abm_illustration

Hayden Atchley   Gregory S. Macfarlane

14 April 2023

Table of contents

List of Figures

3

List of Tables

index

1 Introduction

This is a book created from markdown and executable code.

2 Literature Review

Travel demand modeling in the modern sense has its origins in the 1950s, with the Chicago Area Transportation Study (Chicago Area Transportation Study 1959) being one of the first urban planning studies to use the now-ubiquitous "four-step" modeling framework (McNally 2007). Up until this point, most urban transportation planning used existing demand or uniform-growth travel forecasts to model travel demand, but the Chicago Study used a combination of trip generation, trip distribution, modal split, and network assignment models to more accurately represent travel behavior (Weiner 1997). Since then, there have been numerous studies iterating on the "four-step" (more appropriately termed "trip-based") framework, and models of this type are now used in forecasting travel demand across the United States (Park et al. 2020).

These trip-based models are not without problems, however. Rasouli and Timmermans (2014) give several shortcomings of trip-based models: they use several sub-models that are (implicitly or explicitly) assumed independent, and this can result in a lack of consistency or integrity between sub-models; they are strongly aggregated in nature, which can cause significant aggregation bias; and they lack "behavioral realism"—that is, they do not have a concept of individuals making decisions, which is what travel behavior actually is.

Jones (1979) proposed an alternative to the trip-based paradigm, namely an "activity-based" framework that models daily activity patterns at an individual rather than aggregate level. Activity-based models (ABMs) place the focus on "activities" rather than "trips" as the basic unit of analysis, and predict a sequence of activities for each individual and household, with information such as activity location, start time, and duration, using a high level of temporal and spatial granularity. "Trips" are then the result of getting from one activity to the next (Pinjari and Bhat 2011). By adopting this activity-centric framework, ABMs provide a more consistent and comprehensive representation of travel behavior. They take into account complex dependencies and interactions within the model as a whole and at an individual level. ABMs acknowledge that travel choices are not made in isolation, but rather influenced by the preceding activities. This means that e.g. if an individual takes transit to work, they will not be able to drive home. ABMs therefore present a more conceptually accurate model of actual travel behavior than traditional trip-based models.

Despite these advantages, however, many agencies have yet to adopt ABMs, instead continuing to use trip-based models (Miller 2023). While ABMs are superior in certain aspects, they also have disadvantages, such as requiring more detailed input data and greater computational resources. It is also not always clear if ABMs provide substantially better forecasts than

their trip-based counterparts, nor if this tradeoff is worth it for every agency. This literature review presents an overview of both modeling frameworks, and discusses the advantages and disadvantages of using an ABM.

2.1 Overview of Model Types

Trip-based models are often referred to as "four-step" models due to their four fundamental sub-models: trip generation, trip distribution, mode choice, and network assignment (National Academies 2012, 28). Trip-based models are often more complicated than this, possibly including integration with a land use forecast, iteration between mode and destination choice, etc., but the "four steps" are the central component of any of these models (McNally 2007).

In a typical trip-based model, travel demand is predicted based on aggregate population data, often delineated by transportation analysis zones (TAZs). Each sub-model relies on this aggregate data; for example, the modal split sub-model will often use average TAZ income as an input (National Academies 2012, 14). The output of the first three of the four steps is a set of trip matrices, usually separated by mode and possibly time of day. These matrices list the number of trips between each TAZ, and these trips are then assigned to a transportation network (there are various algorithms for doing so, see Ortúzar and Willumsen (2011)).

Activity-based models differ significantly from this approach. Rather than using aggregate data, ABMs use an actual or synthetic population, with individual person and household data (Vovsha, Bradley, and Bowman 2005). These models use an activity or tour scheduler to assign a daily activity pattern (DAP) of zero or more tours to each individual (*n.b.* a tour is a series of trips that begin and end at home). These DAPs are restricted temporally, spatially, and modally; i.e., each person has a logical and followable sequence of trips and activities (Bowman 1998). A "drive alone" trip from work to lunch, for example, cannot be made if transit was taken to work. ABMs output a list of tours and trips by person, time, location, and type, and these can then be assigned to a transportation network in a similar manner as in a trip-based model.

2.2 Comparison of Modeling Frameworks

In discussing the differences between ABMs and trip-based models, there are really two comparisons that need to be made: how the population data is structured, and how travel is organized. Trip-based models generally use aggregate population data while ABMs use a synthetic population, and trip-based models organize travel into trips while ABMs organize travel into activities and tours. The following sections will explain these different aspects of travel demand modeling and discuss the claimed advantages and disadvantages of each.

2.2.1 Population Data

The aggregate population data used in trip-based models can be varied in origin and level of detail, but the basic concept is the same: the study area is organized into generally small zones, and certain demographic and socioeconomic data is known or obtained for each zone (National Academies 2012, 14). This includes data such as number of households, average household income, population, number of workers, etc. Based on this information, the zone can be segmented along arbitrary variables. For example, since households with more workers and more vehicles tend to make more work trips, it is useful to estimate the distribution of households in each zone along these two variables. Then, average trip production rates are determined for each household category (e.g. for households with each number of workers by vehicles), and the total number of trips produced in a zone is calculated based on these rates (National Academies 2012, 37).

This approach is relatively straightforward: the required input data is usually easy to obtain, the trip generation models are often simple, and it is computationally inexpensive (National Academies 2012). However, any desired grouping for analysis must be a part of this segmentation in each model step. For something like an equity analysis, for example, it is impossible to determine the "winners" and "losers" of a potential policy without using race/gender/etc. as variables in the trip generation and destination and model choice steps (Bills and Walker 2017). This is problematic; while there is data on trip rates, mode choice, etc. based on these variables, if this is coded into the model then the model asserts that these trends will hold no matter the policy implementations.

As an added complication to the difficulty of these analyses, the aggregate population data must from the start be segmented along each analysis variable. An analysis based on parents'/adults' highest received education, for example, would require determining the number of households in each TAZ with each possible combination of education level. This can be done, but more detailed and varied analyses would require more segmentation, greatly increasing the number of classifications needed. Since these variables need to be carried through each model step, trip rates etc. need to be estimated for every classification, and while real-world data may exist, sample sizes approach zero very quickly, and so the estimates have little statistical value (Moeckel et al. 2020; National Academies 2012).

An alternative approach to population data is to use a full synthetic population. A synthetic population takes demographic and socioeconomic data at various levels of detail to create a "population" with generally the same distribution as the study area (National Academies 2012, 93). The goal is to have a population that is functionally similar to the actual population, but without the privacy concerns of using real individual household data. Castiglione et al. (2006) argue that the major advantage with this approach is that the demographic and socioeconomic data is known at the person and household level, rather than the zone level, and this data remains available throughout the modeling process. This allows, for example, an equity analysis to determine the "winners" and "losers" of a proposed development without needing to encode race/gender/etc. into each step of the model.

Bills and Walker (2017) for example used the 2000 Bay Area Travel Survey to create a synthetic population and compare the effects that certain scenarios had on high income and low income populations. With a 20% reduction in travel cost, they saw that high income workers benefited more than low income workers. They did similar comparisons for scenarios involving reduced travel times for different mode choices and were able to see the effects each scenario had on the high and low income workers. These types of analysis, which are not possible with aggregate population data, can be very valuable in transportation planning and policy making.

It is important to note that, while many connect them only with ABMs, synthetic populations can be used in running trip-based models as well. Trip-based models using a synthetic population—often called trip-based microsimulation models—do exist (see Walker (2005) and Moeckel et al. (2020)), but these are relatively rare. If a company is willing to go through the initial work of creating a synthetic population, then they may just use an ABM because they believe it will be most accurate.

Figure 2.1 gives a visualization of an example "information pipeline" for a trip-based model using aggregate data and one using a synthetic population. In the aggregate data model, it is impossible to know which trips are made by e.g. 2-worker, 1-vehicle, low-income households, only which trips are made by households with fewer vehicles than workers. With a synthetic population, however, *individuals* are being modeled, and so each trip can be traced to a specific person. All information is known at each point in the model regardless of which data is used in previous steps.
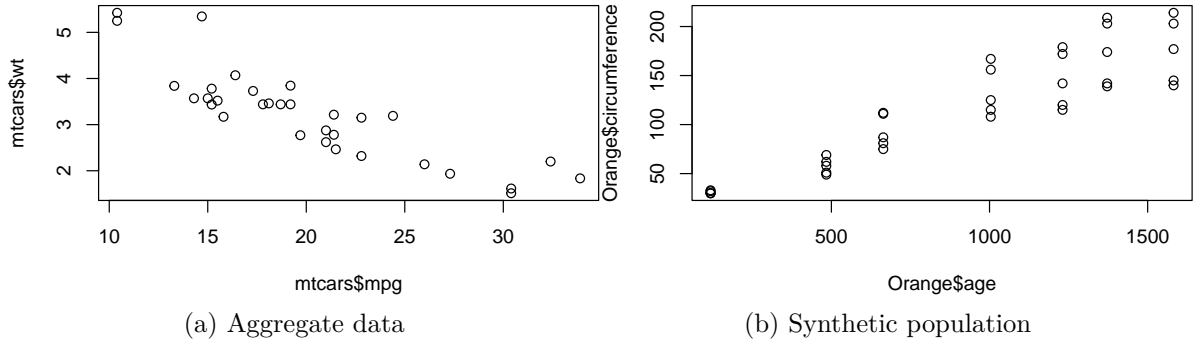

(a) Aggregate data   (b) Synthetic population

Figure 2.1: Example information pipelines for two theoretical trip-based models, one using aggregate data and another using a synthetic population.

2.2.2 Travel Behavior

The other significant difference from trip-based models ABMs have (and the main difference from trip-based microsimulation models) is that ABMs organize travel into "tours", a sequence of trips that begin and end at the home, rather than just trips. In fact, Miller (2023) points out that many current "activity-based" models ought to rather be labelled "tour-based" due to this focus on building tours. This is contrasted with "activity scheduling" models, in which

activity participation is modeled explicitly and trips emerge as the means to get from one activity to the next. However, in practice there are few true "activity scheduling" models, and the term "activity-based" is used to refer to both these and "tour-based" models.

In a typical trip-based model, trips are forecasted based on empirical trip rates, usually by trip purpose and by household type (for example, low-income, 1-vehicle households make a certain number of "home-based work" trips) (McNally 2007). These trips are then assigned an origin and destination, mode, and often a time of day (peak/off-peak, etc.), and the result is a list of trips between each zone by mode and purpose. A trip-based microsimulation model may use regression models rather than aggregate data for some of the model steps (Moeckel et al. 2020), but the end result is similar: a list of trips by person, noting mode and purpose. However, this trip list may be inconsistent, and the forecasted trips may not be physically possible to complete in any sequence, as there is no sense of "trip-chaining". The hope, though, is that over an area the inconsistencies would cancel out, leaving an overall accurate forecast.

ABMs, on the other hand, model *tours* rather than trips. This allows for consistency in trip origins/destinations, mode choice, and time of day: since each trip is a part of a "chain" (tour), the trips within a tour are dependent on each other (Rasouli and Timmermans 2014). The open-source ABM ActivitySim (Association of Metropolitan Planning Organizations 2023), for example, has a tour-scheduling model that determines the number of "mandatory" (work, school, etc.) and "discretionary" tours each individual will make, and performs tour-level mode and destination choice for each tour. After the tour-level decisions are made, trip-level mode/destination choice is done for each trip in the tour, including the possible addition of subtours (see Vovsha, Bradley, and Bowman (2005), fig. 18.1).

Figure 2.2 shows a visual representation of an individual's travel in both a trip-based microsimulation model and an ABM (tour-based). Though each individual's trips are known, in a trip-based model there is no sense of consistency between trips. The activity-based approach adds this consistency by modeling tours, and only generating trips consistent with each tour.

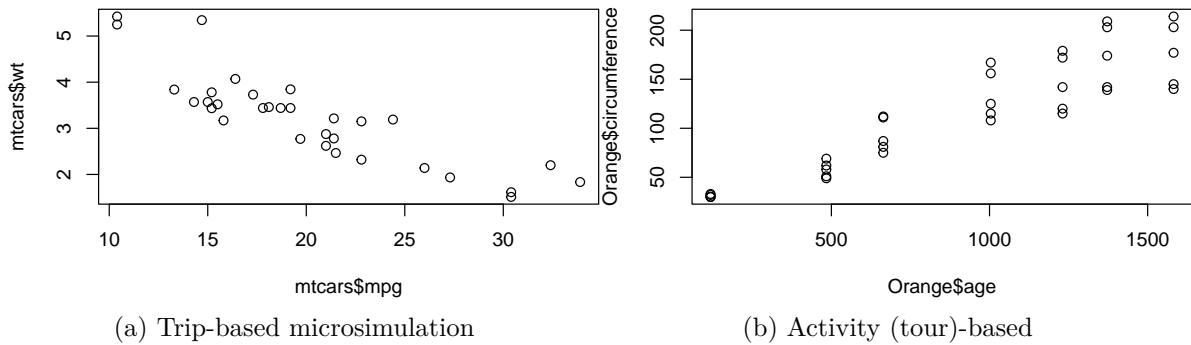(a) Trip-based microsimulation   (b) Activity (tour)-based

Figure 2.2: Example travel pattern for an individual in both a trip-based microsimulation model and an ABM.

In addition to intra-person dependencies, Rasouli and Timmermans (2014) note that ABMs

can model dependencies between members of a household as well. A vehicle can't be used by multiple people in the same household at the same time to travel to different destinations. Because the people within the household will have travel patterns that depend on the patterns of others in the household, a policy affecting one person in the household can affect everyone in the household no matter how directly the policy connects to them (Vovsha, Bradley, and Bowman 2005). These effects aren't possible to forecast in a trip-based model.

Another advantage of organizing travel into tours comes regarding accessibility analyses. Dong et al. (2006) note that when trip-based models are used to analyze accessibility, each zone must be analyzed independently of travel behavior. This approach only analyzes zones' proximity to each other and does not take into account individual travel patterns. They argue that this is a limited view of accessibility, and discuss the "activity-based accessibility measure" (ABA), which is evaluated based on all trips in a day rather than particular trips. As an example, if an individual doesn't live within a 20-minute drive of a grocery store, traditional measures might rate this as poor accessibility. However, if they travel to a park with a nearby grocery store as part of their DAP, then in reality the accessibility should be rated much higher. This proximity may even influence *which* park is chosen. Overall, they found that the ABA predicts more reasonable accessibility outcomes compared to traditional measures.

2.3 Lack of ABM Adoption

Though ABMs have many clear advantages over trip-based models, adoption among agencies has been relatively slow. Many ABMs are implemented in proprietary software, which creates difficulty in maintaining and iterating on the model, Miller (2023) argues. Even in an open-source model like ActivitySim (Association of Metropolitan Planning Organizations 2023), Miller notes several disadvantages of ABMs:

- Computational inefficiency and complicated program design: ABMs take more time, more computing power, and more money to run. This is because the synthetic population needed to run an ABM uses much more data. In areas with thousands of TAZs and millions of people, a super computer is needed, and it will cost much more than what is spent to run trip based models. If a city can see similar results using a trip based model, they may decide not to invest in an ABM.

- Absence of a standard model system: The modeling systems are often designed with different approaches and for specific areas making it hard to transfer from one urban area to another. This also makes it difficult for agencies to determine which approach is the best and decide which to implement. In relation to this, Miller also states that the pressures of publishing unique and ground-breaking research in academia can deter researchers from converging towards best theories and methods.

- Lack of resources: Most of these models were developed in academic settings which often lack resources, and possibly desire, to put them into practice. This leaves it up to

governments and consultants to put the models into practice, but they can be hesitant to promote software development and to invest in new systems.

Because of these reasons, as well as the inertia of current practices, trip-based models remain in use by the majority of agencies and organizations in the US.

2.4 Research Gap

Though there has been much research on ABMs and their theoretical advantages, to the authors' knowledge there is little in the literature about their practicality in the real world. It is often taken as given that ABMs are unilaterally superior to traditional trip-based models due to their better theoretical foundation, but it is not clear if that better foundation always yields better results. Ferdous et al. (2012) compared the trip- and activity-based model frameworks of the Mid-Ohio Regional Planning Commission and found that the ABM performed slightly better at the region level, but about the same at the project level. This is not an overwhelming victory for ABMs, and so based on this an agency may reasonably decide that the increase in data, computational, training, and potentially financial requirements outweigh the potential gains of switching to an ABM.

Ferdous et al. (2012), however, mainly compared the *accuracy* of the two frameworks, but did not address the methodological differences between them. What types of data collection/synthesis are needed for each model type? Are there analyses that can only be done through (or that are made easier by) one of the model types? What would an agency need in order to transition from a trip-based model to an ABM? Are certain types of scenarios suited to one model type? Though some of these questions have been discussed (see e.g. Lemp, McWethy, and Kockelman 2007), a holistic methodological comparison is lacking. Additionally, the answers in the current literature are mainly theoretical, with little use to an agency considering the transition.

This research aims to answer these questions by providing a side-by-side comparison of a potential trip-based and activity-based modeling methodology. Several "proposed development" scenarios are run in each model, and the strengths and weaknesses of each approach are compared. It is important to note that this paper is not focused on model accuracy, as in any model type this can be adjusted dramatically through calibration efforts. Rather, the focus is on the methodological differences between the approaches, and the types of analyses that can be done with each model type.

References

Association of Metropolitan Planning Organizations. 2023. "ActivitySim." *ActivitySim.* https://activitysim.github.io/.

Bills, Tierra S., and Joan L. Walker. 2017. "Looking Beyond the Mean for Equity Analysis: Examining Distributional Impacts of Transportation Improvements." *Transport Policy* 54 (February): 61–69. https://doi.org/10.1016/j.tranpol.2016.08.003.

Bowman, John L. 1998. "The Day Activity Schedule Approach to Travel Demand Analysis." Ph.D. Dissertation, Cambridge, MA: Massachusetts Institute of Technology. https://dspace.mit.edu/bitstream/handle/1721.1/16731/44704881-MIT.pdf?sequence=2.

Castiglione, Joe, Rachel Hiatt, Tilly Chang, and Billy Charlton. 2006. "Application of Travel Demand Microsimulation Model for Equity Analysis." *Transportation Research Record: Journal of the Transportation Research Board* 1977 (1): 35–42. https://doi.org/10.1177/0361198106197700105.

Chicago Area Transportation Study. 1959. *Final Report In Three Parts.* Vol. 1–3. Printed by authority of the State of Illinois. http://archive.org/details/chicagoareatrans01chic.

Dong, Xiaojing, Moshe E. Ben-Akiva, John L. Bowman, and Joan L. Walker. 2006. "Moving from Trip-Based to Activity-Based Measures of Accessibility." *Transportation Research Part A: Policy and Practice* 40 (2): 163–80. https://doi.org/10.1016/j.tra.2005.05.002.

Ferdous, Nazneen, Lakshmi Vana, John L. Bowman, Ram M. Pendyala, Gregory Giaimo, Chandra R. Bhat, David Schmitt, Mark Bradley, and Rebekah Anderson. 2012. "Comparison of Four-Step Versus Tour-Based Models for Prediction of Travel Behavior Before and After Transportation System Changes." *Transportation Research Record* 2303 (1): 46–60. https://doi.org/10.3141/2303-06.

Jones, Peter M. 1979. "New Approaches to Understanding Travel Behaviour: The Human Activity Approach." In *Behavioural Travel Modelling.* Routledge.

Lemp, Jason D., Laura B. McWethy, and Kara M. Kockelman. 2007. "From Aggregate Methods to Microsimulation: Assessing Benefits of Microscopic Activity-Based Models of Travel Demand." *Transportation Research Record: Journal of the Transportation Research Board* 1994 (1): 80–88. https://doi.org/10.3141/1994-11.

McNally, Michael G. 2007. "The Four-Step Model." In *Handbook of Transport Modelling*, edited by David A. Hensher and Kenneth J. Button, 1:35–53. Emerald Group Publishing Limited. https://doi.org/10.1108/9780857245670-003.

Miller, Eric. 2023. "The Current State of Activity-Based Travel Demand Modelling and Some Possible Next Steps." *Transport Reviews* 0 (0): 1–6. https://doi.org/10.1080/01441647.2023.2198458.

Moeckel, Rolf, Nico Kuehnel, Carlos Llorca, Ana Tsui Moreno, and Hema Rayaprolu. 2020.

"Agent-Based Simulation to Improve Policy Sensitivity of Trip-Based Models." *Journal of Advanced Transportation* 2020 (February): e1902162. https://doi.org/10.1155/2020/1902162.

National Academies. 2012. *Travel Demand Forecasting: Parameters and Techniques*. Washington, D.C.: National Academies Press. https://doi.org/10.17226/14665.

Ortúzar, Juan De Dios, and Luis G. Willumsen. 2011. *Modelling Transport*. 1st ed. Wiley. https://doi.org/10.1002/9781119993308.

Park, Keunhyun, Sadegh Sabouri, Torrey Lyons, Guang Tian, and Reid Ewing. 2020. "Intrazonal or Interzonal? Improving Intrazonal Travel Forecast in a Four-Step Travel Demand Model." *Transportation* 47 (5): 2087–2108. https://doi.org/10.1007/s11116-019-10002-0.

Pinjari, Abdul Rawoof, and Chandra R. Bhat. 2011. "Activity-Based Travel Demand Analysis." In *A Handbook of Transport Economics*, edited by André de Palma, Robin Lindsey, Emile Quinet, and Roger Vickerman. Edward Elgar Publishing.

Rasouli, Soora, and Harry Timmermans. 2014. "Activity-Based Models of Travel Demand: Promises, Progress and Prospects." *International Journal of Urban Sciences* 18 (1): 31–60. https://doi.org/10.1080/12265934.2013.835118.

Vovsha, Peter, Mark Bradley, and John L. Bowman. 2005. "Activity-Based Travel Forecasting Models in the United States: Progress Since 1995 and Prospects for the Future." In *Progress In Activity-Based Analysis*, edited by Harry Timmermans, 389–415. https://trid.trb.org/view/759300.

Walker, Joan L. 2005. "Making Household Microsimulation of Travel and Activities Accessible to Planners." *Transportation Research Record: Journal of the Transportation Research Board* 1931 (1): 38–48. https://doi.org/10.1177/0361198105193100105.

Weiner, Edward. 1997. "Urban Transportation Planning in the United States. An Historical Overview." DOT-T-97-24. https://doi.org/10.21949/1526591.