Evaluating the systemic effects of automated mobility-on-demand services via large-scale agent-based simulation of auto-dependent prototype cities

Citation: Jimi B. Oke, Arun Prakash Akkinepally, Siyu Chen, Yifei Xie, Youssef M. Aboutaleb, Carlos Lima Azevedo, P. Christopher Zegras, Joseph Ferreira, Moshe Ben-Akiva, Evaluating the systemic effects of automated mobility-on-demand services via large-scale agent-based simulation of auto-dependent prototype cities, Transportation Research Part A: Policy and Practice, Volume 140, 2020, Pages 98-126 © 2020 Elsevier Ltd

As Published: 10.1016/J.TRA.2020.06.013

Publisher: Elsevier BV

Evaluating the systemic effects of automated mobility-on-demand services via large-scale agent-based simulation of auto-dependent prototype cities

Jimi B. Oke[a,b], Arun Prakash Akkinepally[b], Siyu Chen[b], Yifei Xie[b], Youssef M. Aboutaleb[b], Carlos Lima Azevedo[c], P. Christopher Zegras[d], Joseph Ferreira[d], Moshe Ben-Akiva[b]

[a] *Department of Civil and Environmental Engineering, University of Massachusetts, Amherst, MA 01003, United States*

[b] *Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, United States*

[c] *Department of Management Engineering, Technical University of Denmark, Kgs. Lyngby, 2800, Denmark*

[d] *Department of Urban Studies and Planing, Massachusetts Institute of Technology, Cambridge, MA 02139, United States*

Abstract

As demand for urban mobility continues to grow and given that over 60% of the world's population is expected to be urban by 2050, increasingly innovative solutions must be devised to adequately meet the transportation needs of global metropolitan areas. Our objective is to therefore develop and implement a framework to analyze the systemic impacts of future mobility trends and policies. First, we build on prior work in classifying the world's cities into 12 urban typologies that represent distinct land-use and behavioral characteristics. Second, we introduce a generalized approach for a creating a detailed, simulatable prototype city that is representative of a given typology. Using this, we build and simulate two auto-dependent (largely US-specific) prototype cities via a state-of-the-art agent-based platform, SimMobility, for integrated demand microsimulation and supply mesoscopic simulation. We demonstrate the framework by analyzing the impacts of automated mobility on-demand (AMoD) implementation strategies in the cities based on demand, congestion, energy consumption and emissions outcomes. Our results show that the introduction of AMoD cannibalizes mass transit while increasing vehicle kilometers traveled (VKT) and congestion. In sprawling auto-dependent cities with low transit penetration, the congestion and energy consumption effects under best-case assumptions are similar regardless of whether AMoD competes with or complements mass transit. In dense auto-dependent cities with moderate transit modeshare, the integration of AMoD with transit yields better outcomes in terms of VKT and congestion. Such cities cannot afford to disinvest in mass transit, as this would result in unsustainable outcomes. Overall, this framework can provide insights into how AMoD can be sustainably harnessed not only in low-density and high-density auto-dependent cities, but also in other typologies.

Keywords: automated mobility-on-demand, agent-based simulation, future urban mobility, urban typologies, prototype cities

---

1 Introduction

As demand for urban mobility continues to grow, and given that over 60% of the world's population is expected to be urban by 2050 (United Nations, Department of Economic and Social Affairs, Population Division, 2018), increasingly innovative solutions must be devised to adequately meet the transportation needs of global metropolitan areas. The implementation and provision of these services, however, are significantly constrained by energy, environmental and financial requirements and regulations. On the one hand, private car ownership remains on an upward trajectory, globally (Sperling and Gordon, 2008; Dargay et al., 2007).[1] In 2016, passenger cars alone contributed about 12% (772 $MtCO_2e$)[2] of the United States' GHG emissions (EPA, 2018), while in Europe, passenger cars similarly contributed 11% (494 $MtCO_2e$) total GHG emissions in the same year (Transport & Environment, 2018). On the other hand, the demand for ridesourcing[3] services has been rapidly increasing over the past several years (Jin et al., 2018; Shaheen and Cohen, 2019). The long-term impacts of this trend are yet to be determined. However, it has been generally observed that the growth of on-demand passenger mobility is likely partially responsible for the decline in mass transit ridership in many cities, both in the US and elsewhere, along with impacts on congestion that are yet to be systematically quantified (Clewlow and Mishra, 2017). To both mitigate the deleterious effects of these trends and realize their maximum potential benefit, cities require integrated mobility solutions driven by smart analytical frameworks (Shaheen and Chan, 2016). The insights required to generate these solutions depend on large-scale, representative and detailed simulations of the urban environment. However, given the sheer differences in demographics (population), land-use, demand, supply and supply-demand interaction characteristics across cities worldwide, these innovations must be tailored to each urban area.

Our objective is to therefore understand and analyze globally relevant impacts of future trends and policies on urban mobility outcomes. First, we build on prior work in classifying the world's cities into 12 urban typologies which represent distinct land-use and behavioral characteristics (Oke et al., 2019). This classification used the most recent mobility, economic, demographic, land-use and behavioral data obtained from 331 cities. Second, we propose a generalized approach for a creating detailed, simulatable prototype city that is representative of a given typology. In this paper, we focus on two auto-dependent typologies for carrying out our simulation experiments: *Auto Sprawl* and *Auto Innovative*. These two typologies represent almost all the metropolitan areas in the US and Canada, where the automobile is the dominant mode of transportation. Using our generalized approach, we build two eponymous prototype cities and use a state-of-the-art agent-based platform, SimMobility, for integrated demand microsimulation and supply mesoscopic simulation. We analyze the impacts of automated mobility-on-demand (AMoD) implementation strategies in both prototype cities based on induced demand, modal and location shifts, level of service (congestion), energy consumption and emissions. Our results provide insights into how

[1] Recent statistics indicate that global sales of passenger vehicles dropped to 78.9 million units in 2018 from 79.6 million units in 2017. However, these numbers are still over 40% greater than the annual average number of passenger vehicles sold between 2000 and 2015, which was 54.9 million units. Since 2016, annual passenger vehicles sales have averaged no less than 74 million units (`https://www.statista.com/statistics/200002/international-car-sales-since-1990/`).

[2] 1 $MtCO_2e$ (million metric tonne of CO_2 equivalent) is equivalent to 1 Tg CO_2e (teragram of CO_2 equivalent).

[3] Subsequently, in this paper, we will use the term "mobility-on-demand" (MoD) to refer to the mode comprising taxi and ridesourcing services. We acknowledge that MoD is currently being standardized to encompass integrated and multimodal on-demand mobility beyond passenger ridesourcing (`https://www.transit.dot.gov/regulations-and-guidance/shared-mobility-definitions`). However, for the sake of consistency with the term AMoD, as used in this manuscript, and also with recent papers published in this research stream, we will restrict the term MoD to refer to taxi and ridesourcing.

AMoD can be effectively harnessed in two distinct city types relevant to the US and Canada.

In recent years, several key efforts have been made in understanding the role and impact of AMoD on future urban mobility, yet major gaps still remain in our understanding of how these impacts would vary with urban form or topology. Efforts utilizing agent-based simulation to analyze the introduction of AMoD (Martinez et al., 2015; Fagnant and Kockelman, 2014) have been limited to specific cities and case studies, giving rise to results that are not very generalizable. More recent work investigating the integration of AMoD with mass transit (Wen et al., 2018; Shen et al., 2018; Scheltes and de Almeida Correia, 2017; Liang et al., 2016), while agent-based, have also been narrowly defined for singular cities. Conceptual considerations of on-demand mobility across variations in urban form have also been explored by (Shaheen et al., 2017). Regarding the generation of representative prototype cities for simulating urban typologies, there is no comprehensive extant approach for this process from open-source data to agent-based simulation. Furthermore, only very few studies have accounted for demand-supply interactions (Wen et al., 2018; Azevedo et al., 2016) or incorporated high-resolution activity-based modeling systems (Nahmias-Biran et al., 2019). While attempts to assess the impacts of shared AMoD systems under globally optimal conditions (Alonso-Mora et al., 2017; Vazifeh et al., 2018) have demonstrated the potential of AMoD to reduce congestion in perhaps very specific urban settings, these studies neglect behavioral impacts. There are also no examples in the extant literature of AMoD impact assessment for sprawling auto-dependent cities that are chiefly to be found in North America. The question of whether (A)MoD can better complement or substitute mass transit has been addressed by Basu et al. (2018) and Hall et al. (2018), although the latter use a purely data-driven approach absent of simulation. Current results show that apart from large, dense, transit-oriented cities, (A)MoD could either complement or substitute transit in auto-dependent cities, such as are found in the US, depending on the availability of transit. Our paper thus makes significant contributions to the existing literature on AMoD impacts on future mobility in the following areas:

(a) We introduce a generalized, integrated approach to prototype city generation for representative agent-based simulation of urban typologies
(b) We conduct large-scale simulations of AMoD implementation strategies in prototype cities integrating demand microsimulation and supply mesoscopic simulation
(c) We address demand-supply interactions and compare impacts for two distinct auto-dependent urban typologies

The organization of the remainder of this paper is as follows. First, we describe in Section 2 our experimental framework and methodology. This begins with a background (Subsection 2.1) on the urban typologies and a detailed explanation of our simulation environment. We then discuss the prototype city generation approach (Subsection 2.2), including model calibration and validation steps. Lastly, in this section, we motivate and present the scenarios simulated in this study (Subsection 2.3). Data sources are also indicated throughout this section where relevant. Calibration and simulation results of the AMoD scenarios for both cities are presented and discussed in Section 3. We conclude with a summary of key findings and contributions, outlining steps for further work in Section 4.

2 Materials and methods

In this section, we describe the key elements of our methodology. First, we introduce the reader to the urban typologies that serve as the bases for our prototype city simulations (Subsection 2.1). We also summarize the relevant aspects of the simulation environment used in performing our experiments. Following this background, we discuss our approach to generating the prototype cities (Subsection 2.2). Finally, in Subsection 2.3, we explain the motivation, design and implementation of the AMoD scenarios within the simulation environment. An overview of the analytical framework is given in Figure 1.[4]

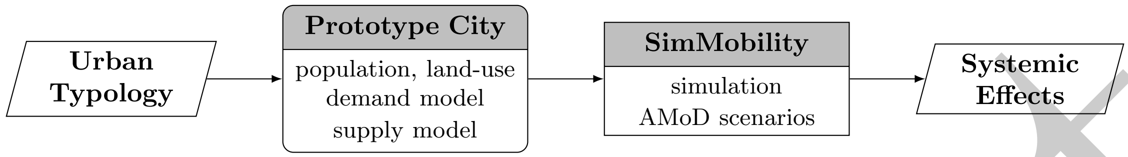


FIGURE 1 Overview of analytical framework

2.1 *Background*

2.1.1 *Urban typologies*

Given the infeasibility of modeling and simulating every city of interest in the world for future urban mobility analyses, we use the approach of simulating prototype cities, each representing a group of similar cities we refer to as an urban typology. Thus, our objective in collecting urban data and subsequently classifying cities is to enumerate typologies with distinct land-use, travel behavior, topology, socio-economic, energy and emissions characteristics. From these typologies, we can generate prototype cities that are representative of their urban typology outcomes. Ultimately, by simulating prototype cities, we can obtain insights that are broadly applicable to member cities of a given typology.

Data was gathered from 331 metropolitan areas worldwide using 69 mobility, demographic, land use, and economic indicators.[5] We used exploratory factor analysis (EFA) as a dimensionality reduction approach in order to obtain an interpretable transformation of the original dataset. This is based on the hypothesis that there exists a latent or underlying structure governing the variables in the data. The EFA yielded an optimal representation of 9 factors: Metro Propensity, Bus Rapid Transit (BRT) Propensity, Bikeshare Propensity, Development, Population, Congestion, Sprawl, Network Density and Sustainability. We named the factors based on the corresponding variables with the most important loadings (contributions).

Hierarchical agglomerative clustering has been demonstrated as an effective unsupervised learning approach for classification. We performed the clustering on the dataset reduced from 69 dimensions (variables) to a rotated subspace of 9 dimensions (factors). In this case, the Ward algorithm gave the best performance. This resulted in 12 urban typologies (Table 1). For a complete description of the typologies and methods, the reader is referred to Oke et al. (2019).

2.1.2 *Archetype city selection*

The generation of a prototype city requires the input of (i) a population sample, (ii) land-use patterns and (iii) a supply network. While some relevant research has been done in the area of synthetic network generation (Zhou et al., 2015; Dai et al., 2016), the level of detail (including population and land-use factors) required for the high-resolution simulation we perform rendered extant approaches infeasible within the scope of our efforts. Thus, to ensure consistent demand-supply interactions unique to an input network, we obtain inputs (i) – (iii) from an *archetype city* of the respective typology. For our purposes, we define an archetype city as one that closely represents

[4] The prototype city data and models described in this section are available at `https://github.com/jimioke/mitei-prototype-cities`.

[5] An interactive dashboard for further exploration of the data and typology results is available at `its.mit.edu/typologies`.

the average observations of its typology. To obtain candidates for the archetype city, we first rank each city by its Euclidean distance to the typology centroid, which is the mean vector of the 9 factor values of that typology (Table 1). We then select the archetype city based on two criteria: (a) nearness to typology centroid, and (b) availability and quality of inputs (i) – (iii). In this paper, the subjects of our AMoD simulation experiments are *Auto Sprawl* and *Auto Innovative*, and their archetype cities are Baltimore and Boston, respectively. We compare them by their mobility factor scores (on a scale of 0 to 1) in Figure 2. As will be described in Subsection 2.2, while we require detailed archetype city inputs (i) – (iii), we calibrate the prototype city to typology averages for the sake of representativeness.

TABLE 1 Top five cities ranked in order of closeness to the centroid of their corresponding typologies

Typology	**Rank 1**	**Rank 2**	**Rank 3**	**Rank 4**	**Rank 5**
Auto Sprawl	Baltimore	Tampa	Raleigh	Cincinnati	Indianapolis
Auto Innovative	Washington DC	Atlanta	Boston	Seattle	Toronto
BusTransit Dense	Rio de Janeiro	Bogota	Leon	Recife	Sao Paulo
BusTransit Sprawl	Mecca	Shiraz	Mashdad	Santa Cruz	Medina
Congested Boomer	Bangalore	Lahore	Kinshasa	Karachi	Chennai
Congested Emerging	Kumasi	Phnom-Penh	Conakry	Bandung	N'Djamena
Hybrid Giant	Sendai	Warsaw	Sofia	Hiroshima	Sapporo
Hybrid Moderate	Cordoba	Havana	Izmir	Valparaiso	Panama City
MassTransit Heavyweight	Berlin	Madrid	Munich	Seoul-Incheon	Singapore
MassTransit Moderate	Jerusalem	Newcastle	Tel-Aviv	Liverpool	Turin
MetroBike Emerging	Shenyang	Harbin	Changshaw	Zhengzhou	Qingdao
MetroBike Giant	Shenzhen	Guangzhou	Shanghai	Chongqing	Beijing

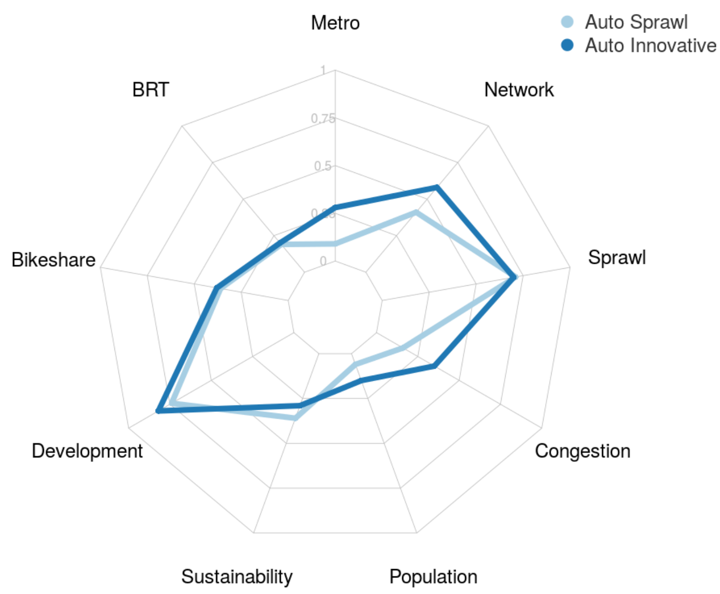


FIGURE 2 Comparison of the *Auto Sprawl* and *Auto Innovative* typologies based on normalized factor scores

The distinguishing characteristics between *Auto Sprawl* and *Auto Innovative* are the Metro Propensity, Population and Network density factors. *Auto Innovative* also suffers more from congestion and inefficiency than does *Auto Sprawl*, which is less dense (both in terms of network and population). We further compare these two typologies on a selected number of input variables (Table 2). From Table 2, we see that *Auto Sprawl* has a higher car mode share, while *Auto Innovative* has a mass transit mode share that is three times greater than that of its counterpart.

Table 2 Selected characteristics of *Auto Sprawl* and *Auto Innovative*

Selected Variable Typology Average	**Value**	
	Auto Sprawl	Auto Innovative
Car mode share (%)	86.4	78.4
Mass Transit mode share (%)	3.6	10.7
Bike mode share (%)	0.5	0.9
Walk mode share (%)	2.8	3.6
Population (million)	1.7	5.3
Population density (thousand/sq. km)	1.48	1.62
Per Capita GDP ($1000)	50.0	60.1
Annual CO_2 emissions per capita (tCO_2e)	16.5	13.7

Auto Innovative also is also three times as populated on average, but only slightly denser, which indicates that *Auto Sprawl* cities have larger areas than those in *Auto Innovative*. Lastly, *Auto Sprawl* generates 20% more per-capita CO_2 emissions than *Auto Innovative*. Given these key differences between the two typologies, our approach of conducting a detailed agent-based simulation of mobility in respectively generated prototype cities can provide typology-specific insights on the impacts of AMoD deployment, particularly with regard to mass transit and energy consumption.

2.1.3 Modeling and simulation environment

Assessing the systemic evolution of a city for behavioral implications and performance impacts of new mobility services requires a highly detailed large-scale simulation environment. We use the platform, SimMobility, which employs an integrated activity- and agent-based framework for analyzing future mobility scenarios by capturing the decisions and movements of a given population (Adnan et al., 2016; Azevedo et al., 2016). SimMobility is modular with three components that operate at different temporal scales. The Short-Term simulator is microscopic, with up to millisecond granularity for small scale network performance analyses. The Mid-Term module simulates daily demand, supply and demand-supply interactions, while the Long-Term component models land development and economic interactions at the annual scale. For the purposes of the AMoD scenario experiments in this paper we use the Mid-Term module (hereafter referred to as *SimMobility-MT*), which requires as an input, a synthetic population and land-use specification, along with a transport supply network specification. SimMobility-MT consists of three submodules, namely: *PreDay*, *WithinDay* and *Supply*. The *PreDay* sumbodule simulates daily travel plans, *WithinDay* simulates modifications to plans and events, while *Supply* concurrently simulates agent events and actions. An important feature of SimMobility-MT is "day-to-day learning," whereby aggregate zonal travel times from *Supply* are updated in *PreDay* to iteratively modify passenger activity decisions based on network performance.

The *PreDay* module consists of an activity-based model (ABM) and an implementation framework (developed in C++ and Lua) which produce a daily activity schedule (DAS) for each individual in the synthetic population. In the *PreDay* ABM, interconnected hierarchical choice models are used to generate the DAS across three levels: *Day Pattern*, *Tour* and *Intermediate Stop*. As indicated by solid lines in Figure 4, higher-level choice models are conditioned on the lower-level models. Disaggregate accessibility measures, capturing the expected utility of all alternatives, from the mode and destination choice models are integrated into the higher-level models, as indicated by the dashed arrows, as well as summarized in Table 4. At the *Day Pattern* level, decisions are made pertaining to the binary choice of leaving or staying at home, the multinomial choice of primary activities to perform (and their sequence), the multinomial choice of stop purpose combinations before or after the primary activity of each tour, and the number of tours. The *Tour* level contains models for usual work choice (binary), mode choice for fixed-location tours, mode-destination

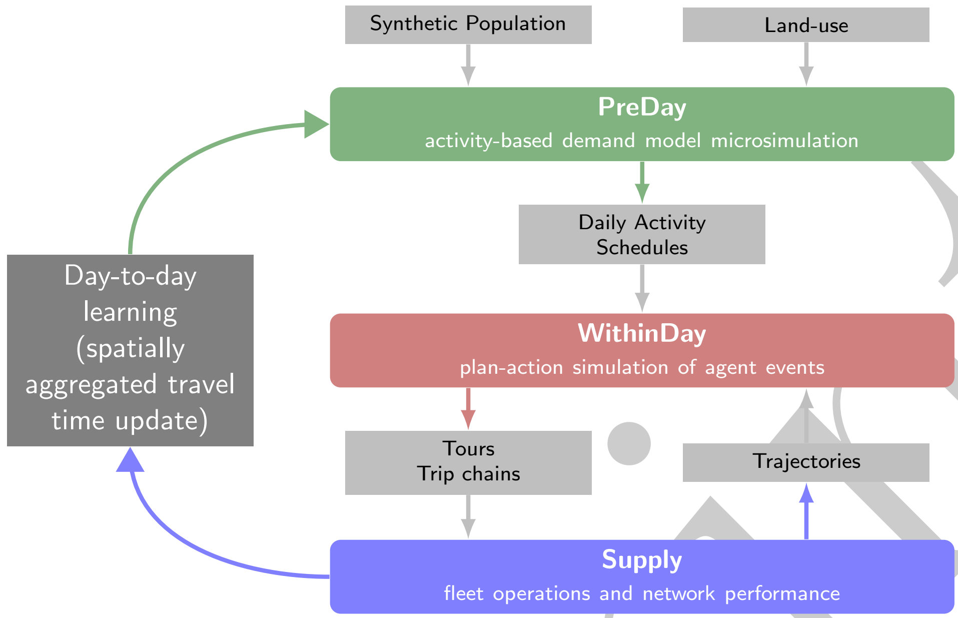


FIGURE 3 SimMobility MidTerm framework

choice for tours without fixed locations, start time and duration of each tour. Work-based subtour decisions are also handled at this level. The lowest level is the *Intermediate Stop*, where the number, mode and destination and departure times of stops made before or after primary activities are decided. Upon simulation of the ABM, the resulting DAS is a detailed plan of each individual's sequence of activity categories, times and locations at a half-hour resolution. A summary of the *PreDay* ABM is given in Table 3. For a more detailed description of the *PreDay* module, the reader is referred to de Lima et al. (2018). For the prototype cities, we limit the activity choice set to the following: *Work*, *Education*, *Shop* and *Other*. Personal errands and recreation are among the purposes captured by the *Other* designation. Utility equations for selected models in each of the levels indicated in Figure 4 are provided in Appendix B.

TABLE 3 *PreDay* activity-based submodels and their activity dimensions across the three levels: *Day Pattern*, *Tour* and *Intermediate Stops*.

Level	**Models**	**Activity dimensions**
Day Pattern	Travel	N/A
	Tour	All
	Intermediate Stops	All
	Number of Tours	All
Tour	Usual Work	Work
	Mode	Work, Education
	Mode-Destination	Work, Education, Shop, Other
	Time-of-Day	Work, Education, Shop, Other
	Work-based Subtour	Work
	Subtour Mode-Destination	Work
	Subtour Time-of-Day	Work
Intermediate Stops	Generation	All
	Mode-Destination	All
	Time-of-Day	All

In the *WithinDay* submodule, the route choice, rescheduling and mode change decisions of the agents are simulated based on the plan-action framework (Ben-Akiva, 2010) in response to

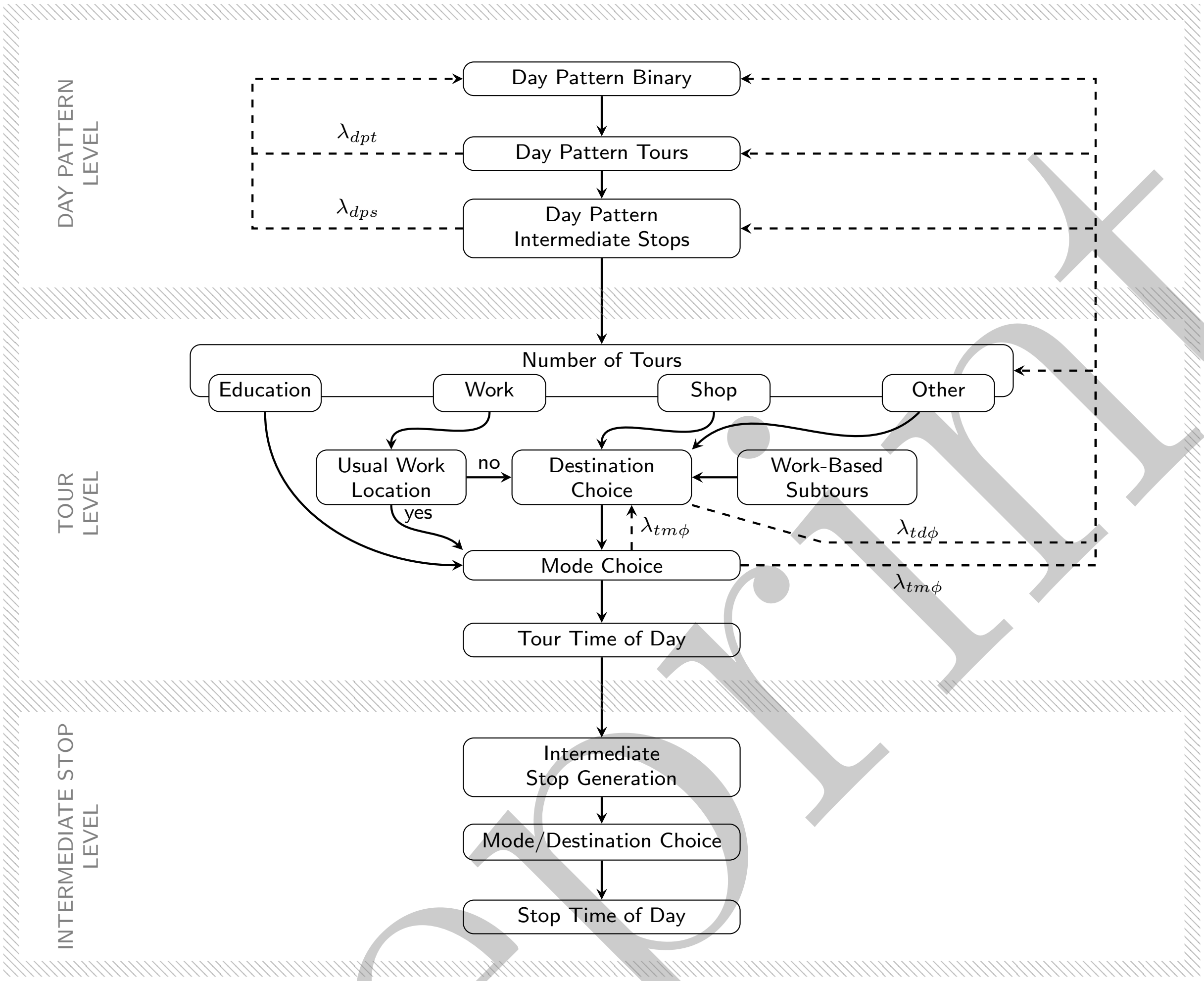


FIGURE 4 SimMobility MidTerm *PreDay* activity-based model structure. Solid lines indicate conditional dependency in the direction of the arrows. Dashed lines indicate accessibility integration in the direction of the arrows. The hatched outlines demarcate the three levels of the model.

TABLE 4 Summary of inclusive values used in *PreDay* model. $\phi = \{\text{Work}, \text{Education}, \text{Shop}, \text{Other}\}$

Symbol	**Model source**	**Model inclusion**
λ_{dpt}	Day Pattern Tour	Day Pattern Binary
λ_{dps}	Day Pattern Stop	Day Pattern Binary
$\lambda_{td\phi}$	Destination Choice	Number of Tours, Day Pattern Intermediate Stops/Tours/Binary
$\lambda_{tm\phi}$	Mode Choice	Destination Choice, Number of Tours, Day Pattern Intermediate Stops/Tours/Binary

information and control. The resulting trip-chains are then performed in the *Supply* submodule. Following the dynamic traffic assignment framework, trip trajectories obtained in *Supply* are used to update link travel times, which in turn modify the routing decisions of the passengers. The *WithinDay* submodule also handles daily events, using a publish-subscribe mechanism. Events are defined across five dimensions: time, information, location, perception and requests.

Public (bus, train, on-demand) and private (car, motorcycle, private bus, walk) modes are simulated in the *Supply* submodule. The network-level simulation is mesoscopic and time-based. Thus, the capacities are determined by lane-groups, and speed-density relationships govern the movements of vehicles in each segment. At each advance interval (up to a five-second resolution), speeds, positions and energy consumption are reported for each vehicle. Vehicle fleet operations for the public modes are also implemented in *Supply*. For public buses and trains, schedules and headways are implemented for route management. A bus controller and train controller track movements and occupancy, and can also handle disruptions. The mobility-on-demand (MoD) services, namely: taxi and ridesourcing service, are also managed by a controller to handle ride matching, vehicle routing and dispatching and rebalancing operations. Powertrain-specific energy consumption is calculated based on models incorporated into SimMobility from Rakha et al. (2011); Fiori et al. (2016); Wang and Rakha (2017a,b).

2.2 Prototype city generation

To assess the impacts of future mobility strategies in different urban typologies, we create representative prototype cities for simulation in SimMobility. The creation of a prototype city involves data-driven approaches for generating population, demand and supply parameters. From the member cities of a typology, a archetype city is chosen near the centroid of the typology characteristics (as discussed in Subsubsection 2.1.2). The archetype city serves as the source of the population data and attributes, supply network, land use, and demand-supply interaction patterns. We create a synthetic population and land-use specification of the archetype city in order to preserve the demand-supply patterns for its network. However, mode shares (including MoD) and activity shares are calibrated to the typology average, in order to capture behavioral impacts at the typology level. Following this, the we calibrate the prototype city to fit archetype city values for average travel times (by activity), link speeds and waiting times in order to ensure consistency with the archetype city network. In each of these steps, we generalized from existing approaches (Strauch et al., 2005; Le et al., 2016; Fournier et al., 2020) and leverage on open data sources.

2.2.1 Population and land-use

We define four levels of spatial resolution for population and land-use specification in a prototype city Figure 5. The urban typology prototype city (UTP), representing the highest level, is given by the metro area of the selected archetype city. The Second Administrative Levels are the highest-level regions into which the UTP is subdivided. Traffic Analysis Zones (TAZs) are obtained for the UTP given the archetype road network. Where unavailable, these are generated as Thiessen polygons. CELLs are obtained by gridding over the entire UTP. The corresponding land use weight at the CELL level is used to determine the *number* and *location* of the households therein, as well as the *numbers* of work (WORK) and education (EDU) locations in each TAZ.

We use the Hierarchical Iterative Proportional Fitting (HIPF) method (Müller and Axhausen, 2012) to generate a synthetic population for each prototype city. Aggregate-level data area obtained from the American Community Survey, while household and individual sample data are obtained from the Public Use Microdata Samples. At the household level, we control for the (i) income and (ii) vehicle ownership variables, while at the individual level, we control for the (iii) age, (iv) gender and (v) employment status variables. (See Figure 8 and Figure 9 for a concise validation summary of these variables.) These individual and household variables, including others such as household composition and education level, are the population-based predictors in the demand modeling framework. Land-use variables are also required for the destination-choice models. These include TAZ area, number of firms/businesses, number of shops, distance to city center, among others. Travel time and cost of modes are also inputs to the mode-choice models. The required variables differ across the various choice models. However, some representative model specifications are described in detail in Appendix B. A summary of the syntheses and allocation interactions is depicted in Figure 6.

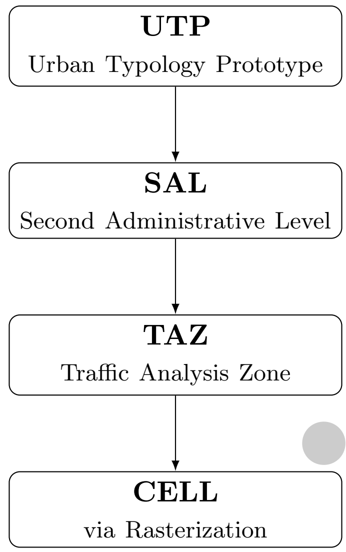


Figure 5 The four levels of spatial resolution in the prototype city development framework.

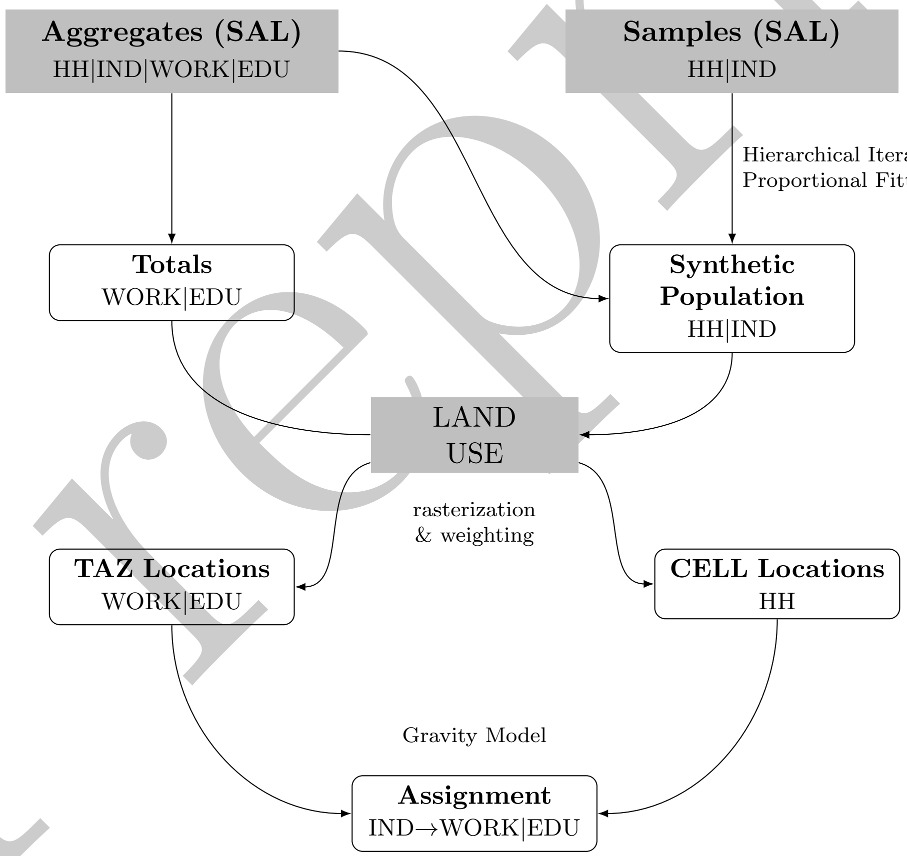


Figure 6 Population and land-use synthesis framework for generation and allocation

Land-use category data enable us to spatially allocate households, places of employment and education. We further apply IPF to obtain TAZ-level employment and education cross-tabulation weights, and then use gravity models to perform the final assignments at the individual level. Further details are provided in Appendix A.

2.2.2 Demand

The *PreDay* ABM was estimated for one of the North American typology archetype cities—Boston. Data were obtained from the Massachusetts Travel Survey (MTS) and the Central Transportation Planning Staff (CTPS) of Boston for the period 2010 to 2011. The estimation was performed using the PythonBiogeme platform. Further details and results on the model and estimation procedure can be found in de Lima et al. (2018). The Boston model served as a starting point for the *Auto Sprawl* and *Auto Innovative* models. In calibrating the ABM for the prototype cities, we adjust the following quantities to match the mode and activity shares for their respective typologies: alternative-specific constants, specific variable coefficients and scale parameters (which control the correlations in the nested logit structure). The calibration approach is formally described in Chen et al. (2019). The activity validation data are obtained from available travel surveys for cities in the respective typologies (see Table 5). These data were not used in the typology discovery, given their sparsity at that scale. Thus, individual reports were searched for the few cities available in these typologies in order to obtain aggregate activity shares. Sources are also detailed in Table 5. Temporal patterns, mode-activity proportions, trip and tour distributions are also checked. Furthermore, we validate our model for consistency with expected commuter distances and travel times. Similar checks are also performed for education trips. We justify the approach of using the Boston model as a starting point based on the fact that it is the archetype city for *Auto Innovative* and that travel patterns are relatively similar across cities in the United States. Details of four representative choice models including the specific variables used in each of them are given in Appendix B.

TABLE 5 Activity share validation data for the prototype cities

Measure		**Auto Sprawl**	**Auto Innovative**
Activity shares (%)	Work	28	26
	Education	12	13
	Shop	16	15
	Other	44	47
Trips per person		3.5	3.5
Survey data sources		Detroit, Edmonton, Indianapolis, Melbourne, Minneapolis-St. Paul, Portland, OR	Atlanta, Boston, Philadelphia, San Francisco, Seattle, Toronto, Vancouver

Furthermore, we fit the model to fuel price elasticities of trip demand based on the literature relevant to US cities (Small and van Dender, 2007; Goodwin et al., 2004). These values are summarized in Table 6.

TABLE 6 Short-term trip elasticities used in validating demand model

Mode	**Elasticity**	**Scope**	**Source**
Car	−0.10	North America & Europe	Goodwin et al. (2004)
	−0.03	USA	Small and van Dender (2007)
	−0.11	Europe	TRACE (1999)
Mass Transit	0.13	Europe	TRACE (1999)
Active Mobility	0.18		

Time and cost are critical variables in the mode and mode-destination choice models in the demand framework. Given that the spatial resolution of the ABM is at the TAZ level, distances, travel times (private and public modes) and fares are estimated for each TAZ OD pair. TAZ-TAZ distances are initially estimated based on the Manhattan separation of the centroids. They are later updated via day-to-day feedback until equilibrium. Travel times for *Walk* are then calculated using an average speed of 5 km/h, while those for *Bike* are computed using a speed of 20 km/h. Mass transit costs are given by fares at the zonal level. These are obtained from the transit agency of the archetype city where available. Otherwise, an average cost factored by the TAZ pair distance is used as an estimate. Mass Transit travel time accounts for both the waiting and travel times, along with access and egress time. For Car, the initial travel time is also obtained from the archetype city agency where available, or estimated based on a network-wide average speed if not. The cost for Car explicitly accounts for that of fuel (gasoline or electricity), a general distance-based operational cost and parking at the destination TAZ. Taxi fares are given by the following function[6]:

$$C_{taxi} = F^b_{taxi} + F^d_{taxi} d + F^w_{taxi} \cdot \max\{0, t - \tfrac{d}{40}\} \tag{1}$$

where F^b_{taxi}, F^d_{taxi} and F^w_{taxi} are the base fare, distance charge (per kilometer) and waiting time charge (per hour), respectively. The trip distance is d, while t is the travel time (same for Car mode). The expression $\max\{0, t - \frac{d}{40}\}$ represents the additional time spent in transporting the passenger given an expected speed of 40 km/h. In the case of ridesourcing, we assume the current cost structure for Uber XL (standard service)[7], where

$$C_{RS} = \max\left\{F^{min}_{RS}, \left(F^s_{RS} + F^b_{RS} + F^d_{RS} d + F^t_{RS} t\right)\right\} \tag{2}$$

F^{min}_{RS} is the minimum fare, while F^s_{RS}, F^b_{RS}, F^d_{RS} and F^t_{RS} are the service fee, base fare, distance charge (per kilometer) and travel time charge (per hour), respectively. The fare values for both services are obtained from the corresponding archetype city, as shown in Table 7.

TABLE 7 Mobility-on-demand fare parameter values for *Auto Sprawl* and *Auto Innovative*

Mode	**Parameter**	*Auto Sprawl* (USD)	*Auto Innovative* (USD)	**Description**
Taxi	F^b_{taxi}	1.80	2.60	Base fare
	F^d_{taxi}	1.38	1.75	Per-km charge
	F^w_{taxi}	0.40	0.47	Per-min charge
Ridesourcing	F^{min}_{RS}	6.85	6.85	Minimum fare
	F^s_{RS}	2.35	1.85	Service charge
	F^b_{RS}	1.10	2.10	Base fare
	F^d_{RS}	1.38	1.35	Per-km charge
	F^w_{RS}	0.12	0.21	Per-min charge

2.2.3 Supply

The transport supply network model development procedure entails the following steps:

(a) obtaining and processing the road network of the archetype city from Open Street Maps and generating inputs suitable for mesoscopic simulation

(b) generating a public transit graph from General Transit Feed Specification (GTFS) sources

(c) calibrating supply parameters to achieve realistic baseline network performance

We developed tools to perform these steps to facilitate simulation in SimMobility-MT.

[6] This represents the existing taxi tariff structure in the two archetype cities of Baltimore and Boston. Fare coefficients were obtained from `https://www.taxi-calculator.com/`.

[7] Fare coefficients available at `http://uberestimate.com/prices/`.

The first step in creating the transport supply network model is obtaining the road network contained within the boundary of the archetype city. This boundary specifies the metropolitan region associated with the city. Within this boundary, the drivable road network, consisting of nodes, links, lanes, segments (and their capacities), lanes, turning connections, and their respective geometries, are extracted from the Open Street Maps database. First, the extracted network is cleaned and simplified in order to remove extraneous nodes at intersections. Following this, we construct segments by subdividing links to account for changes in geometry and number of lanes. We then use connectors to specify lane associations between consecutive segments. Using a greedy assignment, we finally create turning paths to indicate lane connectivity at intersections. In cases where the number of lanes and capacities were unavailable for a certain, we used the link category to infer these values. The road networks for *Auto Sprawl* and *Auto Innovative* are shown in Figure 7. These are the actual networks of the respective archetype cities of Baltimore and Boston.

The second step is the generation of the mass transit portion of the supply network model. This constitutes the rail network system (where available) and the bus network system. Open data based on the GTFS open data on the schedules, shapes and locations of transit trips and routes. For the rail system, we specify the access segments from the road network. In the case of the bus system, we map the stops to segments in the road network, with access and egress nodes specified. From these, we construct a transit route-choice graph which represents the connectivity within the mass transit system (bus and rail), and between the transit and road network systems. In the transit graph, there are three edge types: (i) bus, (ii) rail and (iii) walk. The simulation of the bus and rail systems also requires inventory and control data. Thus, we obtain from the archetype city, where available, fleet and capacity information for the bus and rail routes. We also assume movement parameters including speed and acceleration limits.

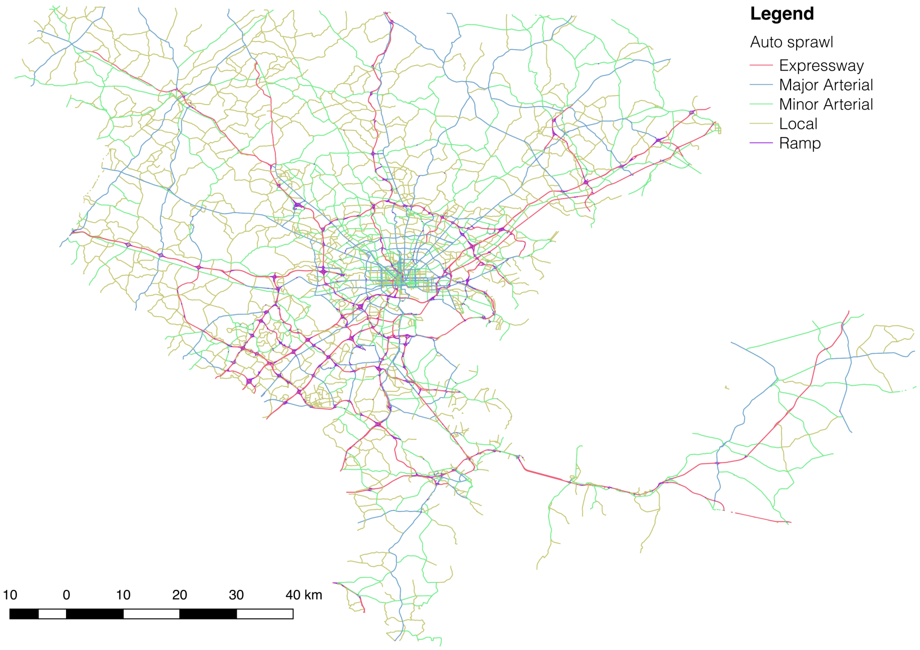
Table 8 Selected transport network supply model parameter specifications for *Auto Sprawl* and *Auto Innovative*

Network components		*Auto Sprawl*	*Auto Innovative*
Road	Nodes	11 410	18 016
	Links	24 133	46 763
	Segments	252 006	164 980
Bus	Routes	139	844
	Stops	797	4 170
Rail	Routes	10	25
	Stations	76	121

Once the transport supply network model has been specified, the final step is to generate pathsets for both the road network and the mass transit network. Pathsets for the transit network are computed based on the transit graph. Separate tables are used to store the road network paths and those for the transit network. For each possible OD pair of nodes, we generate pathsets under various scenarios. Each pathset is characterized by its length, travel time, among others. In the *WithinDay* module, travelers choose their trip paths based on the outcome of the route choice model. We generate the pathsets prior to simulation in order to save computation time. Given that paths cannot feasibly be computed for every possible OD pair, we limit the set of OD's to the initial daily activity schedule that has been generated. Further OD's are generated as needed upon subsequent day-to-day simulations. Beyond this, we generate new pathsets for the ODs that cannot be mapped for the demand case at hand. In time, sufficient pathsets are generated to handle new demand cases.

From the trajectory output of *Supply*, we feedback link-based travel times for *WithinDay* iterations. This allows for route-choice modifications in order to reach an equilibrium in traffic

(A) *Auto Sprawl*



(B) *Auto Innovative*

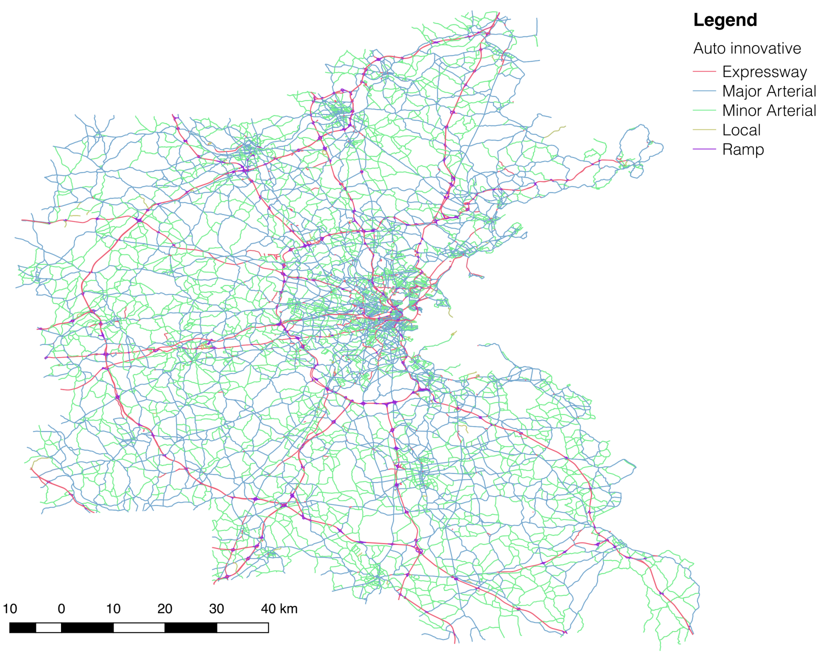


FIGURE 7 The road networks of *Auto Sprawl* and *Auto Innovative* prototype cities

assignment. Once this equilibrium is reached, link-based travel times are aggregated at the zonal level in order to update the costs for day-to-day learning.

2.3 *Scenario design*

The scenarios in this paper have been designed to investigate the future impacts of automated mobility-on-demand (AMoD) across auto-dependent cities (largely to be found in North America). AMoD services are expected to influence behavior and energy consumption on the demand side, while the supply side might register impacts on congestion, emissions and parking. Underlying these shifts will be changes in on-demand service costs, which have been estimated to reduce by up to 50% (from \$1.72 to \$0.92 per km) as a result of automation in Singapore (Pavone, 2015; Spieser et al., 2014). Similar gains are also projected in the US of up to 50%, from \$0.65 per mile to under \$0.30 per mile (Stephens et al., 2016). In Switzerland (Europe), even greater cost savings have been estimated (Bösch et al., 2018). As a moderate estimate, therefore, we implement AMoD service provision based on a 50% reduction in user fare from the existing traditional taxi mode, using the cost structure as given by Equation 1. The ridesourcing services represented by the likes of Uber are already cheaper than traditional taxi.

Table 9 Mode availabilities across scenarios

Mode		Base Case	AMoD Intro	AMoD No Transit	AMoD Transit Integration
Car	Drive Alone	✓	✓	✓	✓
	Pool (2)	✓	✓	✓	✓
	Pool (3)	✓	✓	✓	✓
Mass Transit	Bus	✓	✓	✗	✓
	Train	✓	✓	✗	✓
On-demand	MoD	✓	✗	✗	✗
	AMoD Single	✗	✓	✓	✓*
	AMoD Shared	✗	✓	✓	✓*
Active Mobility	Bicycle	✓	✓	✓	✓
	Walk	✓	✓	✓	✓
Other	Private Bus	✓	✓	✓	✓
	Motorcycle	✓	✓	✓	✓

*Restricted to short trips and first/last mile trips to rail stops.

Table 10 Powertrain assumptions and specifications

Vehicle	**Powertrain**
Private car	Gasoline-powered (conventional and hybrid), battery-electric
Ridesourcing	Gasoline-powered hybrid
AMoD	Battery-electric
Public Bus	Diesel
Train	Electric

2.3.1 *Base Case*

The *Base Case* represents the current state of activity and mobility patterns in the prototype cities for the year 2016. The car modes include "Drive Alone" and "Car Pool". The latter is modeled explicitly for 2 or 3 vehicle occupants. Mass transit (bus and rail) is available based on the public transport supply model, as described in Subsubsection 2.2.3. Mobility-on-demand

(MoD) refers to both taxi and ridesourcing services, which represent smartphone-based ridesourcing services. The cost of ridesourcing is modeled according to the current fare structure of Uber[8].

The Private Bus mode captures the services provided by employers and private companies for transport to workplaces and schools, respectively. Its availability is assumed to be the same as that of mass transit, considering that these private transit services predominantly share the same bus stop infrastructure with their public counterpart.

2.3.2 AMoD Intro

In this scenario, automation is introduced for on-demand services. Thus, MoD is replaced by AMoD, while other modes in the *Base Case* remain unchanged. The choice of single ("AMoD Single") or shared ("AMoD Shared") rides is explicitly modeled. Shared vehicles have a capacity of 4 passengers, while depots are distributed across each TAZ. For every demand case, optimal fleet sizes are computed and vehicles are routed and rebalanced to minimize costs. We refer the reader to Marczuk et al. (2015); Basu et al. (2018) for further background on AMoD supply implementation in SimMobility.

2.3.3 AMoD No Transit

Here, AMoD replaces MoD (as in *AMoD Intro*) but mass transit is also removed. This scenario is an attempt to answer the question of whether AMoD can fully substitute mass transit and under what conditions it can successfully do so in the prototype cities of interest. It simulates a future in which transit is abandoned while AMoD proliferates.

2.3.4 AMoD Transit Integration

We consider the impacts of a government intervention in the operation of AMoD services in the scenario *AMoD Transit Integration*. Rather than have AMoD compete directly with mass transit, it is integrated as an access or egress mode to rail stops. These AMoD rides are necessarily shared and further subsidized by 20% of their original rates.[9] The integration radius is set at 7.5 miles from all rail stops based on the typical commute distances in these cities. Non-integrated AMoD remains available but restricted to local trips within the same 7.5-mile distance threshold. This integration also necessitates the enhancement of the transit route-choice graph, such that the AMoD access and egress links are now included in the transit pathsets. Integrated transit thus has greater availability in each of the prototype cities. At the *PreDay* level, the traditional mass transit mode (with walk-only access/egress) is nested with the AMOD-access/egress transit option.

Since we are simulating a prototype city, we do not delve into operational details here, as these can vary widely from one real city to another. Instead, we focus on simulating what might be regarded as the upper bound (best-case scenario) of transit integration performance. No restrictions are made to the number of passengers that can be picked up or dropped off at a given station.[10] However, travel times are fed back into the demand model, which is iterated in order to satisfactorily incorporate the effects of congestion on demand.

2.4 Calibration and validation

We validate each prototype city for population, demand and supply. The synthetic population is validated using the control variables at the individual (Figure 8) and household (Figure 9) levels.

[8] http://uberestimate.com/prices/

[9] Reports on mobility-on-demand operators (Li et al., 2019) and taxicab operators (Controller, 2005) indicate that 20% is the expected gross earning rate from total revenue. Subsidies at this level thus hypothetically ensure that service provision remains reasonably profitable for the AMoD operator(s).

[10] Furthermore, we assume no curbside restrictions on pickups and dropoffs

Further, we validate these totals at the Second Administrative Level (Figure 10). The spatial distributions across the SALs are also matched.

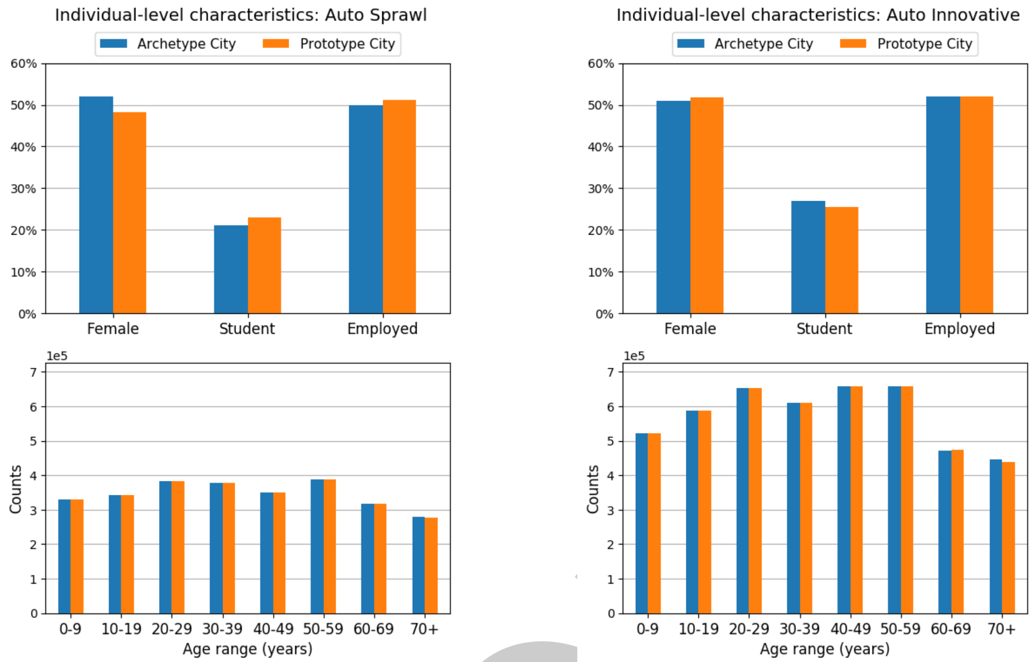


FIGURE 8 Individual-level validation for *Auto Sprawl* and *Auto Innovative*

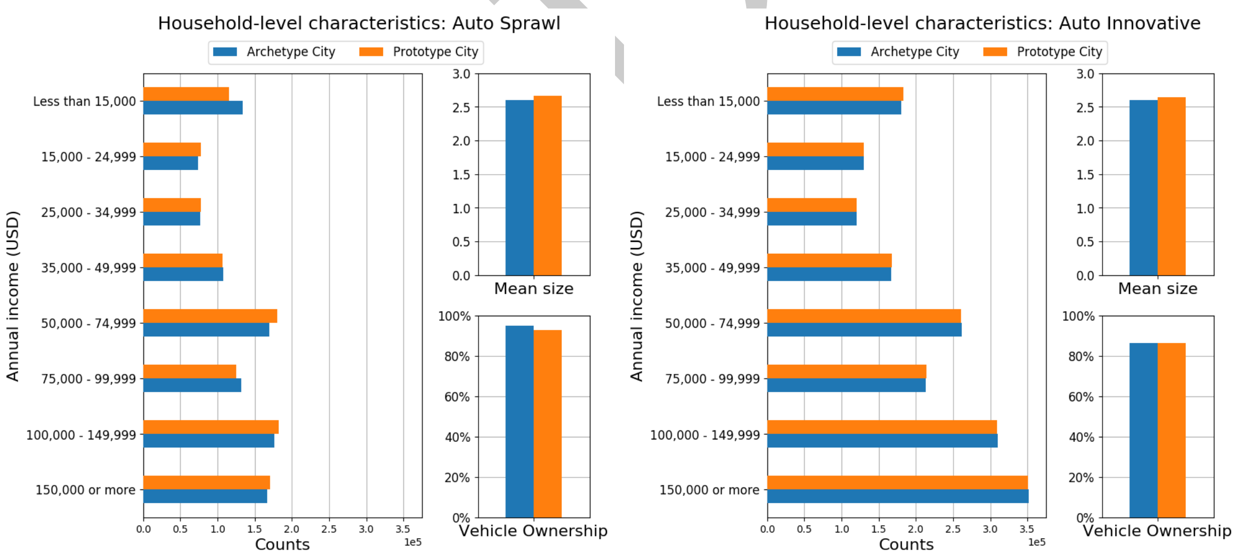


FIGURE 9 Household-level validation for *Auto Sprawl* and *Auto Innovative*

The demand model is calibrated and validated for activity and mode shares. The results are shown in Figure 11 and Figure 12 for both prototype cities. Aggregate activity and mode shares are similar for both cities, except when comparing the shares for Car and Mass Transit. In *Auto Innovative* the Mass Transit share is higher at the expense of that of Car, when compared to *Auto Sprawl*.

Further, we show the fuel cost elasticities of trip demand for both prototype cities in Table 11.

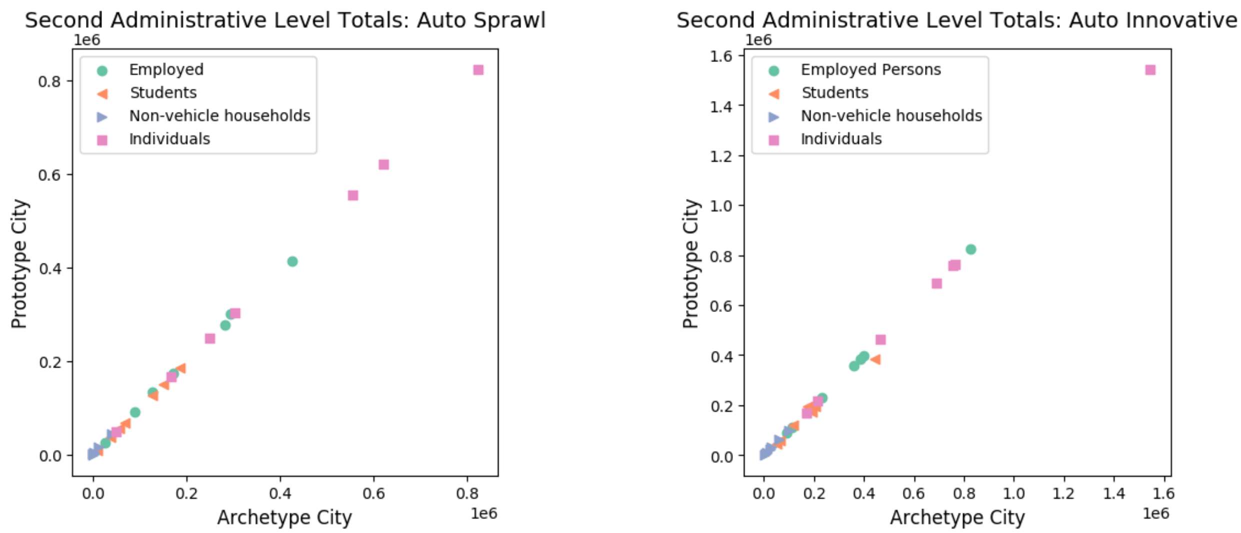


Figure 10 Second Administrative Level (SAL) validation for *Auto Sprawl* and *Auto Innovative*

(a) *Auto Sprawl* (b) *Auto Innovative*

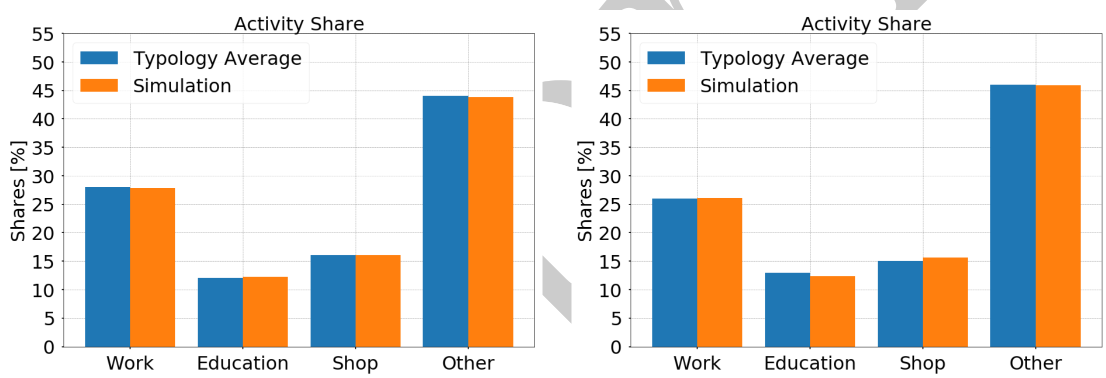


Figure 11 Activity share validation of the prototype city simulations

(a) *Auto Sprawl* (b) *Auto Innovative*

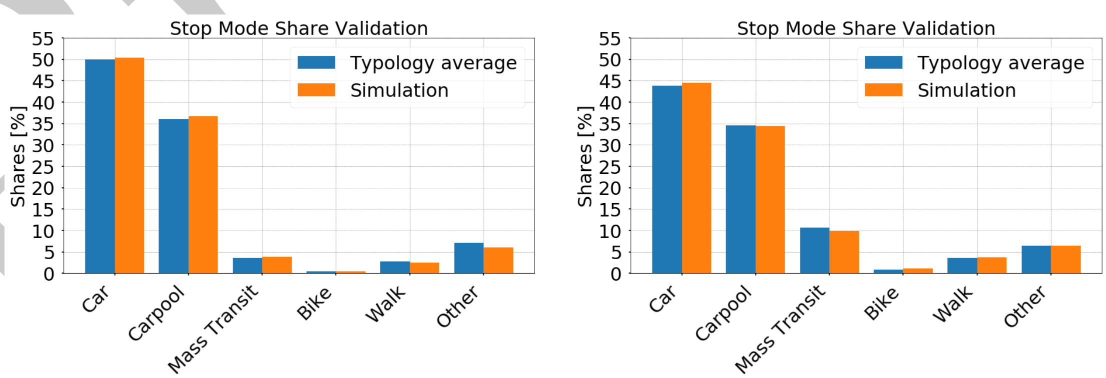


Figure 12 Mode share validation of the prototype city simulations

These values are matched as closely as possible to reference values in Small and van Dender (2007) and TRACE (1999).

Table 11 Base case fuel cost elasticities of trip demand for *Auto Sprawl* and *Auto Innovative*

Mode	*Auto Sprawl*	*Auto Innovative*
Personal Car	-0.03	-0.04
Mass Transit	$+0.06$	$+0.01$
Walk	$+0.01$	$+0.04$
Bike	$+0.08$	$+0.07$
All	-0.005	-0.006

2.4.1 *AMoD fare elasticity*

We then investigate the AMoD cost elasticities in both cities (Figure 13). In *Auto Innovative*, demand for AMoD is more elastic than in *Auto Sprawl*. This result indicates that AMoD meets a greater need in *Auto Sprawl*, and the demand for AMoD is less sensitive to fare increases than in *Auto Innovative*.

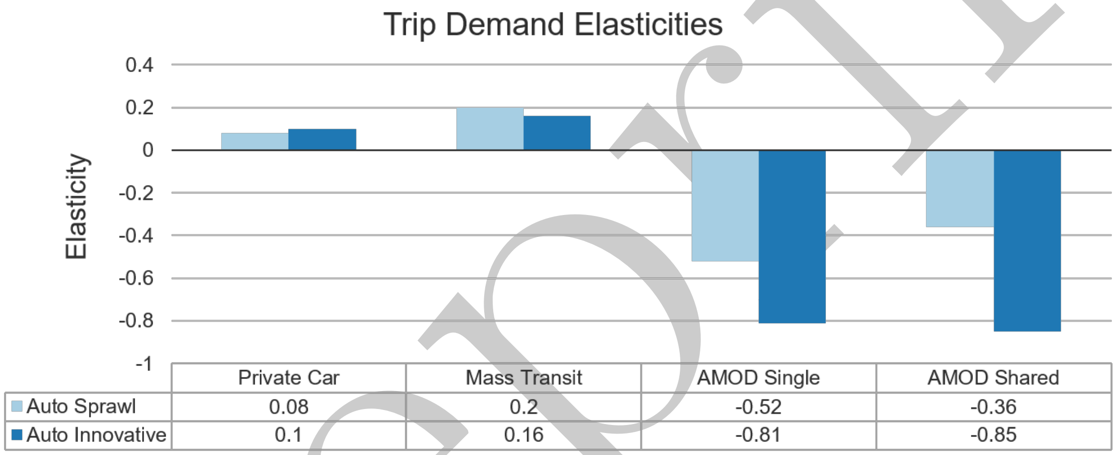


Figure 13 AMoD fare elasticities of trip demand

3 Results and discussion

We present the results of our simulations of the AMoD scenarios earlier described. These results belong to the convergent states after the day-to-day learning procedure. We focus on the demand, network level of service (vehicle kilometers traveled and congestion), energy and emissions impacts. The policy implications in each case are discussed in the respective subsections.

3.1 *Demand impacts*

The introduction of AMoD results in an increased number of trips in both prototype cities. Under the *AMoD Intro* regime, the percentage of induced trips with respect to *Base Case* is 4% in *Auto Sprawl*. This indicates that *AMoD* potentially satisfies an existing latent demand for mobility that the existing base levels do not meet. However, in *Auto Innovative*, induced demand is only 2% of base demand. In terms of absolute numbers, however, more trips are still generated in *Auto Innovative*, compared to *Auto Sprawl*, given its larger population and density.

When mass transit is removed under *AMoD No Transit*, the number of induced trips remains around the same level in *Auto Sprawl*. In *Auto Innovative*, the induced share becomes negative, indicating that fewer trips are made. This is a reflection of the importance of the mass transit system in this typology and the congestion effects of servicing previously transit-based trips by AMoD.

The mode shares are shown for *Auto Sprawl* and *Auto Innovative* in Figure 14. Trip modal shifts in both cities are shown for *Base Case* → *AMoD Intro* → *AMoD No Transit* in Figure 15, and for *Base Case* → *AMoD Intro* → *AMoD Transit Integration* in Figure 16. A summary of the demand impacts of the AMoD scenarios is given in Table 12.

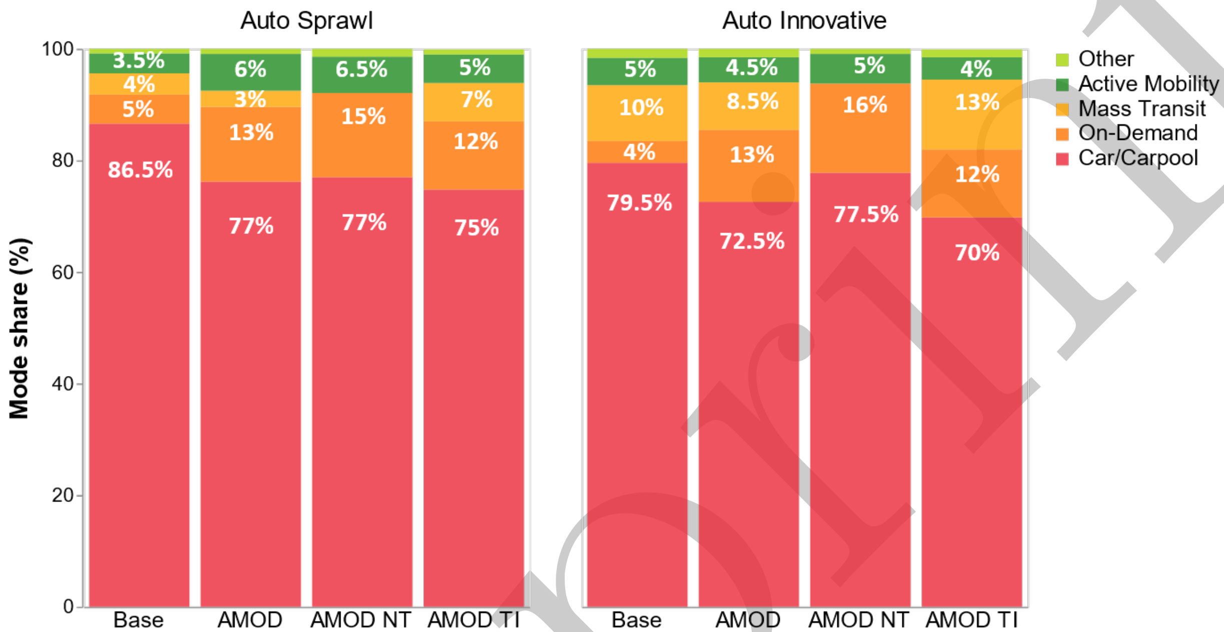


FIGURE 14 Mode shares across the four scenarios in *Auto Sprawl* and *Auto Innovative*

Under *AMoD Intro* in *Auto Sprawl*, on-demand trips increase by 170%, private car trips decrease by 7%, while transit ridership decreases by 21%. In *Auto Innovative*, however, on-demand trips increase by 240%, while private car trips and transit ridership decrease by 7% and 13%, respectively. Were transit to be abandoned (*AMoD No Transit* scenario), on-demand trips would only increase by 200% while private car trips would decrease by 5% in *Auto Sprawl*. On-demand trips would increase by 310% in *Auto Innovative* however, while private car trips would reduce by 1.3%. The lower impact on private car usage in *Auto Innovative* compared to *Auto Sprawl* under *AMoD No Transit* is due to the greater shift to Carpool from Mass Transit in *Auto Innovative*.

With the integration of AMoD and transit, on-demand trip request increases are moderated while transit cannibalization is reversed. In *Auto Sprawl*, on-demand trips increase by 140% and private car trips decrease by 13% but transit ridership increases by 88%. In *Auto Innovative*, on-demand trips increase by 220%, private car trips decrease by 11%, while transit ridership increases by 28% under *AMoD Transit Integration*.

Thus, we observe that a low-cost AMoD service results in a significant initiation of trips that would otherwise not exist due to the increased accessibility of the automated service. This effect is more pronounced in *Auto Sprawl* (less dense, minimal mass transit availability) than in *Auto Innovative* (dense, congested, moderate mass transit). Given that *Auto Sprawl* cities already have a low mass transit modeshare, they could potentially dis-invest in mass transit with a manageable demand for AMoD. In *Auto Innovative*, however, the removal of transit does not appear sustainable, as we find that AMoD cannot efficiently service all the erstwhile transit trips. Adopting a strategy whereby AMoD complements mass transit moderates the demand for AMoD while boosting transit

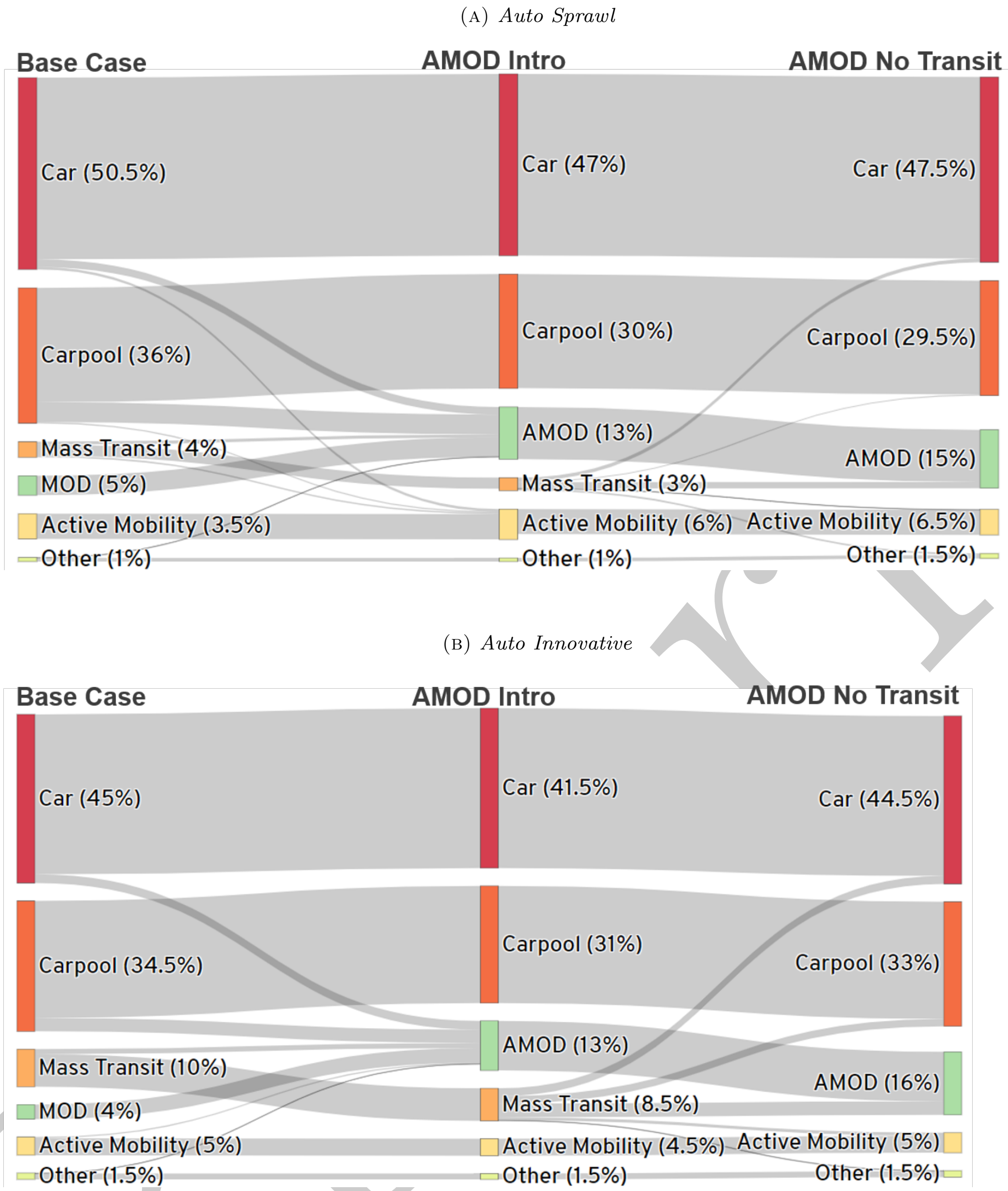


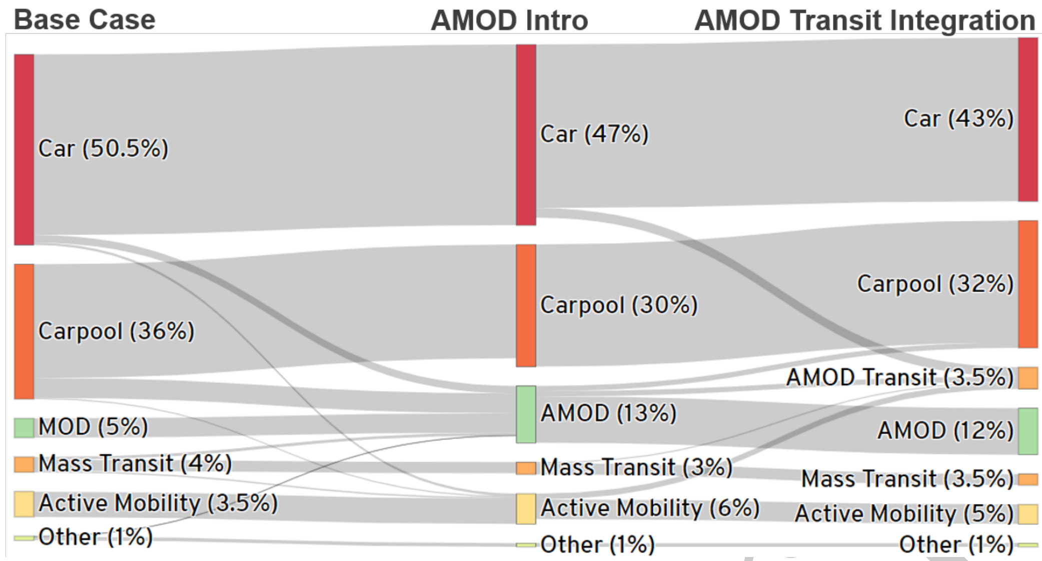
FIGURE 15 Modal shifts from *Base Case* → *AMoD Intro* → *AMoD No Transit* in (A) *Auto Sprawl* AND (B) *Auto Innovative*

ridership. Thus, transit integration promises to be a more viable path than pure competition between transit and AMoD, especially for *Auto Innovative* cities such as Boston, Chicago and Toronto. Overall, AMoD results in a reduction of private car usage. Another benefit we find from the presence of AMOD is that it increases active mobility usage in *Auto Sprawl*. People who formerly drove alone, for instance, on switching to AMoD, would now walk for subtours (such as going to lunch at work) as their cars would no longer be available.

3.2 Network level of service

From the high-fidelity supply simulations, we compute the following systemic metrics to assess the impact of AMoD under various scenarios on vehicle kilometers traveled (VKT) and congestion,

(A) *Auto Sprawl*



(B) *Auto Innovative*

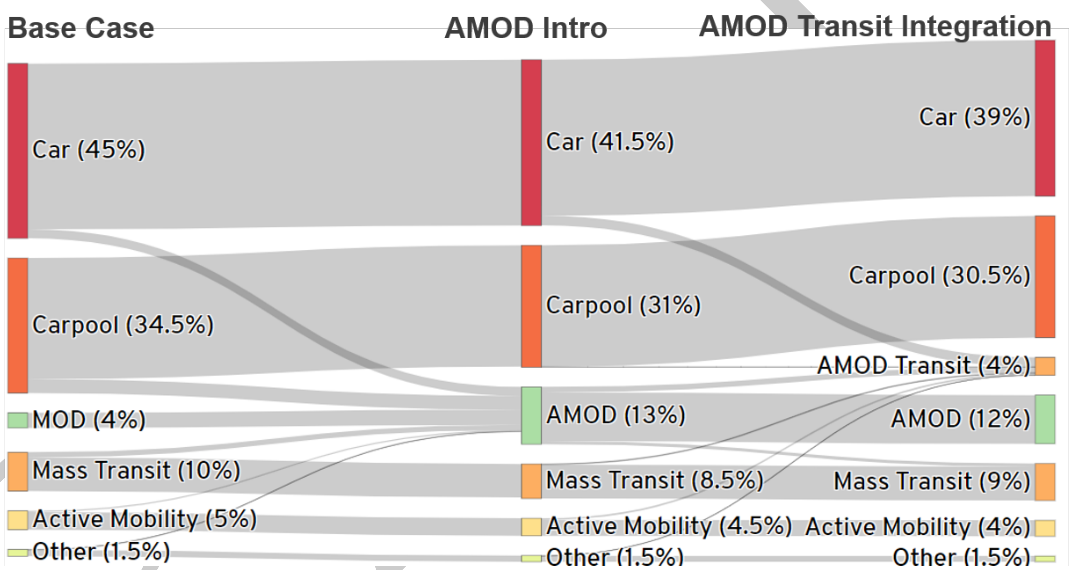


FIGURE 16 Modal shifts from *Base Case* → *AMoD Intro* → *AMoD Transit Integration* in (A) *Auto Sprawl* AND (B) *Auto Innovative*

as measured by the travel time index (TTI).

3.2.1 Vehicle kilometers traveled

The VKT is calculated as the distance traveled by all passenger vehicles (private cars and on-demand fleet) over the course of an entire day (Figure 17). In general, the impacts of AMoD are greater in *Auto Innovative* than in *Auto Sprawl*, due to the differences in density and, consequently, demand. Private car usage decreases more significantly in *Auto Sprawl*, however.

Under *AMoD Intro*, VKT increases by 9% in *Auto Sprawl* and by 26% in *Auto Innovative*. With the removal of transit under *AMoD No Transit*, VKT increases by 14% in *Auto Sprawl*. Under the

Table 12 Trip impacts of the *AMoD Intro*, *AMoD No Transit* and *AMoD Transit Integration* scenarios with respect to *Base Case* in *Auto Sprawl* and *Auto Innovative*.

Scenario	Trip type	*Auto Sprawl*		*Auto Innovative*	
		% Change	No. trips ($\times 10^3$)	% Change	No. trips ($\times 10^3$)
AMoD Intro	Person	2.9	375	1.7	262
	Passenger-vehicle	4	246	4	433
	Mass transit	−21	−77	−13	−211
AMoD No Transit	Person	3.6	345	−0.1	−21
	Passenger-vehicle	5	353	12	1270
	Mass transit	−100	−366	−100	−1581
AMoD Transit Integration	Person	2.5	248	2.3	366
	Passenger-vehicle	0.2	13	2.8	290
	Mass transit	88	322	28	450

same scenario in *Auto Innovative*, the VKT increases by 39%, as on-demand VKT triples. This result further highlights the deleterious impacts under this scenario in a dense auto-dependent city with a mild reliance on mass transit. In the *AMoD Transit Integration* case, VKT increases by 13% in *Auto Sprawl*, which is not significantly different than under *AMoD No Transit*. Similarly, in *Auto Innovative*, VKT increases by 29% under *AMoD Transit Integration* (compared to an increase of 26% under *AMoD Intro*). Thus, integration does not reduce VKT impacts in auto-dependent cities.

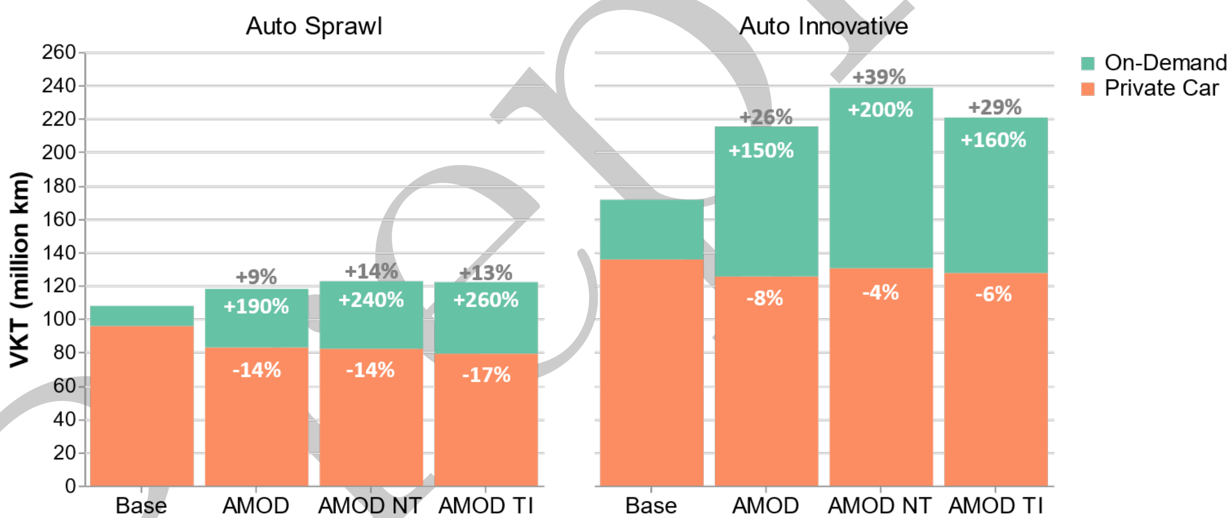


Figure 17 Vehicle kilometers traveled

3.2.2 *Travel Time Index*

In these simulations, the TTI is computed as the distance-weighted average of the ratio of in-simulation trip time to free-flow trip time for all passenger-vehicle trips:

$$TTI = \frac{\sum_{i \in T} d_i \frac{tt_i^{\text{sim}}}{tt_i^{\text{ff}}}}{\sum_{i \in T} d_i}, \tag{3}$$

where d_i, tt_i^{sim} and tt_i^{ff} are the distance, in-simulation time and free-flow time, respectively, for each passenger vehicle trip i within time period T. Thus, in the hypothetical case where the TTI is 1, average daily traffic is similar to free-flow and there is consequently no congestion. A TTI of 1.55 over the entire day would imply that on average, trips would require 55% more time than under free-flow conditions. The TTI realized by time-of-day for both cities is shown in Figure 18. In Figure 19, we depict the TTI trends (averaged over the entire day).

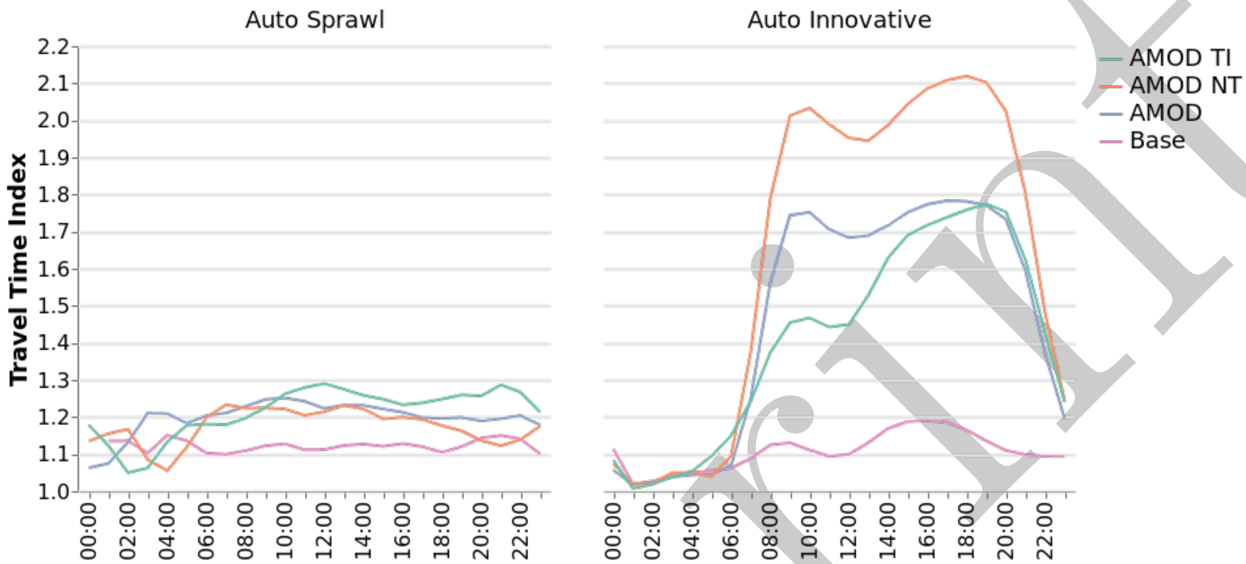


FIGURE 18 Travel time index (distance-weighted average by the hour).

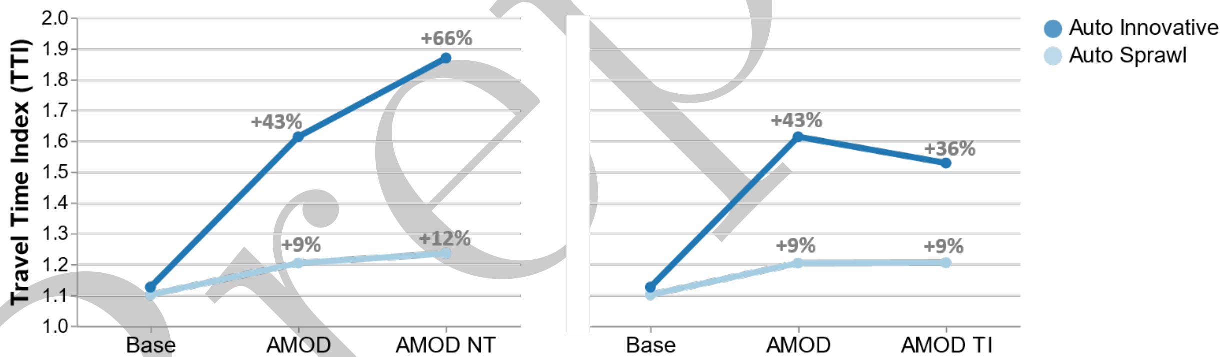


FIGURE 19 Travel time index (distance-weighted average over the entire day).

Under *AMoD Intro*, congestion (as measured by TTI) increases in both cities—by 9% in *Auto Sprawl* and by 43% in *Auto Innovative*. The impact is thus nearly five times greater in *Auto Innovative* due to its density. With the abandonment of transit in the *AMoD No Transit* scenario, congestion increases by 12% in *Auto Sprawl*, which is a mild change compared to the 9% increase under *AMoD No Transit*. In *Auto Innovative*, however, the abandonment of transit leads to unsustainable gridlock, as congestion increases by 66% and travel times on average are nearly double those under free-flow conditions.

Given the limited availability of transit in *Auto Sprawl*, integration does not modify the effects of AMoD on congestion. However, transit integration mitigates the congestion impacts of AMoD in the denser *Auto Innovative*, as the TTI increases by 36% under *AMoD Transit Integration* (compared to 43% under *AMoD Intro*).

In summary, we find that AMoD under any strategy will increase VKT and congestion in either city type. However, *Auto Sprawl* experiences lower impacts. In all three AMoD scenarios in this prototype city, VKT does not increase beyond 14% and congestion does not increase beyond 12%. In *Auto Innovative*, the removal of transit is not viable, as it potentially results in gridlock. Across the AMoD scenarios studied, *Auto Innovative* fares best under *AMoD Transit Integration*. This further suggests that dense metropolitan areas with significant mass transit infrastructure are better off implementing policies that encourage AMoD to complement mass transit in order to mitigate the deterioration of network performance. However, sprawling cities with low transit penetration can afford to not integrate these services while still maintaining a reasonable network performance.

3.3 Energy and emissions

As earlier discussed, the energy is computed for each vehicle during simulation, using the speeds and accelerations in each successive timeframe. For the AMoD scenarios, we have assumed the on-demand fleet is fully electrified. This represents an optimistic environmentally-friendly future where such a regulation is imposed.

Following each simulation, we calculate the well-to-wheels (primary) energy by factoring the production, transmission and distribution losses. In both cities, we use the U.S. average energy-to-fuel ratio of 1.17 and 1.05 for gasoline and diesel, respectively. For electricity, the primary-to-generated energy ratio is 2.99. These ratios were obtained from the analyses conducted by Gençer and O'Sullivan (2019). The simulation output (secondary) energy consumption results are then multiplied by the respective ratios for each fuel to obtain the primary energy consumption:

$$E_j^{\text{pri}} = \sum_{j \in F} \sum_{i \in T} \alpha_j E_{ij}^{\text{sec}} \tag{4}$$

where F is the set of energy sources {gasoline, diesel, electric}, α_j the respective primary energy factors, E_j^{pri} the primary energy in GWh and E_{ij}^{sec} the secondary energy in GWh for each energy source j, each passenger vehicle, bus or train trip i, and over time period T.

For reporting purposes, we consider the primary energy for both gasoline- and diesel-powered vehicles as "Fuel" and that for electric vehicles as "Electricity". The primary energy across all four scenarios in both cities is shown in Figure 20.

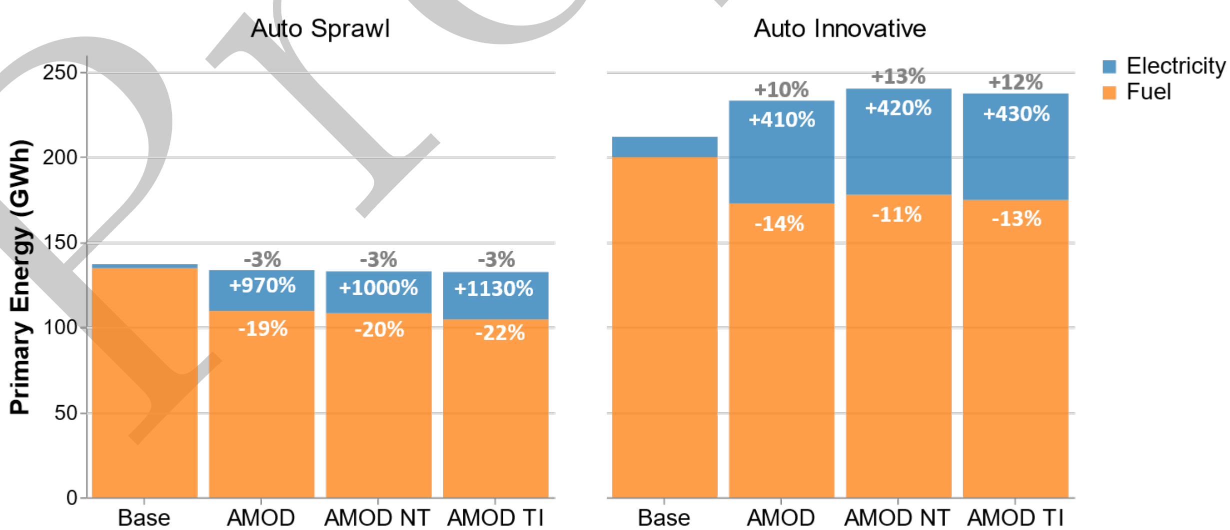


FIGURE 20 Primary (well-to-wheels) energy consumption across the scenarios. In both cities, we assume the AMoD fleet are fully electrified.

Similarly, to obtain the greenhouse gas (GHG) emissions, we multiply primary energy values by U.S. nationwide average energy source-specific GHG emissions intensities (accounting for both generation and usage). The values are 331 gCO_2e/kWh (gasoline intensity) and 438 $gCO2_2e/kWh$ (electricity intensity) (Gençer and O'Sullivan, 2019). We make the assumption that the gasoline intensity is valid for diesel. Thus, we obtain:

$$GHG_j = \sum_{j \in F} \beta_j E_j^{\text{pri}} \tag{5}$$

where β_j are the emissions intensities $\{331, 331, 438\}$ gCO_2e/kWh for gasoline, diesel and electricity, respectively, while GHG_j are the emissions in $MTCO_2e$ for each energy source j. Again, we report gasoline and diesel emissions in the "Fuel" category and electric vehicle-based emissions in the "Electricity" category (Figure 21).

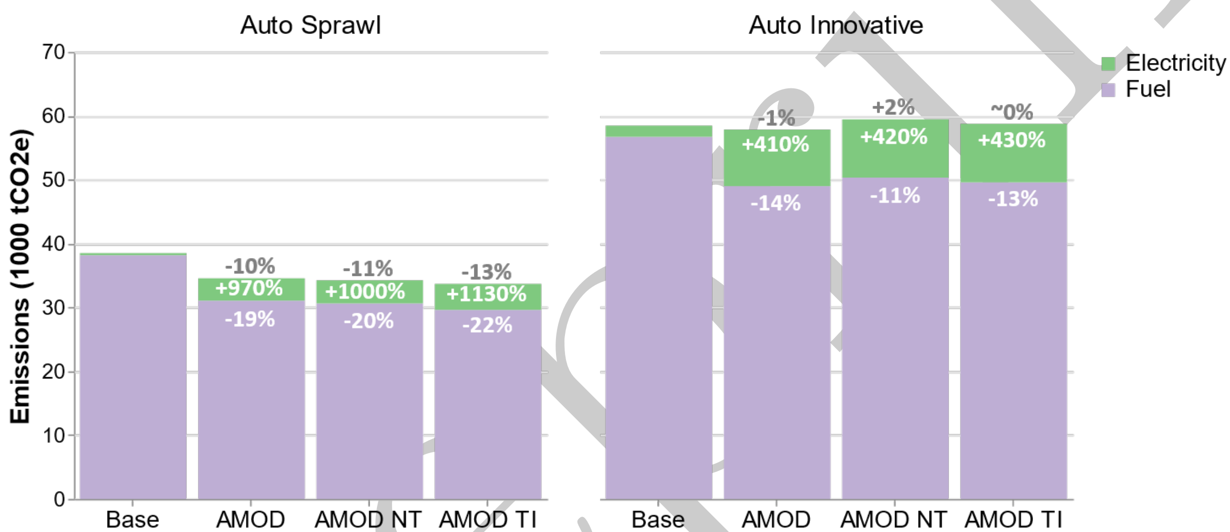


FIGURE 21 Total (well-to-wheels) emissions impacts across the scenarios. In both cities, we assume the AMoD fleet are fully electrified.

Under *AMoD Intro*, primary energy consumption reduces by 3% in *Auto Sprawl*. Conversely, in *Auto Innovative*, energy consumption increases by 10%. In *Auto Sprawl*, energy consumption is similarly impacted under all the AMoD Scenarios—a 3% reduction compared to *Base Case*. However, under *AMoD No Transit*, energy consumption increases by 13% in *Auto Innovative*. Transit integration only slightly increases the energy expenditure in *Auto Innovative* (12%), compared to under *AMoD Intro* (10%).

The trends in GHG emissions are similar to those for energy. In *Auto Sprawl*, emissions are reduced by 10%, 11% and 13% under *AMoD Intro*, *AMoD No Transit* and *AMoD Transit Integration*, respectively. In *Auto Innovative*, the emissions impacts are less significant at -1%, 2% and 0%.

Thus, we see that in the optimistic case of total AMoD electrification, overall energy consumption increases significantly in dense auto-dependent cities. Meanwhile, electrification does lead to energy savings and even greater reductions in emissions in low-density auto-dependent cities. Further reductions in the emissions intensities of electricity generation would improve outcomes in *Auto Innovative* cities.

4 Conclusion

We have developed an approach for generating prototype cities that are representative of urban typologies discovered from recent data mining effort. The prototype city generation methodology consists of the synthesis, modeling and calibration of: (i) population and land-use, (ii) demand, and (iii) road and transit network supply. Using this method, we generated two prototype cities, *Auto Sprawl* and *Auto Innovative*, which represent most cities in the United States: auto-dependent and wealthy. While *Auto Sprawl* is less dense (in terms of population and network) with a higher car mode share, *Auto Innovative* is denser, more populated and has greater mass transit availability and mode share.

Our novel integrated framework allows for high-fidelity large-scale agent-based simulation which incorporates demand-supply integration via "day-to-day learning". Using SimMobility MidTerm, we simulated three automated mobility-on-demand (AMoD) strategies in both of these cities. The *AMoD Intro* scenario addressed a future where AMoD replaces existing MoD services. In *AMoD No Transit*, abandonment of transit was explored in contrast to the *AMoD Transit Integration* scenario, where AMoD was restricted to local circulation and access/egress to mass transit.

The simulation results we obtained indicate that merely introducing AMoD in place of existing ridehailing and ridesourcing will be detrimental to mass transit, as it cannibalizes transit ridership by up to 21%. It also increases VKT by 9% in *Auto Sprawl* but by 26% in *Auto Innovative*. While AMoD reduces the energy consumption and emissions in *Auto Sprawl* under the assumption of full electrification, it increases the energy consumption in *Auto Innovative*. We also found that while AMoD can substitute mass transit in *Auto Sprawl*, it cannot sustainably do so in denser cities, as it would result in gridlock. Ultimately, the impact of AMoD can be mitigated through policy interventions, such as transit integration. Under such a scenario, transit cannibalization is reversed in *Auto Innovative*, while transit ridership is even further boosted in *Auto Sprawl*. Further mitigating effects on congestion are clearly observed in *Auto Innovative*.

Even with significant AMoD-transit integration, AMoD still increases VKT and congestion. From these initial results, it appears that transit removal could be a viable strategy for *Auto Sprawl* cities, as it would have the greatest environmental impact. For *Auto Innovative* cities, integrating transit with AMoD provides the best outcome. With further cost-benefit analyses, however, informed policy recommendations that would be useful to cities in the respective typologies could be obtained.

As earlier noted, several studies have attempted to quantify the demand, congestion, energy and emissions impacts of AMoD under various strategies with mixed results. Our results here indicate that urban form, population and behavior are critical to AMoD outcomes. The simulation of possible scenarios in a representative prototype city can potentially save costs for a metropolitan planning agency considering viable approaches in coming to terms with an imminent AMoD future. While outcomes in a specific city are not guaranteed to be exactly the same as those obtained from a corresponding prototype city, they can be expected to follow a similar pattern.

The strength of our prototype city simulation approach lies in the fact that we can analyze the impacts of various urban mobility scenarios by typology, rather than providing results from a single city that might be irrelevant to other cities. Thus, even further simulations of interest can be examined in a given prototype in order to gain further insight into the impacts of AMoD on cities in the corresponding typology. In future work, we plan to simulate other typologies and also explore other AMoD-related interventions. In order to further validate the archetype city approach of selecting a network from a real city near the typology centroid, we will explore a quantification of the uncertainty of impact of the network structure. We are currently investigating the effects of possible changes in future private vehicle ownership as AMoD becomes more prevalent. Reductions in car ownership can potentially mitigate the VKT impacts of AMoD. However, further concurrent policies might be required to manage network performance for best outcomes. Thus, congestion pricing interventions are also of interest for future research, as these hold promise for effectively tackling congestion and consequent GHG emissions not only in tomorrow's cities but also in those of the present day.

Acknowledgments

This work was supported through the MIT Energy Initiative's Mobility of the Future study. The authors would also like to thank collaborators at the Singapore MIT Alliance for Research and Technology for their technical support in adapting the SimMobility platform for our needs. We are grateful to the anonymous reviewers whose comments and suggestions increased the quality of this manuscript.

References

Adnan, M., Pereira, F.C., Lima Azevedo, C.M., Basak, K., Lovric, M., Raveau, S., Zhu, Y., Ferreira, J., Zegras, C., Ben-Akiva, M.E., 2016. Simmobility: A multi-scale integrated agent-based simulation platform, in: Transportation Research Board 95$^{\text{th}}$ Annual Meeting.

Alonso-Mora, J., Samaranayake, S., Wallar, A., Frazzoli, E., Rus, D., 2017. On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment. Proceedings of the National Academy of Sciences 114, 462–467. URL: http://www.pnas.org/content/114/3/462, doi:10.1073/pnas.1611675114, arXiv:http://www.pnas.org/content/114/3/462.full.pdf.

Azevedo, C.L., Marczuk, K., Raveau, S., Soh, H., Adnan, M., Basak, K., Loganathan, H., Deshmunkh, N., Lee, D.H., Frazzoli, E., et al., 2016. Microsimulation of demand and supply of autonomous mobility on demand. Transportation Research Record: Journal of the Transportation Research Board , 21–30.

Basu, R., Araldo, A., Akkinepally, A.P., Nahmias-Biran, B.H., Basak, K., Seshadri, R., Deshmukh, N., Kumar, N., Azevedo, C.L., Ben-Akiva, M., 2018. Automated mobility-on-demand vs. mass transit: A multi-modal activity-driven agent-based simulation approach. Transportation Research Record 0, 0361198118758630. URL: https://doi.org/10.1177/0361198118758630, doi:10.1177/0361198118758630, arXiv:https://doi.org/10.1177/0361198118758630.

Ben-Akiva, M., 2010. Planning and action in a model of choice, in: Hess, S., Daly, A. (Eds.), Choice Modelling: The State-of-the-Art and the State-of-Practice: Proceedings from the Inaugural International Choice Modelling Conference, Emerald Group Publishing Limited. pp. 19–34.

Bösch, P.M., Becker, F., Becker, H., Axhausen, K.W., 2018. Cost-based analysis of autonomous mobility services. Transport Policy 64, 76–91. URL: http://www.sciencedirect.com/science/article/pii/S0967070X17300811, doi:10.1016/j.tranpol.2017.09.005.

Chen, S., Prakash, A.A., Azevedo, C.L.D., Ben-Akiva, M., 2019. Formulation and solution approach for calibrating activity-based travel demand model-system via microsimulation. URL: https://www.researchgate.net/publication/329781783_Formulation_and_solution_approach_for_calibrating_activity-based_travel_demand_model-system_via_microsimulation. in Review.

Clewlow, R.R., Mishra, G.S., 2017. Disruptive Transportation: The Adoption, Utilization, and Impacts of Ride-Hailing in the United States. resreport UCD-ITS-RR-17-07. Institute of Transportation Studies, University of California, Davis.

Controller, O., 2005. Taxicab industry report: Rates of fare and gate fees. Technical Report. City and County of San Francisco. URL: https://sfcontroller.org/ftp/uploadedfiles/controller/reports/rptTaxiDec2005.pdf.

Dai, L., Derudder, B., Liu, X., 2016. Generative network models for simulating urban networks, the case of inter-city transport network in southeast Asia. Cybergeo: European Journal of Geography doi:10.4000/cybergeo.27734.

Dargay, J., Gately, D., Sommer, M., 2007. Vehicle ownership and income growth, worldwide: 1960-2030. The Energy Journal Volume 28, 143–170. URL: https://EconPapers.repec.org/RePEc:aen:journl:2007v28-04-a07.

EPA, 2018. Fast Facts: U.S. Transportation Sector Greenhouse Gas Emissions, 1990-2016. Technical Report EPA-420-F-18-013. United States Environmental Protection Agency. URL: https://nepis.epa.gov/Exe/ZyPDF.cgi?Dockey=P100USI5.pdf.

Fagnant, D.J., Kockelman, K.M., 2014. The travel and environmental implications of shared autonomous vehicles, using agent-based model scenarios. Transportation Research Part C: Emerging Technologies 40, 1–13. URL: http://www.sciencedirect.com/science/article/pii/S0968090X13002581, doi:10.1016/j.trc.2013.12.001.

Fiori, C., Ahn, K., Rakha, H.A., 2016. Power-based electric vehicle energy consumption model: Model development and validation. Applied Energy 168, 257–268. URL: http://www.sciencedirect.com/science/article/pii/S030626191630085X, doi:10.1016/j.apenergy.2016.01.097.

Fournier, N., Christofa, E., Akkinepally, A.P., Azevedo, C.L., 2020. Integrated population synthesis and workplace assignment using an efficient optimization-based person-household matching method. Transportation , 1–27doi:10.1007/s11116-020-10090-3.

Gençer, E., O'Sullivan, F.M., 2019. A framework for multi-level life cycle analysis of the energy system, in: Kiss, A.A., Zondervan, E., Lakerveld, R., Ŏzkan, L. (Eds.), 29th European Symposium on Computer Aided Process Engineering. Elsevier. volume 46 of *Computer Aided Chemical Engineering*, pp. 763 – 768. URL: http://www.sciencedirect.com/science/article/pii/B9780128186343501284, doi:https://doi.org/10.1016/B978-0-12-818634-3.50128-4.

Goodwin, P., Dargay, J., Hanly, M., 2004. Elasticities of road traffic and fuel consumption with respect to price and income: A review. Transport Reviews 24, 275–292. doi:10.1080/0144164042000181725, arXiv:https://doi.org/10.1080/0144164042000181725.

Hall, J.D., Palsson, C., Price, J., 2018. Is Uber a substitute or complement for public transit? Journal of Urban Economics 108, 36–50. URL: http://www.sciencedirect.com/science/article/pii/S0094119018300731, doi:10.1016/j.jue.2018.09.003.

Jin, S.T., Kong, H., Wu, R., Sui, D.Z., 2018. Ridesourcing, the sharing economy, and the future of cities. Cities 76, 96 – 104. URL: http://www.sciencedirect.com/science/article/pii/S0264275117311952, doi:https://doi.org/10.1016/j.cities.2018.01.012.

Le, D.T., Cernicchiaro, G., Zegras, C., Ferreira, J., 2016. Constructing a synthetic population of establishments for the SimMobility microsimulation platform. Transportation Research Procedia 19, 81–93. URL: http://www.sciencedirect.com/science/article/pii/S2352146516308560, doi:10.1016/j.trpro.2016.12.070. transforming Urban Mobility. mobil.TUM 2016. International Scientific Conference on Mobility and Transport. Conference Proceedings.

Li, S., Tavafoghi, H., Poolla, K., Varaiya, P., 2019. Regulating tncs: Should uber and lyft set their own rules? Transportation Research Part B: Methodological 129, 193 – 225. URL: http://www.sciencedirect.com/science/article/pii/S0191261519300669, doi:https://doi.org/10.1016/j.trb.2019.09.008.

Liang, X., de Almeida Correia, G.H., van Arem, B., 2016. Optimizing the service area and trip selection of an electric automated taxi system used for the last mile of train trips. Transportation Research Part E: Logistics and Transportation Review 93, 115–129. URL: http://www.sciencedirect.com/science/article/pii/S1366554516300552, doi:10.1016/j.tre.2016.05.006.

de Lima, I.V., Danaf, M., Akkinepally, A., Azevedo, C.L.D., Ben-Akiva, M., 2018. Modeling framework and implementation of activity- and agent-based simulation: An application to the Greater Boston Area. Transportation Research Record 2672, 146–157. URL: https://doi.org/10.1177/0361198118798970, doi:10.1177/0361198118798970, arXiv:https://doi.org/10.1177/0361198118798970.

Marczuk, K.A., Hong, H.S.S., Azevedo, C.M.L., Adnan, M., Pendleton, S.D., Frazzoli, E., Lee, D.H., 2015. Autonomous mobility on demand in simmobility: Case study of the central business district in Singapore, in: 2015 IEEE 7$^{\text{th}}$ International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM), pp. 167–172. doi:10.1109/ICCIS.2015.7274567.

Martinez, L.M., Correia, G.H.A., Viegas, J.M., 2015. An agent-based simulation model to assess the impacts of introducing a shared-taxi system: An application to lisbon (Portugal). Journal of Advanced Transportation 49, 475–495. URL: www.scopus.com. cited By :32.

Müller, K., Axhausen, K.W., 2012. Multi-level fitting algorithms for population synthesis. Arbeitsberichte Verkehrs-und Raumplanung 821.

Nahmias-Biran, B., Oke, J.B., Kumar, N., Basak, K., Araldo, A., Seshadri, R., Akkinepally, A., Azevedo, C.L., Ben-Akiva, M., 2019. From traditional to automated mobility on demand: A comprehensive framework for modeling on-demand services in simmobility. Transportation Research Record 0, 0361198119853553. URL: https://doi.org/10.1177/0361198119853553, doi:10.1177/0361198119853553, arXiv:https://doi.org/10.1177/0361198119853553.

Oke, J.B., Aboutaleb, Y.M., Akkinepally, A., Azevedo, C.L., Han, Y., Zegras, P.C., Ferreira, J., Ben-Akiva, M.E., 2019. A novel global urban typology framework for sustainable mobility futures. Environmental Research Letters 14, 095006. URL: https://doi.org/10.1088%2F1748-9326%2Fab22c7, doi:10.1088/1748-9326/ab22c7.

Pavone, M., 2015. Autonomous mobility-on-demand systems for future urban mobility, in: Autonomes Fahren. Springer, pp. 399–416.

Rakha, H.A., Ahn, K., Moran, K., Saerens, B., den Bulck, E.V., 2011. Virginia tech comprehensive power-based fuel consumption model: Model development and testing. Transportation Research Part D: Transport and Environment 16, 492–503. URL: http://www.sciencedirect.com/science/article/pii/S1361920911000782, doi:10.1016/j.trd.2011.05.008.

Scheltes, A., de Almeida Correia, G.H., 2017. Exploring the use of automated vehicles as last mile connection of train trips through an agent-based simulation model: An application to delft, Netherlands. International Journal of Transportation Science and Technology 6, 28–41. URL: http://www.sciencedirect.com/science/article/pii/S2046043016300296, doi:10.1016/j.ijtst.2017.05.004. connected and Automated Vehicles: Effects on Traffic, Mobility and Urban Design.

Shaheen, S., Chan, N., 2016. Mobility and the sharing economy: Potential to facilitate the first-and last-mile public transit connections. Built Environment 42, 573–588.

Shaheen, S., Cohen, A., 2019. Shared ride services in North America: definitions, impacts, and the future of pooling. Transport Reviews 39, 427–442. URL: https://doi.org/10.1080/01441647.2018.1497728, doi:10.1080/01441647.2018.1497728, arXiv:https://doi.org/10.1080/01441647.2018.1497728.

Shaheen, S., Cohen, A., Yelchuru, B., Sarkhili, S., Hamilton, 2017. Mobility on demand operational concept report. Technical Report FHWA-JPO-18-611. United States Department of Transportation. URL: https://rosap.ntl.bts.gov/view/dot/34258.

Shen, Y., Zhang, H., Zhao, J., 2018. Integrating shared autonomous vehicle in public transportation system: A supply-side simulation of the first-mile service in Singapore. Transportation Research Part A: Policy and Practice 113, 125–136. URL: http://www.sciencedirect.com/science/article/pii/S096585641730681X, doi:10.1016/j.tra.2018.04.004.

Small, K., van Dender, K., 2007. Long Run Trends in Transport Demand, Fuel Price Elasticities and Implications of the Oil Outlook for Transport Policy. OECD/ITF Joint Transport Research Centre Discussion Papers 2007/16. OECD Publishing. URL: https://EconPapers.repec.org/RePEc:oec:itfaaa:2007/16-en.

Sperling, D., Gordon, D., 2008. Two billion cars transforming a culture. TR News , 3–9.

Spieser, K., Treleaven, K., Zhang, R., Frazzoli, E., Morton, D., Pavone, M., 2014. Toward a systematic approach to the design and evaluation of automated mobility-on-demand systems: A case study in Singapore, in: Road vehicle automation. Springer, pp. 229–245.

Stephens, T.S., Gonder, J., Chen, Y., Lin, Z., Liu, C., Gohlke, D., 2016. Estimated Bounds and Important Factors for Fuel Use and Consumer Costs of Connected and Automated Vehicles. Technical Report NREL/TP-5400-67216. National Renewable Energy Lab. (NREL), Golden, CO (United States). doi:10.2172/1334242.

Strauch, D., Moeckel, R., Wegener, M., Gräfe, J., Mühlhans, H., Rindsfüser, G., Beckmann, K.J., 2005. Linking transport and land use planning: the microscopic dynamic simulation model ilumass. Geodynamics , 295–311.

TRACE, 1999. Final Report for Publication. Technical Report Contract No: RO-97-SC.2035. European Commission. URL: https://trimis.ec.europa.eu/sites/default/files/project/documents/trace.pdf.

Transport & Environment, 2018. CO2 Emissions From Cars: the facts. Technical Report. European Federation for Transport and Environment AISBL. Brussels, Belgium.

United Nations, Department of Economic and Social Affairs, Population Division, 2018. World Urbanization Prospects: The 2018 Revision. Technical Report. United Nations. URL: https://population.un.org/wup/.

Vazifeh, M.M., Santi, P., Resta, G., Strogatz, S., Ratti, C., 2018. Addressing the minimum fleet problem in on-demand urban mobility. Nature 557, 534.

Wang, J., Rakha, H.A., 2017a. Convex fuel consumption model for diesel and hybrid buses. Transportation Research Record: Journal of the Transportation Research Board , 50–60.

Wang, J., Rakha, H.A., 2017b. Electric train energy consumption modeling. Applied Energy 193, 346–355. URL: http://www.sciencedirect.com/science/article/pii/S0306261917301861, doi:10.1016/j.apenergy.2017.02.058.

Wen, J., Chen, Y.X., Nassir, N., Zhao, J., 2018. Transit-oriented autonomous vehicle operation with integrated demand-supply interaction. Transportation Research Part C: Emerging Technologies 97, 216–234. URL: http://www.sciencedirect.com/science/article/pii/S0968090X18300378, doi:10.1016/j.trc.2018.10.018.

Zhou, E., McGlaughlin, A., Turan, D., 2015. Generating synthetic road networks from various reduced dimension representations.

Appendix A Population and Land-use synthesis

The tools we have created for generating a prototype city are available upon request at https://github.com/jimioke/virtual-city-generator. The outputs are formatted to SimMobility specifications.

Appendix A.1 Assignment of household and work/education locations

We aggregate land use categories into these: low residential (L), high residential (H), commercial (C), industrial (I), education (E), open land (O). More categories can be used if available, but the above is the highest level specification in our generalized approach.

Households w are then assigned as follows:

1. Allocate weights:[11]

$$HH(w_L, w_H, w_C, w_I, w_E, w_O) = (8, 10, 4, 1, 0, 0) \tag{A.1}$$

2. Grid the map and assign weights to each cell w_c^{HH} given its prevailing land use category

3. Normalize cell weights $p_{c(SAL)}^{HH}$ in each second administrative level (SAL)

$$p_{c(SAL)}^{HH} = \frac{w_c^{HH}}{\sum_{c \in SAL} w_c^{HH}} \tag{A.2}$$

[11]We note that if totals for work, education and households by TAZ or other level are available, the weights do not have to be arbitrarily assigned. For the numbers shown (*Auto Sprawl*), we used a linear program to obtain the household and employment weights, given that the unit totals were available.

4. Number of households in each cell given by:

$$N_{HH}^{c(SAL)} = p_c^{SAL} \cdot N_{HH}^{SAL} \tag{A.3}$$

where N_{HH}^{SAL} is the number of households in each SAL

5. Randomly sample to locate households in cell centroids for all SAL, while controlling for SAL totals.

Work and education allocation to grid cells in network:

1. We obtain the numbers of firms and schools in each SAL then assign weights as follows:

$$WORK(w_L, w_H, w_C, w_I, w_E, w_O) = (1, 2, 10, 5, 3, 1) \tag{A.4}$$
$$EDU(w_L, w_H, w_C, w_I, w_E, w_O) = (0, 0, 0, 0, 1, 0) \tag{A.5}$$

2. Assign weights to cells for work and education: w_c^{WORK}, w_c^{EDU}

3. Find $p_{c(SAL)}^{WORK}$ and $p_{c(SAL)}^{EDU}$ as before

4. Find $N_{WORK}^{c(SAL)}$ and $N_{EDU}^{c(SAL)}$ as before while controlling for total student and worker totals.

The partitioning of *Auto Sprawl* is shown in Figure A.22 and Figure C.25. The education locations are shown in Figure A.24.

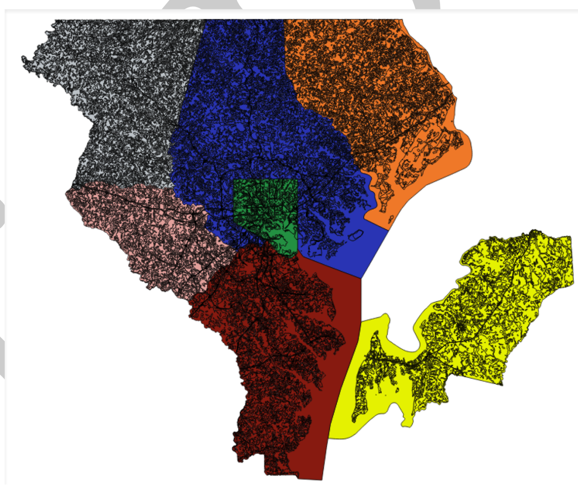


FIGURE A.22 Second Administrative Levels in *Auto Sprawl*

Appendix A.1.1 Cost estimates

There are three key inputs at the zonal level required for running SimMobility *PreDay*: zonal attributes (size, location, points of interest, etc.) and zone-zone travel time and cost data (skim matrices). Estimates of these were obtained for each archetype city. The operating cost and travel time variables in the skim matrices are described below in Table B.17. In cases were these data are not publicly available, we can readily estimate values from the network attributes.

FIGURE A.23 Gridding *Auto Sprawl* for household, employment and education allocation

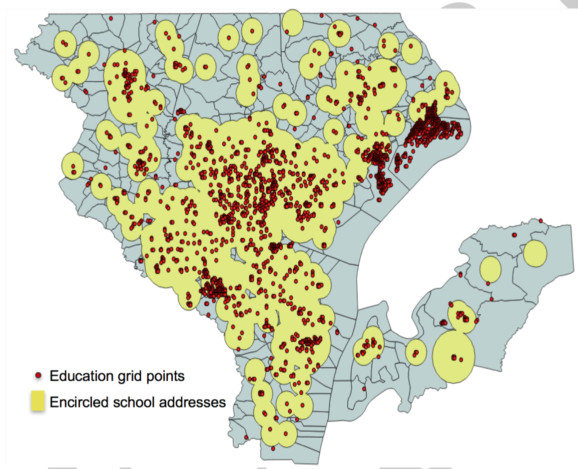


FIGURE A.24 Education allocation for *Auto Sprawl*

Parameters	**Description**	**Notes**
`distance`	Zone-zone distance (km)	
`car_cost_erp`	Road pricing cost (monetary units)	
`car_ivt`	Car in-vehicle time (hours)	
`pub_ivt`	Public transit in-vehicle time (hours)	
`pub_walkt`	Public transit access-egress walking time (hours)	
`pub_wtt`	Public transit waiting time (hours)	
`pub_cost`	Public transit cost between zones (monetary units)	
`avg_transfer`	Average number of public transit transfers between zones	
`pub_out`	Public transit out-of-vehicle travel time (hours)	Sum `pub_walkt` and `pub_wtt`

TABLE A.13 Operating cost parameters (Skim matrix elements) in SimMobility

Appendix B PreDay Models

Appendix B.0.1 Example: Day Pattern Binary Choice Model (dpb)

The Day Pattern Binary model determines if the individual makes any tours in a given day. The model takes the form of a binary logit, where the choices are either to travel and to stay at home. It takes personal characteristics of the individual and the inclusive value of the Day Pattern Tours model as inputs. A decision to travel will lead to Day Pattern Tours model. The utility specification of the *dpb* model is specified as Equation B.1.

$$
\begin{aligned}
&V_{not\ travel} = 0 \\
&V_{travel} = \beta_{ASC} + \beta_{g,m} S_g + \sum_{i=1}^{6} \beta_{emp,i,m} S_{emp,i} + \sum_{i=1}^{5} \beta_{edu,i,m} S_{edu,i} + \sum_{i=1}^{6} \beta_{age,i,m} S_{age,i} + \beta_I I_{dpt}
\end{aligned}
\tag{B.1}
$$

Variable Name	**representation**	**Domain**
Travel time of mode m	T_m	continuous
Cost of mode m	C_m	continuous
Gender	S_g	binary
Employment		
full-time worker	$S_{emp,1}$	binary
part-time worker	$S_{emp,2}$	binary
retired	$S_{emp,3}$	binary
disabled	$S_{emp,4}$	binary
homemaker	$S_{emp,5}$	binary
unemployed	$S_{emp,6}$	binary
Type of student		
preschool	$S_{edu,1}$	binary
Kindergarten - 8^{th} grade	$S_{edu,2}$	binary
9^{th} grade - 12^{th} grade	$S_{edu,3}$	binary
undergraduate	$S_{edu,4}$	binary
graduate	$S_{edu,5}$	binary
Age category		
under 20	$S_{age,1}$	binary
20 - 25	$S_{age,2}$	binary
25 - 35	$S_{age,3}$	binary
36 - 50	$S_{age,4}$	binary
51 - 65	$S_{age,5}$	binary
65 +	$S_{age,6}$	binary
Logsum from the day pattern tour model	I_{dpt}	continuous

Table B.14 Day Pattern Binary Choice Model Parameters

Appendix B.0.2 Example: Work Tour Mode Choice Model (tmw)

The utility specification of the *tmw* model is specified as Equation B.2. The subscript m stands for the available modes in the choice set. The general availability are set according to scenarios as introduced in Table 9. Private bus is only available to students and workers to their usual workplaces (if any) in Central Business District. Individual specific availability of driving alone and motorcycle are set according to vehicle ownership and license.

$$V_m = \beta_{ASC,m} + \beta_{T,m} T_m + \beta_{C,m} C_m + \beta_{g,m} S_g + \beta_{inc,m} S_{inc} + \beta_{tran,m} S_{tran} + \sum_{i=1}^{6} \beta_{emp,i,m} S_{emp,i} + \sum_{i=1}^{5} \beta_{edu,i,m} S_{edu,i} + \sum_{i=1}^{4} \beta_{veh,i,m} S_{veh,i} \tag{B.2}$$

Variable Name	**Symbol**	**Domain**
Travel time of mode m	T_m	continuous
Cost of mode m	C_m	continuous
Gender	S_g	binary
Household income	S_{inc}	continuous
Transit card	S_{tran}	binary
Employment		
full-time worker	$S_{emp,1}$	binary
part-time worker	$S_{emp,2}$	binary
retired	$S_{emp,3}$	binary
disabled	$S_{emp,4}$	binary
homemaker	$S_{emp,5}$	binary
unemployed	$S_{emp,6}$	binary
Type of student		
preschool	$S_{edu,1}$	binary
Kindergarten - 8^{th} grade	$S_{edu,2}$	binary
9^{th} grade - 12^{th} grade	$S_{edu,3}$	binary
undergraduate	$S_{edu,4}$	binary
graduate	$S_{edu,5}$	binary
Number of household vehicles		
No vehicle	$S_{veh,1}$	binary
1 vehicle	$S_{veh,2}$	binary
2 vehicle	$S_{veh,3}$	binary
3 and 3+ vehicle	$S_{veh,4}$	binary

Table B.15 Work Tour Mode Choice Model Parameters

Appendix B.0.3 Example: Shopping Tour Mode Destination Choice Model (tmds)

The utility specification of the *tmds* model is specified as Equation B.3. The subscript m stands for the available modes in the choice set and d stands for the available destination in the choice set. At most there would be (*number of modes* $\times$ *number of TAZ*) alternatives in the choice set. The general availability of modes are set according to scenarios as introduced in Table 9. Private bus is not available. Individual specific availability of driving alone and motorcycle are set according to vehicle ownership and license. Mass transit is set according to the network.

$$V_{m,z} = \beta_{ASC,m} + \beta_{T,m} T_m + \beta_{C,m} C_m + log(Z_{area,d} + \beta_{emp} Z_{emp,d} + \beta_{pop} Z_{pop,d}) \tag{B.3}$$

Appendix B.0.4 Example: Intermediate Stop Generation (isg)

The utility specification of the *isg* model is specified as Equation B.4. The Intermediate Stop Generation model is a nested logit quit/no-quit model, whereby a no-quit choice results in a new

Variable Name	Symbol	Domain
Travel time of mode m	T_m	continuous
Cost of mode m	C_m	continuous
Area of zone d	$Z_{area,d}$	continuous
Employment of zone d	$Z_{emp,d}$	integer
Population of zone d	$Z_{pop,d}$	integer

Table B.16 Shop Tour Mode Destination Choice Model Parameters

intermediate stop. While one of the nests includes only the quit option, the other includes the other available activity purposes. Individual characteristics, tour purpose, and remaining time window—determined based on the start or end time of the primary activity, and preceding or successive stops—are included as variables in the model. Availability of stop purposes is determined based on the outputs of the Day Pattern Stops model. Stops are scheduled sequentially. V_{work} is an example of the structural utility of performing a work activity. On the other hand, V_{quit} is the structural utility of stopping intermediate stops and is set to zero for reference.

$$\begin{aligned} &V_{quit} = 0 \\ &V_{work} = \beta_{ASC,work} + \beta_{g,work} S_g + \beta_{in,work} Z_{in} Z_{window} + \beta_{out,work} Z_{out} Z_{window} + \\ &\quad \sum_{i=1}^{3} \beta_{in,stop,i} Z_{in} Z_{stop,i} + \sum_{i=1}^{3} \beta_{out,stop,i} Z_{out} Z_{stop,i} + \sum_{i=1}^{4} \beta_{prim,i,work} Z_{prim,i} \end{aligned} \tag{B.4}$$

Variable Name	Symbol	Domain
Gender	S_g	binary
Whether the half tour is inbound	Z_{in}	binary
Whether the half tour is outbound	Z_{out}	binary
Available time window	Z_{window}	continuous
The stop to be decided would be		
the first stop	$Z_{stop,1}$	binary
the second stop	$Z_{stop,2}$	binary
the third stop	$Z_{stop,3}$	binary
Primary activity of the tour		
work	$Z_{prim,1}$	binary
education	$Z_{prim,2}$	binary
shop	$Z_{prim,3}$	binary
other	$Z_{prim,4}$	binary

Table B.17 Work Tour Mode Choice Model Parameters

Appendix C Day-to-day learning

In this section, we explain how we calibrate link travel time information in detail. Our simulator, SimMobility, uses two travel time tables during simulation: default link travel time and historical link travel time. The default link travel time provides link travel time calculated by free flow speed and historical link travel time, including travel time at a 5-min interval, is from previous simulation result. SimMobility will try to look for travel time information at a certain interval from historical link travel time first. If it cannot find historical link travel time, it will use free flow travel time from default link travel time. The problem with historical link travel time is that its information is not necessarily accurate for current demand. For some high congestion links, their travel time in historical link travel time table may be very low, which will cause more people choosing routes containing those links and thus resulting in gridlock.

The purpose of the day-to-day learning process is to update historical link travel times. The historical link travel times are an input to the supply simulation and are used in the private traffic and public transit route-choice models. Having accurate and representative historical link travel times is important for the simulation results to match the actual observations.

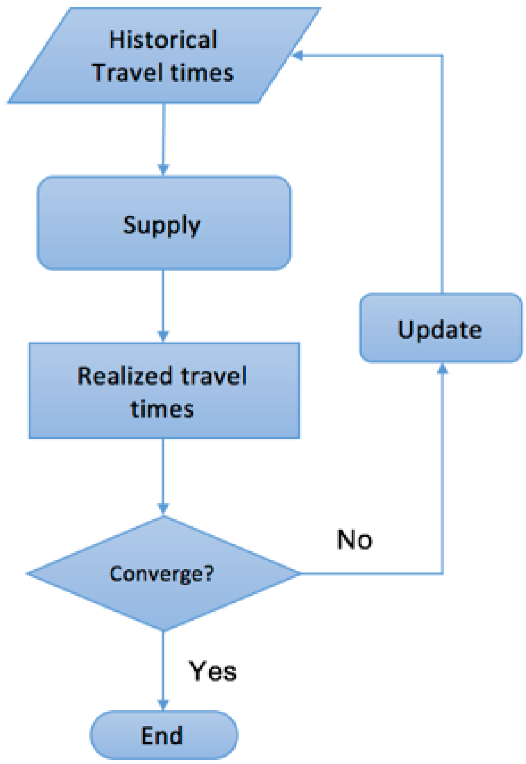


FIGURE C.25 The workflow of day-to-day learning. We first do a simulation with current historical link travel time table and get realized link travel time table as result. The new estimate of historical link travel times are obtained via a weighted combination of the initial historical link travel times and realized link travel times. Specifically, we compare two tables, weight the realized link travel time as 0.25 and initial historical link travel time as 0.75 to calculate a new estimate of the historical link travel time. Finally, we upload this new estimate of the historical link travel time table to the database to replace the older historical link travel time table and then start the next iteration. This process is repeated until the link travel times from the simulation are consistent with the historical link travel times.