-term analyses as an antidote

The central tenet of the near-term analysis focused on users proposed here is that, in order to understand public transit system performance, one must have a well-developed understanding of current public transit riders and their travel patterns so that changes in trip characteristics, overall utility, or consumer surplus can be calculated. This perspective is surprisingly absent from contemporary transit performance analyses, despite the ongoing importance of assessing quantities such as changes in travel time and utility/consumer welfare across planning practice (e.g., Manaugh and El-Geneidy, 2012; Niemeier, 1997; Martens and Di Ciommo, 2017).

This absence can be explained in part by the withering criticism leveled at measures derived from revealed or modeled behavior from various sources in the literature. Without exception, these sources argue for access to opportunities approaches without acknowledging their

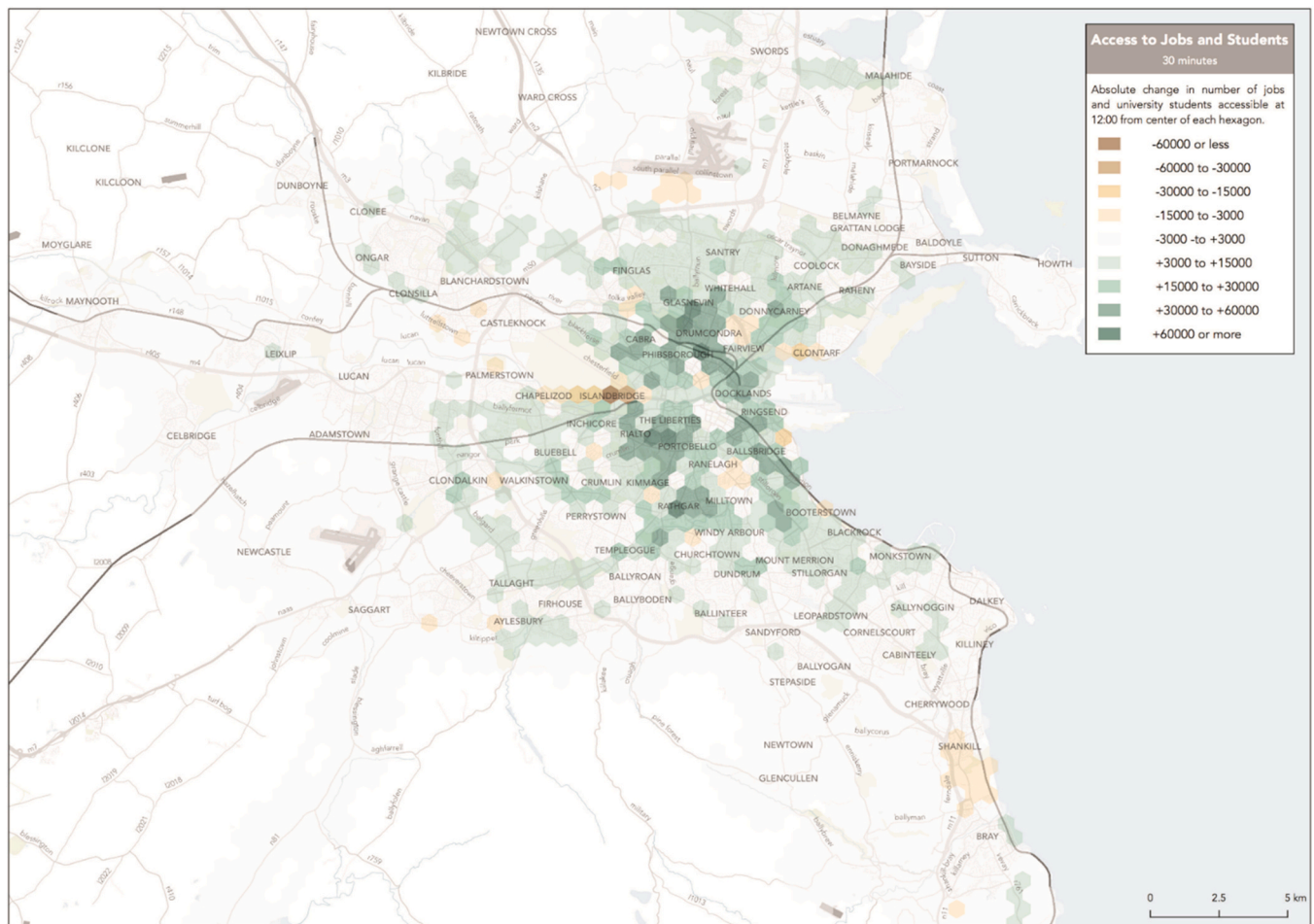


Fig. 1. Change in number of jobs and students accessible within 30 min due to the Dublin bus network redesign (Jarrett Walker + Associates, 2018, 162).

inherent limitations. For example, Pereira et al. (2017) highlight the important role of constraints that exist outside of individual/household control. These constraints (residential location, income, family structure, mandatory activity locations, etc.) interact with preferences to determine revealed travel behavior outcomes. They argue that revealed behaviors that seem to indicate disadvantage, such as long commute distances, can only be considered problematic to the extent that they arise because of constraints.

The notion of constraints versus preferences is important but can be thoughtfully addressed. One simply needs to determine the appropriate criteria that can be used to segment the population and home in on the group of interest. For example, it has long been understood that some public transit users have no other option while others freely choose to ride (Garrett and Taylor, 1999; Taylor and Morris, 2015). Other work has differentiated between car-less (constrained) and car-free (choice) households using responses to attitudinal questions on a household activity survey (Brown, 2017). That work has demonstrated both socioeconomic as well as travel behavioral differences between households that choose to not have a vehicle and those that wish to but cannot. Simply grouping all zero-vehicle households or all public transit users together to calculate performance measures is inappropriate, but with care, relatively homogeneous groups can be identified, and their characteristics studied.

Another similar issue that is sometimes raised is the problem of comparing baseline conditions between groups. Levine et al. (2019, 9) describe a hypothetical situation in which women are found to have shorter commute times than men. In this case, they state that we should not conclude that women derive outside benefits from the existing transportation system. This is correct; differences between groups must

be understood in context. More travel at baseline can be positive or negative. But a proposed system change that reduced commute times for women more than men given existing travel patterns, or that enhanced available travel choices for women at a high rate, would provide valuable information to the public and decision-makers about transportation system benefits. Changes are more straightforward to interpret than baseline conditions.

Martens and Golub (2018) and Pereira et al. (2017) raise issues related to travel patterns shifting over time as locations of family, friends, and mandatory activity locations also shift. Walker (2018, 125) makes the point most strongly, asking:

When you went shopping at a particular store, does it matter that you *could have* gone shopping somewhere else, or shopped online while in bed, or embraced an ascetic spiritual path of buying as little as possible? A study of freedom would be intensely interested in that, while conventional planning would merely record what you did and use that to predict what you, despite your illusion of freedom, will continue to do. (emphasis in original).

This line of argumentation misapprehends the nature of planning-related analysis and falls prey to the individualistic nature of the isochrone. When assessing regional transportation plan performance, outcomes for *individuals* are never of concern. Instead, analysts aggregate results over population groups, places, or the entire region to determine performance. Even though an individual's travel patterns will likely change as circumstances change, including many different people and their (revealed) travel within a single performance measure ensures that the measures capture these variations in behavior. For the same reason, travel surveys collected on a single day are used to

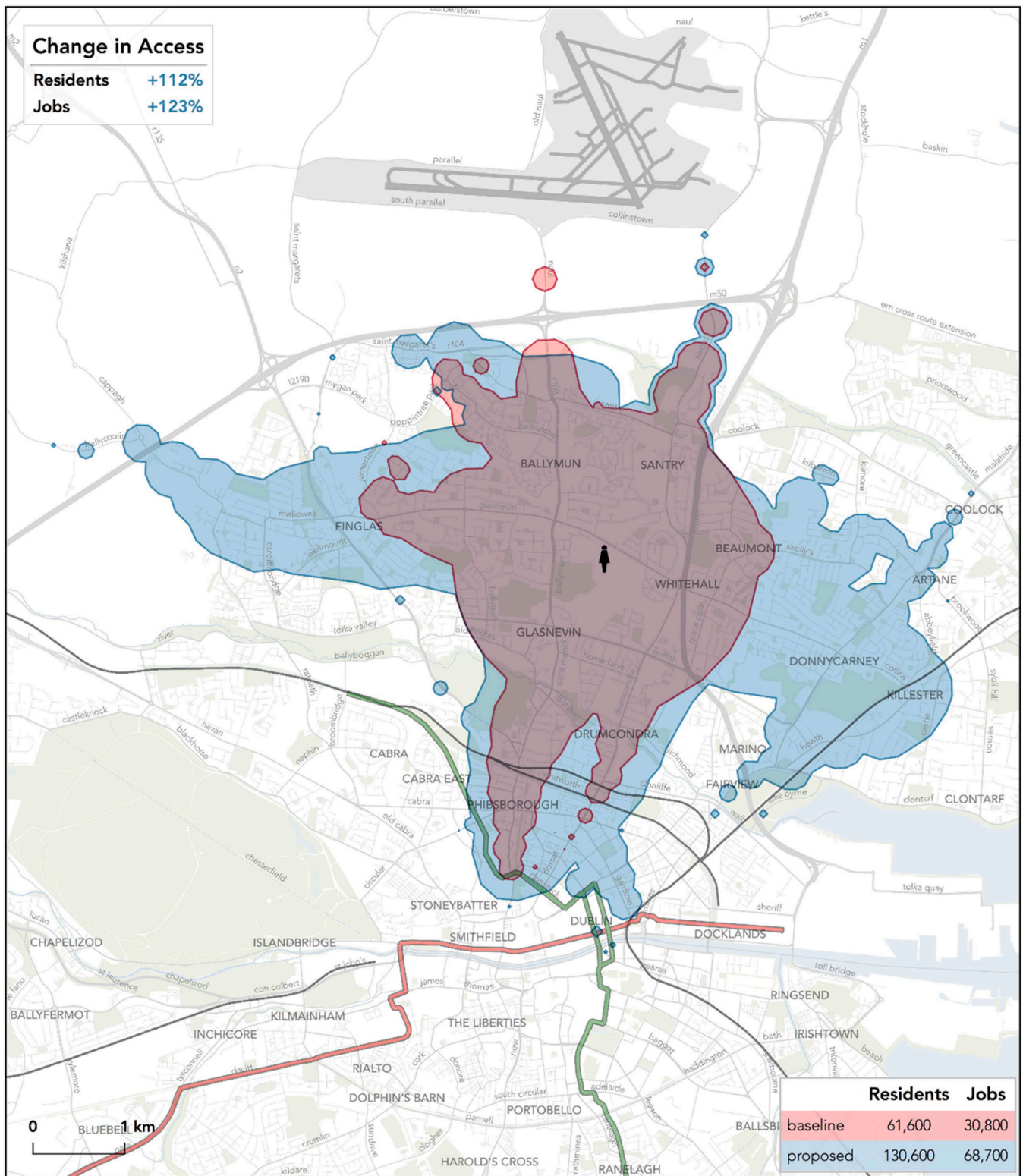


Fig. 2. Isochrone map showing gained (blue), lost (red), and retained (overlapping blue and red) 30-min cumulative opportunities accessibility from Dublin City University – Main Campus under the Dublin bus network redesign (Jarrett Walker + Associates, 2018, Appendix A1). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

estimate regional travel demand models. Even though individual travel patterns change from day to day, a random sample of individuals drawn for a single day will reproduce regional travel patterns in all their complexity.

Another argument against revealed behavior measures is that they do not consider foregone trips or travel demand that might be

suppressed (Lovejoy, 2012; Nordbakke and Schwanen, 2015; Pereira et al., 2017). This is also a fair point; given different circumstances, individuals facing transportation-related constraints would be likely to make different decisions. But this is where high-quality travel models can play an important role by supporting near-term analysis (either for the current year or one to five years from the present). The effects of

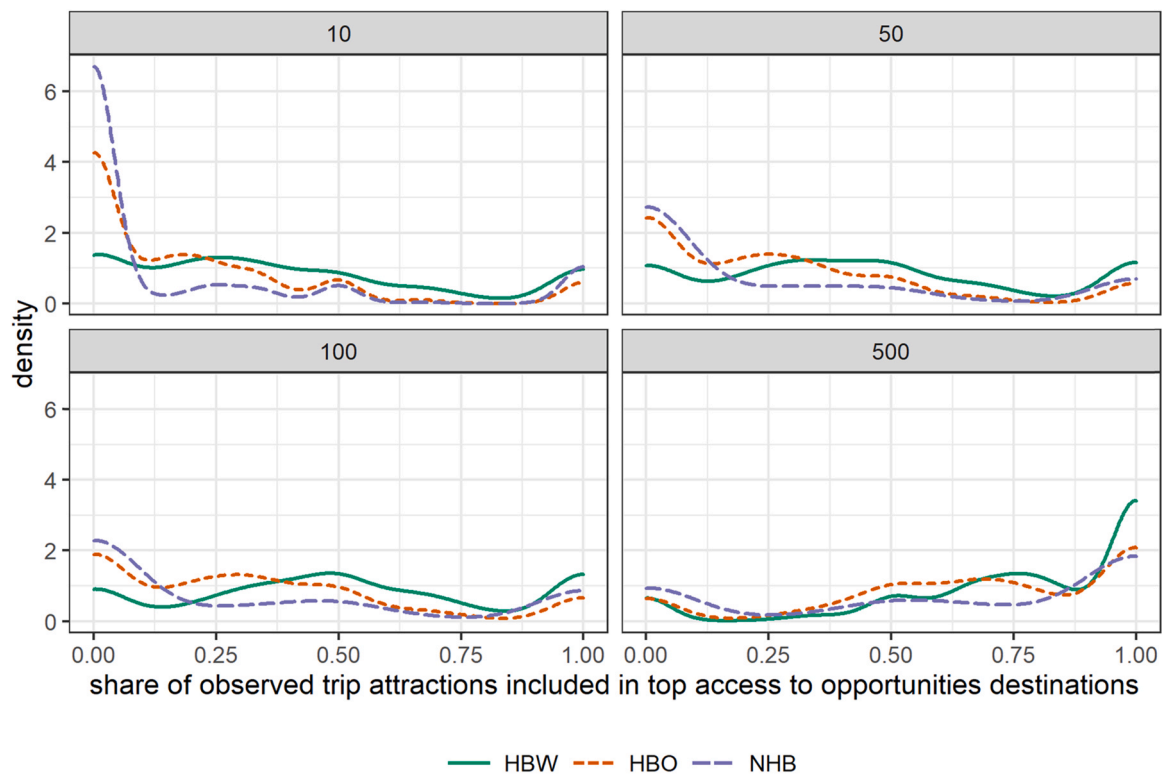


Fig. 3. Kernel density plots illustrating the share of observed trip attractions (taken from the transit rider survey) found in the top N destinations that compose the access to opportunities score. Each panel shows results for N values of 10, 50, 100, and 500. HBW = home-based-work, HBO = home-based-other, NHB = non-home-based.

changes in level of service on various modes or changes in socioeconomic conditions can be simulated. What if all low-income households were suddenly given cars? What if local bus service suddenly became more attractive? How would these changes shift travel patterns and reveal potentially suppressed demand? At the end of the day, individuals have to make decisions about where to travel and when.

In addition to these critiques, there are also real challenges associated with collecting revealed behavior data. Cost is a key factor. Obtaining high-quality, statistically controlled transit rider survey data can be expensive. A completed survey record can cost upwards of \$50 when costs associated with survey design, data collection, cleaning, and expansion are considered (Schaller, 2005). Other issues that plague transportation data collection, such as survey fatigue, are less relevant here since prospective survey respondents can easily be found on transit vehicles, as well as at stops and stations. Nevertheless, transit rider survey data collection methods continue to advance, and such data collection is mandated by FTA for large public transit agencies, as part of their Title VI program (FTA, 2012). Passively collected data harvested from cell phone and global positioning system traces can also be used to evaluate various aspects of public transit service and ridership (Welch and Widita, 2019), but collected data are not representative of the general population and demographic information is almost always missing (Chen et al., 2016; Wang et al., 2018). Smartphones also show some promise for collecting data traditionally gleaned through travel or activity diaries (Faghih Imani et al. 2020).

There are also two major benefits associated with people-focused near-term analyses. The first is that public and transportation equity advocates appear to be quite interested in how transportation infrastructure investments and policy choices will affect the traveling public given their current travel patterns (Marcantonio and Karner, 2014). From the perspective of these advocates, inequities already exist within the current system. These are reflected in racial and class differences in terms of mode share, level of service, exposure to pollution, noise, and other transport-related externalities (Karner et al. 2018). Without an

understanding of current inequities, the thinking goes, transport policy and infrastructure decision-making cannot be expected to mitigate them. The second is that near-term analyses mitigate many of the shortcomings known about long-range forecasting. The difficulty of projecting future land use characteristics and demographics, for example, is not at all relevant when conducting near-term analyses since those characteristics can reasonably be expected to be similar to those that prevail today. Further, a near-term forecast can supplement and inform long-range efforts by providing valuable comparative information about model uncertainty.

An empirical example

To demonstrate some of the fundamental differences between revealed behavior and access to opportunities measures, we calculated a performance measure using multiple data sources representing travel behavior and public transit level of service within the Metropolitan Transit Authority of Harris County (colloquially known as Houston METRO), roughly encompassing Harris County, Texas (USA) and the city of Houston. The performance measure assesses the correspondence between an origin's most important destinations (as revealed using a high-quality transit rider survey) and the destinations contributing most to an origin's access to opportunity score. Full details regarding data sources and methods are provided in the Appendix.

In concept, the performance measure calculation proceeded as follows. For each block group in Harris County, a cumulative opportunities measure was calculated. In this case, the measure was the total number of jobs accessible using public transit within 45 min, considering average performance across AM peak period, midday, and PM peak period travel times. For each of these three periods, four randomly selected departure times were used to estimate total travel time by public transit (including walking to the transit stop, waiting for the vehicle, in-vehicle travel time, any necessary transfers, and walking to the destination) to each reachable destination. This step resulted in

12 matrices representing travel time between all regional origins and destinations at different times of day. Job totals were associated with each destination block group. For each origin, at each time slice, the top N destinations contributing most to that origin's accessibility score were identified. For the purposes of demonstration, N was set to 10, 50, 100, and 500.

Each block group with a transit trip production as identified in the rider survey was associated with its attraction block groups. To be clear, while the cumulative opportunities measure identifies a destination as reachable if it is possible to travel there within 45 min on public transit, the rider survey data only identifies a production-attraction interchange where a trip was actually made.

Once these two lists were assembled—the top N destinations contributing to the cumulative opportunities score and all attraction trip ends visited from a production end—we calculated the share of observed trip attractions accounted for in the top access to opportunities destinations for three trip types: home-based-work, home-based-other, and non-home-based. Fig. 3 summarizes these results using kernel density plots for the shares of observed trip attractions found in the top N destinations contributing to an origin's accessibility score. Results are shown for four values of N (10, 50, 100, and 500). The area under the curve between any two points along the x-axis represents the proportion of all trip origins that have shares within that range.

The results shown in Fig. 3 achieve face validity. As N increases, shares tend toward one. This is sensible; as the list of destinations contributing to an accessibility score grows, the likelihood that actually chosen destinations will be included in the list also grows. In general, non-home-based and home-based-other trips see fewer of their attraction locations represented in the most important cumulative opportunities destinations. The density curves for these two trip types plot higher on the y-axis at zero, indicating that there are relatively higher numbers of block groups whose destinations are not reflected in the most important contributors to access to opportunities. This too appears reasonable, as home-based-work trips will tend toward the major employment destinations that contribute most to accessibility scores. Other trip purposes are more likely to be dispersed and not well reflected in the aggregate nature of a cumulative accessibility measure.

These results demonstrate a conceptual mismatch between the places to which people need to travel, and the locations identified as most important within a typical accessibility analysis. The purpose of this empirical example is not to discredit commonly used accessibility measures, but rather to highlight the value of an alternative approach grounded in real-world travel behavior, aligned with trips actually made.

Discussion and conclusions

Assessing transportation system performance under both base conditions and alternative scenarios has been of interest to the profession for decades. Yet not all performance analysis methods are created equal. Within public transit performance analysis, the literature and practice seem to be converging on the consensus that cumulative opportunities approaches are the preferred method both for communicating information to the public and determining whether certain investments are worth pursuing. However, the downsides of these analyses are almost never reported.

Importantly, the discussion presented in this paper should make

Appendix

The data used in the empirical example presented in this paper were gleaned from multiple sources, including:

- A production-attraction matrix from the 2017 transit rider survey conducted by the Houston-Galveston Area Council (H-GAC) and Houston METRO. These data were collected between February and May 2017, two years after Houston's comprehensive bus network redesign.
- Job location information from the US Census Bureau's 2017 Longitudinal Employer-Household Dynamics Origin-Destination Employment Statistics.

clear that simply increasing the accuracy, precision, or behavioral fidelity of a specific measure is not a desired end. Rather, our emphasis is on developing performance measures that are *meaningful*. One key dimension of meaning, as applied in this paper, is the extent to which a performance measure is linked to impacts on people and their travel. Whereas access to opportunities measures the extent to which people can in principle reach opportunities, such measures are substantially limited in their ability to reveal how—across an entire region—the benefits and burdens of a transportation system change are likely to be distributed or felt. Near-term analyses should also be preferred in that they produce results subject to much less uncertainty than longer term forecasts.

No performance measure or analytical approach will ever be completely accurate or provide absolutely definitive information about whether one course of action should be chosen over another. Transportation decision-making processes are intensely political exercises, involve the expenditure of substantial amounts of public funds, and create winners and losers. There is potentially a role for each of the analytical approaches described in this paper during different types of planning activities.

One additional limitation of a near-term analysis focused on people is that any proposed changes to public transit in a region could have real impacts on future-year origin destination travel patterns. Clearly, an analysis focused on current-year travel will not capture these. A future-year forecast could capture these types of changes, but access to opportunities proponents would argue that long-range ridership forecasting will always come with high uncertainties and questions about likely accuracy, so their approach is to be preferred (Walker, 2018). But even access to opportunities approaches build in uncertainty arising from how precisely zone-to-zone travel times will change, incomplete pedestrian networks, and potential demographic changes.

While it is indeed true that travel forecasting models have performed poorly in the past, this realization is an argument for improving travel models so that they are more relevant to the decision-making process, not ignoring the insights they provide. If long-range person-based analyses are performed, an analyst should also prepare a near-horizon or even current-year analysis, as this may help planners, analysts, and the public understand the uncertainties associated with long-range forecasting, by examining the differences between observed and predicted behavior.

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.

Acknowledgments

This work was supported by the Federal Transit Administration under cooperative agreement TX-2017-087-00. Note that the opinions expressed herein are those of the author; they are not endorsed by any government agency or third party. Ken Cervenka provided valuable insights and feedback on early drafts. Three anonymous reviewers provided helpful comments that improved the paper. Any remaining errors or omissions are the author's alone.

- Travel times between all census block group origins and destinations in Harris County, Texas, calculated using OpenTripPlanner. To represent travel on a typical weekday, we randomly sampled four departure times within 15-minute windows across three one-hour periods to represent the morning peak (8–9 am), midday (12–1 pm), and the evening peak (5–6 pm). This resulted in 12 separate estimates of travel times that can be combined to produce representative accessibility estimates.

We employed the cumulative opportunities accessibility measure described in Eqs. (A1)–(A3):

$$A_{ik} = \sum_{j=1}^J O_j f(c_{ijk}) \quad (\text{A1})$$

$$f(\bullet) = \begin{cases} 1 & \forall c_{ijk} \leq T \\ 0 & \forall c_{ijk} > T \end{cases} \quad (\text{A2})$$

$$A_i = \text{med}(A_{ik}) \quad \forall k \quad (\text{A3})$$

where A_{ik} is the accessibility score for block group i at departure time k ; O_j is the number of jobs at each destination block group j ; J is the total number of block groups; c_{ijk} is the travel time between block groups i and j at departure time k (considering walk access, waiting, in-vehicle, transfer, and walk egress times); T is the desired travel time threshold; and med is the median operator. For this application, T is set to 45 min.

This formulation of A_i is widely used in the literature and practice and is readily interpretable as the number of opportunities that can be reached from each origin location within the travel time threshold. Calculating accessibility first for each departure time, and then subsequently taking the median to represent an origin's score, is preferable to first averaging travel times between origins and destinations and then calculating accessibility. The former approach yields a more reasonable approximation of travel time between origins and destinations; locations with an infrequent but fast connection may generate only a single travel time within the set of 12 because an upper threshold on calculated travel times is typically set by the analyst to make the computations tractable (e.g., two hours). If travel times are simply averaged, origins and destinations with infrequent but fast connections will appear to have a higher level of service than is actually the case.

The final performance measure summarized in the empirical example requires identifying sets of destinations using both the transit rider survey data and the accessibility dataset. We used a block-group trip table in production-attraction format to mimic how access to opportunities is typically associated with residential locations. In this format, the location of an attraction (e.g., a person's workplace) would be associated with two trips from the person's production location (i.e., home). For home-based-work and home-based-other trips, this approach ensures identification of the most important block groups that people visit from their homes. For non-home-based trips, the production location is where the trip originates, and the attraction location is the trip destination. In all cases, the fully expanded, linked trip weights were used to tally trips, ensuring that the results represented a typical day of travel on Houston METRO in 2017.

An intermediate step in the accessibility calculation requires identifying all destinations reachable from each origin. Define E_{xN} as the set of N destinations contributing the most to a block group's accessibility score. Defining set E in this way is equivalent to sorting the list of destinations for each origin block group in descending order and taking the top N elements. For each production block group x , define D_x as the set of all attractions visited. With these definitions, the summary performance measure for evaluating the congruence between the people-focused measures and access to opportunities measures can be defined using Eq. (A4):

$$P_i = \frac{|D_i \cap E_{iN}|}{|D_i|} \quad (\text{A4})$$

where P_i is the share of revealed trip attractions that are accounted for in the top N destinations contributing to a block group's accessibility score. As N increases, P_i approaches unity.

References

- Allen, J., Farber, S., 2019. Sizing up transport poverty: a national scale accounting of low-income households suffering from inaccessibility in Canada, and what to do about it. *Transp. Policy* 74, 214–223.
- Brown, A.E., 2017. Car-less or car-free? Socioeconomic and mobility differences among zero-car households. *Transp. Policy* 60, 152–159.
- Chen, C., Ma, J., Susilo, Y., Liu, Y., Wang, M., 2016. The promises of big data and small data for travel behavior (aka human mobility) analysis. *Transp. Res. Part C Emerg. Technol.* 68, 285–299.
- Chorus, C.G., Timmermans, H.J.P., 2009. Measuring user benefits of changes in the transport system when traveler awareness is limited. *Transp. Res. Part A Policy Pract.* 43 (5), 536–547.
- de Jong, G., Daly, A., Pieters, M., van der Hoorn, T., 2007. The logsum as an evaluation measure: review of the literature and new results. *Transportation Research Part A: Policy and Practice, Selected Papers on Applications of Discrete Choice Models Presented at the European Regional Science Conference, Amsterdam, August 2005*, 41(9), pp. 874–889.
- Faghih Imani, A., Harding, C., Srikukenthiran, S., Miller, E.J., Habib, K.N., 2020. Lessons from a large-scale experiment on the use of smartphone apps to collect travel diary data: the 'city logger' for the greater Golden Horseshoe area. *Transp. Res. Rec.* 2674 (7), 299–311.
- Farber, S., Fu, L., 2017. Dynamic public transit accessibility using travel time cubes: comparing the effects of literature and new results. *Transportation Research Part A: Policy and Practice, Selected Papers on Applications of Discrete Choice Models Presented at the European Regional Science Conference, Amsterdam, August 2005*, 41(9), pp. 874–889.
- Farber, S., Ritter, B., Fu, L., 2016. Space-time mismatch between transit service and observed travel patterns in the Wasatch Front, Utah: a social equity perspective. *Travel Behav. Soc.* 4, 40–48.
- Flyvbjerg, B., Skamris Holm, M.K., Buhl, S.L., 2005. How (In)Accurate Are Demand Forecasts in Public Works Projects?: The Case of Transportation. *J. Am. Plan. Assoc.* 71 (2), 131–146.
- FTA (Federal Transit Administration), 2012. Circular FTA C 4702.1B: Title VI Requirements and Guidelines for Federal Transit Administration Recipients. US Department of Transportation, Washington, DC.
- FTA (Federal Transit Administration), 2016. Final Interim Policy Guidance Federal Transit Administration Capital Investment Grant Program. US Department of Transportation, Washington, DC.
- Garrett, M., Taylor, B.D., 1999. Reconsidering social equity in public transit. *Berkeley Plan. J.* 13, 6–27.
- 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.
- Golub, A., Martens, K., 2014. Using principles of justice to assess the modal equity of regional transportation plans. *J. Transp. Geogr.* 41, 10–20.
- Golub, A., Robinson, G., Nee, B., 2013. Making accessibility analyses accessible: a tool to facilitate the public review of the effects of regional transportation plans on accessibility. *J. Transp. Land Use* 6 (3), 17–28.
- Hamer, A.M., 1976. *The Selling of Rail Rapid Transit: A Critical Look at Urban Transportation Planning*. Lexington Books.
- Handy, S., 2008. Regional transportation planning in the US: an examination of changes in technical aspects of the planning process in response to changing goals. *Transp. Policy* 15 (2), 113–126.
- Handy, S., Niemeier, D., 1997. Measuring accessibility: an exploration of issues and alternatives. *Environ. Plan. A* 29 (7), 1175–1194.
- Howe, E., Kaufman, J., 1979. The ethics of contemporary american planners. *J. Am. Plan. Assoc.* 45 (3), 243–255.
- Jarrett Walker + Associates, 2018. Dublin Area Bus Network Redesign: Public Consultation Report. National Transport Authority, Dublin, Ireland.

- Kain, J.F., 1990. Deception in Dallas: strategic misrepresentation in rail transit promotion and evaluation. *J. Am. Plan. Assoc.* 56 (2), 184–196.
- Karner, A., 2018. Assessing public transit service equity using route-level accessibility measures and public data. *J. Trans. Geogr.* 67, 24–32.
- Karner, A., Golub, A., 2015. Comparison of two common approaches to public transit service equity evaluation. *Transp. Res. Rec. J. Transp. Res. Board* 2531, 170–179.
- Karner, A., Golub, A., Martens, K., Robinson, G., 2018. Transportation and environmental justice: history and emerging practice. In: Holifield, Ryan, Chakraborty, Jayajit, Walker, Gordon (Eds.), *The Routledge Handbook of Environmental Justice*. Routledge, New York, pp. 400–411.
- Klosterman, R.E., 2013. Lessons learned about planning. *J. Am. Plan. Assoc.* 79 (2), 161–169.
- Levine, J., Grengs, J., Merlin, L.A., 2019. *From Mobility to Accessibility: Transforming Urban Transportation and Land-Use Planning*. Cornell University Press, Ithaca, NY.
- Levinson, D., 2013. *Access Across America*. Department of Civil Engineering, University of Minnesota, Minneapolis, MN.
- Lovejoy, K., 2012. *Mobility Fulfillment among Low-Car Households: Implications for Reducing Auto Dependence in the United States* (Ph.D. dissertation). University of California, Davis.
- Manaugh, K., El-Geneidy, A., 2012. Who benefits from new transportation infrastructure? Using accessibility measures to evaluate social equity in public transport provision. In: Geurs, Karst T., Krizek, Kevin K., Reggiani, Aura (Eds.), *Accessibility Analysis and Transport Planning*. Edward Elgar, Cheltenham, UK.
- Marcantonio, R., Karner, A., 2014. Disadvantaged communities teach regional planners a lesson in equitable and sustainable development. *Poverty Race* 23 (1), 5–12.
- Martens, K., 2012. Justice in transport as justice in accessibility: applying walzer's 'spheres of justice' to the transport sector. *Transportation* 39 (6), 1035–1053. <<http://link.springer.com/journal/volumesAndIssues/11116>>.
- Martens, K., 2016. *Transport Justice: Designing Fair Transportation Systems*. Routledge, New York.
- Martens, K., Di Ciommo, F., 2017. Travel time savings, accessibility gains and equity effects in cost-benefit analysis. *Transp. Rev.* 37 (2), 152–169.
- Martens, K., Golub, A., 2018. A fair distribution of accessibility: interpreting civil rights regulations for regional transportation plans. *J. Plan. Educ. Res.*, 0739456X1879101.
- Niemeier, D.A., 1997. Accessibility: an evaluation using consumer welfare. *Transportation* 24 (4), 377–396.
- Nordbakke, S., Schwanen, T., 2015. Transport, unmet activity needs and wellbeing in later life: exploring the links. *Transportation* 42 (6), 1129–1151.
- Owen, A., Levinson, D., 2014. *Access Across America: Transit 2014*. Accessibility Observatory, University of Minnesota, Minneapolis, MN.
- Páez, A., Scott, D.M., Morency, C., 2012. Measuring accessibility: positive and normative implementations of various accessibility indicators. *J. Transp. Geogr.* 25, 141–153.
- Pereira, R.H.M., 2019. Future accessibility impacts of transport policy scenarios: equity and sensitivity to travel time thresholds for bus rapid transit expansion in Rio de Janeiro. *J. Transp. Geogr.* 74, 321–332.
- Pereira, R.H.M., Schwanen, T., Banister, D., 2017. Distributive justice and equity in transportation. *Transp. Rev.* 37 (2), 170–191.
- Pickrell, D.H., 1992. A desire named streetcar. *J. Am. Plan. Assoc.* 58 (2), 158–176.
- Puget Sound Regional Council, 2018. *The Regional Transportation Plan*. <<https://www.psrc.org/sites/default/files/rtp-may2018.pdf>>.
- Rowangould, D., Karner, A., London, J., 2016. Identifying environmental justice communities for transportation analysis. *Transp. Res. Part A Policy Pract.* 88, 151–162.
- Schaller, B., 2005. TCRP Synthesis 63: On-Board and Intercept Transit Survey Techniques. Transit Cooperative Research Program, Washington, DC.
- Smith, S.K., Tayman, J., Swanson, D.A., 2013. *A Practitioner's Guide to State and Local Population Projections*. Springer, New York.
- Stewart, A.F., 2017. Mapping transit accessibility: possibilities for public participation. *Transp. Res. Part A Policy Pract.* 104 (Suppl. C), S150–S166.
- Taylor, B.D., Morris, E.A., 2015. Public transportation objectives and rider demographics: are transit's priorities poor public policy? *Transportation* 42 (2), 347–367.
- Virginia Department of Transportation, 2020. *Smart Scale Technical Guide*. <<http://www.vasmartscale.org/documents/2020documents/rd4-ss-technical-guide-draft-final-06012020.pdf>>.
- Voulgaris, C.T., 2019. Crystal balls and black boxes: what makes a good forecast? *J. Plan. Lit.* 34 (3), 286–299.
- Voulgaris, C.T., 2020a. What is a forecast for? Motivations for transit ridership forecast accuracy in the federal new starts program. *J. Am. Plan. Assoc.* 86 (4), 458–469.
- Voulgaris, C.T., 2020b. Trust in forecasts? Correlates with ridership forecast accuracy for fixed-guideway transit projects. *Transportation* 47, 2439–2477.
- Wachs, M., 1989. When planners lie with numbers. *J. Am. Plan. Assoc.* 55 (4), 476.
- Wachs, M., 2001. Forecasting versus envisioning: a new window on the future. *J. Am. Plan. Assoc.* 67 (4), 367–372.
- Walker, J., 2018. To predict with confidence, plan for freedom. *J. Public Transp.* 21 (1), 119–127.
- Wang, Z., He, S.Y., Leung, Y., 2018. Applying mobile phone data to travel behaviour research: a literature review. *Travel Behav. Soc.* 11, 141–155.
- Welch, T.F., Widita, A., 2019. Big data in public transportation: a review of sources and methods. *Transp. Rev.* 39 (6), 795–818.

Alex Karner is an associate professor in the graduate program in community and regional planning at the University of Texas at Austin. He conducts research on transportation equity and accessibility from multiple disciplinary and methodological perspectives.