Transportation Research Part C 97 (2018) 216–234

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

Transit-oriented autonomous vehicle operation with integrated demand-supply interaction☆

Jian Wen[1], Yu Xin Chen[1], Neema Nassir, Jinhua Zhao*

Massachusetts Institute of Technology, Cambridge, MA 02139, United States

A R T I C L E I N F O

Keywords:
Shared mobility-on-demand system
Autonomous vehicles
Public transportation
Agent-based simulation
Mode choice model

A B S T R A C T

Autonomous vehicles (AVs) represent potentially disruptive and innovative changes to public transportation (PT) systems. However, the exact interplay between AV and PT is understudied in existing research. This paper proposes a systematic approach to the design, simulation, and evaluation of integrated autonomous vehicle and public transportation (AV + PT) systems. Two features distinguish this research from the state of the art in the literature: the first is the transit-oriented AV operation with the purpose of supporting existing PT modes; the second is the explicit modeling of the interaction between demand and supply.

We highlight the transit-orientation by identifying the synergistic opportunities between AV and PT, which makes AVs more acceptable to all the stakeholders and respects the social-purpose considerations such as maintaining service availability and ensuring equity. Specifically, AV is designed to serve first-mile connections to rail stations and provide efficient shared mobility in low-density suburban areas. The interaction between demand and supply is modeled using a set of system dynamics equations and solved as a fixed-point problem through an iterative simulation procedure. We develop an agent-based simulation platform of service and a discrete choice model of demand as two subproblems. Using a feedback loop between supply and demand, we capture the interaction between the decisions of the service operator and those of the travelers and model the choices of both parties. Considering uncertainties in demand prediction and stochasticity in simulation, we also evaluate the robustness of our fixed-point solution and demonstrate the convergence of the proposed method empirically.

We test our approach in a major European city, simulating scenarios with various fleet sizes, vehicle capacities, fare schemes, and hailing strategies such as in-advance requests. Scenarios are evaluated from the perspectives of passengers, AV operators, PT operators, and urban mobility system. Results show the trade off between the level of service and the operational cost, providing insight for fleet sizing to reach the optimal balance. Our simulated experiments show that encouraging ride-sharing, allowing in-advance requests, and combining fare with transit help enable service integration and encourage sustainable travel. Both the transit-oriented AV operation and the demand-supply interaction are essential components for defining and assessing the roles of the AV technology in our future transportation systems, especially those with ample and robust transit networks.

☆ This article belongs to the Virtual Special Issue on "AI in ATM".

* Corresponding author.
E-mail address: jinhua@mit.edu (J. Zhao).

[1] Both authors contributed to this paper equally.

https://doi.org/10.1016/j.trc.2018.10.018
Received 15 January 2018; Received in revised form 19 October 2018; Accepted 21 October 2018
Available online 01 November 2018

1. Introduction

The autonomous vehicle (AV) is poised to be one of the most disruptive technologies in the transportation industry. The forthcoming adoption and commercialization of AVs are expected to have extensive impacts on our road networks, congestion, safety, land use, and more. The urban public transportation (PT) system is not an exception.

In the past few years, the online-enabled mobility-on-demand services have attracted considerable mode share in hundreds of cities around the world, substituting private cars, taxis, and even buses and trains (Rayle et al., 2016; Hall et al., 2017). The convenience of hailing from a smartphone, ease of transactions, and economic efficiency of crowd-sourcing the rides have made these services very attractive to consumers. It is anticipated that the AV technology may further improve the economics of such services by reducing the operational costs, and therefore adversely impact the ridership, revenue and general viability of public transportation systems (Smith, 2012; Anderson et al., 2014; Gruel and Stanford, 2016). Such impact can be considered a threat to not only PT itself, but also the social values that bus and rail protect in our societies, such as equity, accessibility, and environmental sustainability.[2]

To identify and create synergistic opportunities between AV and PT, and to plan, regulate and leverage the AV technologies towards supporting the PT systems is essential for cultivating an affordable, equitable, sustainable and efficient urban mobility system in the future. However, the extent and nature of AV's impacts on PT are still uncertain, and the interplay between AV and PT is also understudied. Existing research on autonomous mobility-on-demand (AMoD) systems (Fagnant and Kockelman, 2014, 2015; Marczuk et al., 2015; Martinez et al., 2015; Kloostra and Roorda, 2017; Zhu and Kornhauser, 2017) have provided little insight into the future of PT systems. Studies on integrated AV + PT solutions have just begun very recently (Liang et al., 2016; Vakayil et al., 2017; Shen et al., 2018).

This paper proposes a systematic approach to the design, simulation, and evaluation of integrated autonomous vehicle and public transportation (AV + PT) systems. It also illustrates and evaluates concrete scenarios of the integrated AV + PT service designs in a case study based on an agent-based simulation platform. Two features distinguish this research from the state of the art in the literature: the first is the transit-oriented AV operation with the purpose of supporting existing PT modes; the second is the explicit modeling of the interaction between demand and supply.

We highlight the transit-orientation by identifying the synergistic opportunities between AV and PT, which makes AVs more acceptable by all the stakeholders and respects the social-purpose considerations such as maintaining service availability and ensuring equity. Specifically, in this research AV serves first-mile connections to rail stations and provides efficient and affordable shared mobility (door-to-door) in low-density suburban areas that are typically inefficient to serve by conventional fixed-route PT services.

The interaction between demand and supply is modeled using a set of system dynamics equations and solved as a fixed-point problem through an iterative procedure. We develop an agent-based simulation platform of service and a discrete choice model of demand as two subproblems. Using the level of service as a feedback measure between supply and demand, the system enables interaction between the decisions of the service operator and those of the travelers to model the choices of both parties. The simulation platform is capable of incorporating critical operational decisions including fleet sizes, vehicle capacities and sharing policies, fare schemes, and hailing strategies such as in-advance and on-demand requests. It also includes a heuristic method to optimize the request-vehicle assignment and an optimal rebalancing policy. For demand estimation, a nested logit model is utilized to estimate the mode shares based on modal attributes, with the AV mode attributes simulated in the agent-based simulation model.

The design decisions for AV + PT reflect the interest of multiple stakeholders in the system. We test our approach in a major European city, simulating scenarios with various fleet sizes, vehicle capacities, fare schemes, and hailing strategies such as in-advance requests; and evaluating them from the perspectives of passengers, AV operators, PT operators, and urban mobility system as a whole.

The remainder of the paper is organized as follows. Section 2 presents a literature review and identifies the research gaps. Section 3 presents the AV + PT service design in three aspects: operating mode, fleet management, and fare policy, and outlines the key stakeholders, their interests and the associated performance indicators for each. Section 4 presents the methodology for formulating the supply-demand interaction as well as the demand prediction and agent-based simulation tools. Section 5 discusses the results and Section 7 concludes the paper.

2. Literature review

Spieser et al. (2014) are among the first to conceptualize the shared AMoD system as an enabling technology for future urban mobility. Based on the analytical models, they prove that, ideally, shared AMoD would reduce the total number of vehicles in the system to one third. Agent-based simulation has recently become popular in AMoD research for its advantages in capturing individual behaviors, enabling dynamic operations and accounting for stochasticity. The successive works by Fagnant and Kockelman (2014, 2015) are representative for agent-based simulation applications, in which issues such as dynamic ride-sharing, fleet sizing, and operational costs have been discussed from the perspective of operators. Similar simulation framework could also be found in the applications in Singapore (Marczuk et al., 2015), Lisbon (Martinez and Viegas, 2017; Martínez et al., 2017; Vasconcelos et al., 2017), Toronto (Kloostra and Roorda, 2017) and New Jersey (Zhu and Kornhauser, 2017). The scale and scope of research are being expanded gradually to include emissions, congestion and parking implications.

[2] In this paper, the scope of PT includes principally bus and rail. We do not consider taxis a part of public transportation for the reasons that they do not represent the spirit of equity and sustainability.

The dispatching efficiency of real-time shared AMoD systems is critical to the operation. Jung et al. (2016) propose a hybrid simulated annealing algorithm to dynamically assign requests to vehicles. Levin (2017) studies the same problem while considering congestion, in which traffic flow is modeled through the Link Transmission Model. However, ride-sharing is not allowed in this study. Levin et al. (2017) and Rossi et al. (2017) also develop agent-based simulation models for AMoD services with embedded traffic congestion modules. Ma et al. (2017) formulate a linear programming model to efficiently solve the optimal AV assignment problem and fleet size determination for AV rental services, but they do not consider shared rides. Liang et al. (2018) present an integer programming model for vehicle assignment with a rolling horizon scheme, which is able to handle both in-advance requests and on-demand ones. The model is then tested using a case study in Delft. Alonso-Mora et al. (2017) devise a more general mathematical model for optimal assignment and argue that ideally taxi fleet size in New York City could be reduced by 75% when a shared AV fleet with a capacity of four riders serves the travel demand.

Zhang and Pavone (2016) also focus on New York City taxis. Many of their efforts have been made to solve an optimal rebalancing problem. Jorge et al. (2014) formulate an online rebalancing policy with the objective of maximizing profit and prove it effective using Lisbon travel data. Marczuk et al. (2016) then test the optimal rebalancing problem in Singapore. The results show that 23% fewer vehicles are required to guarantee the same level of service when online rebalancing is in use. Deng and Cardin (2018) integrate rebalancing strategies into the planning of a station-based MoD system. Numerical results demonstrate that rebalancing is imperative to reduce the parking spots required for the system. However, it is also worth noting that rebalancing could induce more vehicle distance by design. Vasconcelos et al. (2017) report that, although the added travel distance does not necessarily impair the profitability, it does have negative environmental impact with regard to CO_2 and NO_x and emissions.

The studies above demonstrate the potential productivity of shared AMoD systems. Meanwhile, some other researchers take a different perspective by investigating travel behavior from the demand side. Panagiotopoulos and Dimitrakopoulos (2018) develop a technology acceptance modelling process for the AV technology based on a stated preference survey. Their results show that perceived usefulness, ease to use, trust and social influence are significant predictors for AV acceptance, but perceived usefulness has the strongest impact. Krueger et al. (2016) develop a stated choice model and estimated users' willingness to pay for shared autonomous vehicles and realized that service attributes including travel cost, travel time and waiting time may be critical determinants of the demand. Levin and Boyles (2015) adopt the four-step model to include non-shared AV as a competing mode. A nested logit model is used to predict the mode shares. The results indicate that AV trips will sharply increase while transit ridership declines and road congestion increases. Childress et al. (2015) use activity-based travel demand simulation and reach the same conclusion. Chen and Kockelman (2016) and Qiu et al. (2018) argue that, if shared AV is used instead and pricing strategies are designed deliberately, AMoD could capture significant market share to be profitable without inducing extra traffic.

As shared AMoD grows, it will take some of the market share of public transportation, unless planned on a mutually complementary basis. The idea of integrated AV + PT systems is first illustrated by Lenz and Fraedrich (2016) as "broadening service options of public transport" by providing multimodal service in less dense areas. Liang et al. (2016) use integer programming models to study AV as a last-mile connection to train trips. Vakayil et al. (2017) then develop an AV + PT hybrid system and emphasize its potential for reducing total vehicle miles traveled and the corresponding negative externalities such as congestion and emissions. Shen et al. (2018) use agent-based simulation to explore the idea of supporting bus operation and planning with AV service as a complement. In their paper, high-demand bus routes are preserved while low-demand ones are re-purposed and shared AV comes in as an alternative. Results indicate that the integrated system would benefit both AV and PT operators. Yan et al. (2018) develop a joint stated and revealed preference choice model and studied the acceptance of integrated ride-sharing service (such as Uber and Lyft) with public transit system. They included ride-sharing for first and last miles to transit and also for replacement of low-demand routes. They realized that ride-sharing service integration leads to best results for first and last mile to transit. Yap et al. (2016) are the first to survey the preference of travelers regarding the integrated AV + PT system. Based on the mode choice model, they successfully estimate the sensitivity of travelers towards different services.

However, none of the existing papers incorporate the demand estimation into the AV + PT simulation, nor do they model the interaction between demand and supply as traveler behavior changes in reaction to the level of service change. Also, existing works apply generic AV system design and ignore the uniqueness of transit system and its social-purpose considerations, such as maintaining service availability and ensuring equity, accessibility and affordability.

3. Service design and evaluation framework

When AV and PT are integrated, the AV service design can embrace a broad set of societal objectives such as focusing on system-level performance, enhancing environmental sustainability, improving livability, and caring for the connectivity of all citizens (Freemark et al., 2018). A variety of organizational structures can be imagined with respect to the ownership of the transit and AV operators, the interaction between the operators, and the degree of regulation and intervention from the public authority (Shen et al., 2018).

This paper focuses on the operation design, but we make the following assumptions on the institutional settings: The AV + PT service is either directly operated by the public authority or contracted with and regulated by the public authority, including the service areas, the level of service, and the fare and ticketing. Overall the AV + PT service provider acts in the interest of the overall mobility system. Refer to (Shen et al., 2018) for a detailed discussion of the institutional settings and organizational relationships among the AV technology provider, AV operator, PT systems, and travelers in an integrated AV + PT system.

The section below summarizes the AV + PT service design in three parts: operating mode, fleet management, and fare policy.

The operating mode includes two decisions: sharing policy and hailing policy. The sharing policy involves determining whether sharing is optional and how rides are paired. These decisions affect the system efficiency at large, and the level of service as experienced by the users, such as wait time and travel time. To reflect the transit-orientation and since the AV portion is regarded as part of the public transit service, we impose that trips are shared whenever doing so improves the system performance, i.e., passengers by default should expect sharing the vehicles with other passengers just like when they board a bus or a train. The hailing policy includes the choice between on-demand and in-advance requests. On-demand services are desirable for users because it provides flexibility both in time and space. However, the dispatching of vehicles would be more efficient and economical if the requests are known in advance. Therefore, there is a trade-off between service flexibility and operational efficiency that can be captured in service design via hailing policy decisions. In addition to the two operating modes reflected in the proposed framework, other operational decisions can also be incorporated and tested similarly to achieve other social goals. For example, to guarantee that services will be provided to all and especially the disadvantaged, para-transit services can also be incorporated as another operating mode.

The fleet management decisions include fleet size, vehicle capacity and dispatching strategies of the AV service. The choice of vehicle capacity and fleet size can dramatically impact system performance as well as operational cost. Traditional operators tend to use larger vehicles such as buses to minimize driving cost. The AV technology eliminates the cost of driving, therefore enables the use of smaller vehicles for more flexible service when compared to buses. It also raises the question of whether door-to-door service or trunk service (or a mix) would be more beneficial. AV provides another advantage over human drivers because it fully complies with dispatching directions, which is not always the case with human drivers. As a result, it is possible to improve overall efficiency through the implementation of sophisticated real-time dispatching algorithms that optimize the request-vehicle assignment, routing, and re-balancing at the system level.

Fare policy is a powerful tool to achieve various efficiency, equity and financial goals. On one hand, the farebox revenue is important to cover the operational cost of the AV fleet. On the other hand, pricing can be used for demand management to promote sustainable modes of transport or achieve system objectives such as accessibility and mobility equity (Stuntz et al., 2017). In this paper, we use a cost-based fare structure including three components: base fare, per-unit-distance fare, and per-unit-time fare. To leverage pricing and promote sustainable travel behavior, a sharing discount and a transit transfer discount are also considered. A base price is also implemented with the objective to discourage short trips usually made by active modes like walking and cycling. Recent research has explored the dynamic service pricing as a demand management leverage for mobility on demand and carsharing services, which is found to be very effective in improving certain system performance metrics (Chen and Kockelman, 2016; Jorge et al., 2015; Xu et al., 2018). It is beyond the scope of this paper to discuss dynamic pricing, but it is identified as one avenue for future research.

We model and evaluate each of the three parts in the AV + PT service design using the proposed simulation tool with the demand and supply interaction.

The design decisions reflect the interest of multiple stakeholders in the system. Table 1 identifies four key stakeholders: travelers, AV operators, PT operators, public authority; and lists the key evaluation indicators from their perspectives. This list does not intend to enumerate the comprehensive list of interests for all stakeholders.

For travelers, the level-of-service indicators include availability and total travel time. Availability is represented by service rate, the percentage of travelers being served given their time constraints. Total travel time consists of wait time and in-vehicle travel time. The in-vehicle travel time, when shared, is proportional to the detour factor, defined as the ratio of actual in-vehicle travel time when

Table 1
Stakeholders, interests, metrics and the indicators used in this paper.

Stakeholder	Interests	Aspects	Indicators
Traveler	Level of service	Availability	Service rate
		Total travel time	Wait time
			Detour factor
	Travel cost	Fare	Fare
AV Operator	Financial viability	Cost	Vehicle distance traveled
			Distance-based average load
		Revenue	AV mode share
PT Operator	Performance	Ridership	PT mode share
		Availability	N/A
		Punctuality	N/A
	Financial viability	cost/revenue	N/A
Public Authority	Public equity	Availability	Service rate
		Accessibility	N/A
	Sustainability	Motorized traffic	Vehicle distance traveled
		Non-motorized trips	Active mode share

“N/A” implies indicators not applicable in the scope of this simulation. These indicators (and those omitted in this table) are important and will be explored more in-depth in following studies.
We assume PT operator to be not profit-driven.

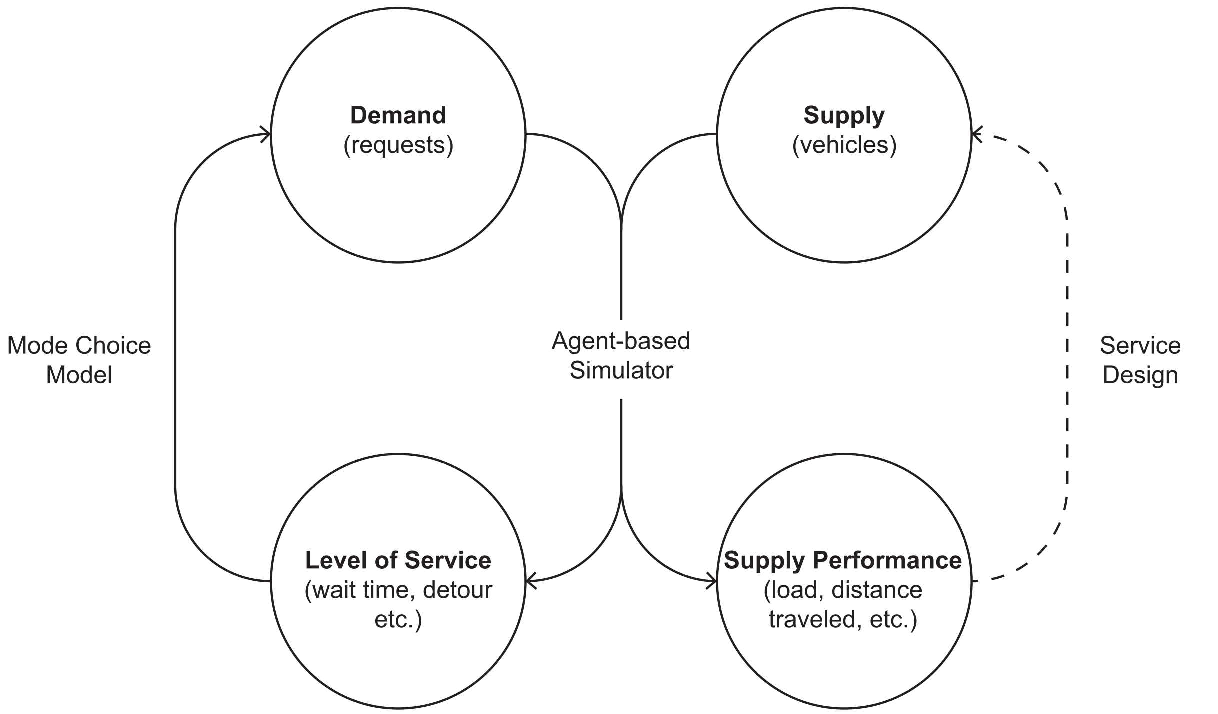


Fig. 1. The demand-supply interaction mechanism.

sharing a ride to the shortest travel time when riding alone. As for the AV operator, the supply performance is evaluated by the operational cost. Specifically, the cost of operating one vehicle is a function of the vehicle distance traveled, including both the service and re-balancing distance. Another indicator related to supply performance is the distance-based average load: the average number of travelers on board weighted by the distance traveled. For PT operator, ridership, availability, punctuality, and cost and revenue are important indicators, but we limit our discussion to PT mode share in this paper. From the public authority's perspective, we can consider service availability and accessibility as indicators for equity concerns, and motorized traffic and active travel modes as indicators of sustainability. Based on these indicators, the design decisions such as fleet size, vehicle capacity, fare scheme and sharing/hailing policies can be simulated and evaluated.

4. Supply-demand interaction

4.1. Formulation and fixed-point solution

In this section, we model the interaction between demand and supply in the AMoD system. As shown in Fig. 1, the interaction mechanism consists of two loops representing the choices of travelers and operators respectively. This mechanism is designed for (a) formulating the fixed-point problem for demand equilibrium when level-of-service indicators are unknown beforehand; and (b) supporting the AV + PT system design by providing supply performance analysis.

Using the level of service as the return value, the iterative loop on the left side enables explicit feedback for travelers. To start, the level-of-service indicators are set to arbitrary initial values (in Section 6 we demonstrate that the simulation results after converging to the fixed point solution are not sensitive to these initial value).

Based on the historical mode choice observations, the simulated level of service and assumptions with regard to fare and preference to AV service, the mode choice model predicts the AV + PT demand volumes for each OD pair. The simulation platform then evaluates the system performance based on the predicted demand matrix, predefined supply settings, as well as other system assumptions including the dispatching strategies. It outputs both the level-of-service indicators and the supply performance indicators. The former is returned to the mode choice model as feedback. The results of the demand prediction are updated accordingly and the iterative loop is completed.

The formulation of the problem is presented as below:

$$\begin{cases} \mathbf{D} = \text{ModeChoice}(\mathbf{T}, \mathbf{L}, \mathbf{V}_d, \mathbf{S}_d) \\ \mathbf{L} = \text{Simulation}(\mathbf{D}, \mathbf{V}_s, \mathbf{S}_s) \end{cases} \tag{1}$$

MODECHOICE is the demand prediction subproblem and SIMULATION is the simulation subproblem. MODECHOICE takes the matix of total current trips $\mathbf{T}$, level-of-service indicators $\mathbf{L}$, demand decision variables $\mathbf{V}_d$ (e.g. fare) and other demand assumptions $\mathbf{S}_d$ (e.g.

alternative specific constant for AV) as input. The matrix of total current trips $\mathbf{T}$ that is used as an input in MODECHOICE subproblem is derived from the sample of trips on all modes recorded in the household travel survey and expanded to match the population. The MODECHOICE output $\mathbf{D}$ is a vector of predicted OD-specific demand for AV + PT service. Symmetrically, SIMULATION takes in predicted demand $\mathbf{D}$, supply decision variables $\mathbf{V}_s$ (e.g. vehicle capacity, fleet size, hailing policy) and system assumptions $\mathbf{S}_s$ and gives estimates of level of service $\mathbf{L}$ and supply performance $\mathbf{P}$.

To find the solution to the system of equations in (1) we apply an interactive fixed-point solution approach. The pseudo code to the proposed algorithm is shown in Algorithm 1. To update the solution from one iteration to the next, we use the method of successive averages (MSA) (Sheffi and Powell, 1982) as shown in line 9 of the algorithm. The procedure keeps updating $\mathbf{D}$, $\mathbf{L}$ and $\mathbf{P}$ iteratively until $\mathbf{D}$ converges to a fixed-point solution and the demand-supply reaches balance and SOLVEFIXEDPOINT returns the estimated performance indicators at convergence, for travelers and operators.

Algorithm 1. Fixed-point Solution

1: **procedure** SOLVEFIXEDPOINT($\mathbf{T}$, $\mathbf{V}_d$, $\mathbf{S}_d$, $\mathbf{V}_s$, $\mathbf{S}_s$)
2: let level-of-service indicators be initial values $\mathbf{L}^{(0)}$
3: $\mathbf{D}^{(0)} = \text{MODECHOICE}(\mathbf{T}, \mathbf{L}^{(0)}, \mathbf{V}_d, \mathbf{S}_d)$
4: let step counter $i = 0$
5: **do**
6: $i = i + 1$
7: $\mathbf{L}^{(i)}, \mathbf{P}^{(i)} = \text{SIMULATION}(\mathbf{D}^{(i-1)}, \mathbf{V}_s, \mathbf{S}_s)$
8: $\mathbf{D}^{(i)} = \text{MODECHOICE}(\mathbf{T}, \mathbf{L}^{(i)}, \mathbf{V}_d, \mathbf{S}_d)$
9: $\mathbf{D}^{(i)} = \frac{1}{i}\mathbf{D}^{(i)} + \frac{i-1}{i}\mathbf{D}^{(i-1)}$
10: **while** $\|\mathbf{D}^{(i)} - \mathbf{D}^{(i-1)}\| > \delta$
11: **return** $\mathbf{D}^{(i)}$, $\mathbf{L}^{(i)}$, $\mathbf{P}^{(i)}$

On the right side of Fig. 1, the system design decisions are represented by a dashed line as we do not model the loop explicitly. According to the discussion in Section 3, designing an AV + PT service involves multiple stakeholders and the system design decisions should be made to reflect the interest of each party. In practice, the comprehensive performance metrics should be defined and examined on a case-by-case basis. Due to limited space, this paper only evaluates the most important system design decisions from the perspective of key stakeholders summarized in Table 1.

Section 4.2 provides the details on the simulation subproblems. Sections 4.3 and 4.4 introduce the case study and describe the mode choice model. We collect all the parameters and variables used in the paper in Table 2 and classify them as "input", "output", "decision", "assumption" and "intermediate". The system design "decision" variables and the values of the "assumption" variables will be discussed in Section 5.

4.2. Simulation platform

The level-of-service indicators in $\mathbf{L}$ that have impact on the mode choice behavior are service rate SR_{av}, wait time WT_{av} and detour factor DF_{av}. $\mathbf{L}$ is largely dependent on system design and operational strategies and should be consequently studied together with the supply side. In order to understand the demand-supply interaction, we cast the subproblem $\mathbf{L} = \text{SIMULATION}(\mathbf{D}, \mathbf{V}_s, \mathbf{S}_s)$ into a continuous-time agent-based simulation platform. The platform is able to simulate door-to-door AV + PT service with a fixed-size fleet of dedicated autonomous vehicles (defined by vehicle capacity K_{av} and fleet size V_{av} in $\mathbf{V}_s$), requests (reflecting the predicted demand $\mathbf{D}$) and the necessary operational models and dispatching algorithms at operator's disposal (as assumptions in $\mathbf{S}_s$). The pseudo code for SIMULATION is shown in Algorithm 2.

Algorithm 2. Agent-based Simulation

procedure SIMULATION($\mathbf{D}$, $\mathbf{V}_s$, $\mathbf{S}_s$)
initialize the system according to $\mathbf{V}_s$ and $\mathbf{S}_s$
$t = 0$, $t_a = 0$, $t_r = 0$
while $t < T$ **do**
generate next request with arrival interval Δt and push into queue
$t = t + \Delta t$
if $t > t_a$ **then**
pop all pending requests from the queue
for each request **do**
if $\exists$ vehicle satisfying dispatching constraints **then**
assign the best vehicle to the request
else if t still in the wait time window **then**
push the request back to queue
else
request times out

```
        t_a = t_a + T_a
    if t > t_r then
        rebalance the idle vehicles
        t_r = t_r + T_r
    route the vehicles to t
return L, P based on service performance during T_s
```

There are three modules in the agent-based simulation platform: request generator, vehicle dispatcher and routing server.

The demand generator draws requests from the predicted OD matrix **D** (Eq. 1). During the period of simulation, the arrival of the requests follows a Poisson process of constant arrival rate, which is proportional to the OD-specific demand volume. Depending on hailing policy, a request can be either on-demand or in-advance and have specific constraints including maximum wait time (MWT_{av}) and maximum detour factor (MDF_{av}). Earliest possible departure time and latest possible arrival time also apply accordingly.

Table 2
Parameters and variables.

Vector	Parameter/Variable	Type
Level of service (**L**)	AV service rate (SR_{av})	Output
	AV wait time (WT_{av})	Output
	AV detour factor (DF_{av})	Output
Supply performance (**P**)	AV vehicle distance traveled (VMT_{av})	Output
	AV distance-based load (L_{av})	Output
Supply decision variables ($\mathbf{V}_s$)	AV vehicle capacity (K_{av})	Decision
	AV fleet size (V_{av})	Decision
Supply assumptions ($\mathbf{S}_s$)	AV maximum wait time (MWT_{av})	Assumption
	AV maximum detour factor (MDF_{av})	Assumption
	Period of simulation (T)	Assumption
	Period of study (T_s)	Assumption
	Period of warm-up (T_w)	Assumption
	Period of cool-down (T_c)	Assumption
	Interval of assignment (T_a)	Assumption
	Interval of rebalancing (T_r)	Assumption
Demand decision variables ($\mathbf{V}_d$)	AV base fare (c_{base})	Decision
	AV per-unit-time fare (c_{time})	Decision
	AV per-unit-distance fare (c_{dist})	Decision
	AV leg minimum price (MP_{av})	Decision
	Discount for sharing ($DC_{sharing}$)	Decision
	Discount for transfer ($DC_{transfer}$)	Decision
Predicted demand matrix (**D**)	(see Section 4.2)	Intermediate
Total current trips (**T**)	(see Section 4.3)	Input
Demand assumptions ($\mathbf{S}_d$)	Preference to AV (ASC_{av})	Assumption
	AV penalty wait time (PWT_{av})	Assumption
	(others, see Section 4.4)	Assumption
Supporting variables	coef for AV travel time ($\beta_{T,av}$)	Intermediate
	coef for PT travel time ($\beta_{T,pt}$)	Intermediate
	coef for AV + PT cost (β_C)	Intermediate
	coef for number of transfers (β_X)	Intermediate
	Cost of AV leg (C_{av})	Intermediate
	AV actual travel time (T_{av})	Intermediate
	AV actual travel distance (D_{av})	Intermediate
	AV shortest travel time (ST_{av})	Intermediate
	AV shortest travel distance (SD_{av})	Intermediate
	Cost of PT leg (C_{pt})	Intermediate
	Cost of AV + PT service (C_{av+pt})	Intermediate
	AV adjusted wait time (AWT_{av})	Intermediate
	Utility of AV + PT service (U_{av+pt})	Intermediate

Types “input” and “output” represent inputs and outputs of the model respectively. Type “intermediate” represents the intermediate variables in the model. Type “decision” represents the decision variables. Type “assumption” represents the assumptions made for the model.

SR_{av}, WT_{av} and DF_{av} require being initialized to start the iteration.

On-demand requests are then dynamically assigned to vehicles by the central dispatcher based on insertion heuristics (Jung et al., 2016). The objective of the heuristic method is to insert the request to the job list of a vehicle in the system that satisfies the constraints and minimizes the incremental cost in terms of the total estimated travel time for all travelers. This includes the remaining in-vehicle travel times for those on board, as well as the sum of the wait times and the in-vehicle travel times for those pending to be paired. If a request cannot be served within the wait time window due to vehicle availability and request constraints, the request is rejected and the traveler is assumed to "walk away". The service rate is then defined as the percentage of requests being served. The insertion heuristic also applies to in-advance requests. The only distinguishing characteristic is that in-advance requests are known to the dispatcher beforehand and travelers will be notified of their assignment 30 min before the earliest departure time.

The dispatcher also rebalances idle vehicles periodically to regain the balance between demand and supply and plan for the upcoming requests in the near future. To do so, we use an optimal rebalancing strategy, which is formulated as a mixed-integer nonlinear programming (MINLP) problem. The optimal rebalancing strategy maximizes the service availability during the following T_r seconds by taking into account the predicted demand volumes from each origin. For cost efficiency, the solution is also subject to constraints including maximum vehicle rebalancing distance traveled, which is proven to be an effective strategy in Wen et al. (2017). Rebalancing reduces the average wait time for on-demand travelers by more than 15% at the cost of increasing the vehicle distance traveled by less than 20%.

In order to balance the trade-off between optimality and computational efficiency, the request-vehicle assignment is performed every T_a simulated seconds and rebalancing is performed every T_r simulated seconds.

After dispatching, the routing server updates the shortest routes in real-time. Each vehicle moves accordingly to pickup/drop-off travelers or to rebalance the supply. The travel times between any origin and destination is accessed from OpenStreetMap.

The simulation runs for T seconds, of which T_s seconds makes the period of study. Requesting generated during T_s are used in the evaluation to calculate level of service **L** and supply performance **P**. **L** includes service rate, wait time and detour factor. **P** consists of vehicle distance traveled VMT_{av} and distance-based load L_{av}. The rest of the simulation time before and after T_s are warm-up T_w and cool-down T_c buffers. $T = T_w + T_s + T_c$.

4.3. Case study area

To present the mode choice model, a case study from a major European city is reported in this section. This city has an extensive and developed transportation network in which public transport has a high mode share (45% in 2015). Commuter rail travel has shown strong growth over the past decade, providing good service from the outskirt of the city to the downtown area.

The case study region is a spread-out residential area located about 25 km outside the downtown. It is centered around a commuter rail station with frequent and high-speed train service to downtown. However, bus service in this area is infrequent and not economically efficient as a result of the low residential density. Consequently, local trips are particularly car-dependent. The area is chosen as the case study area because (a) it possesses a significant first-mile demand to the train station, (b) the inefficient bus service requires improvement, and (c) it has an appropriate density for initial AV trials. We choose the case study area (CSA) of 15 km × 10 km and about 159 thousand residents, which is larger than the local administrative boundary in order to cover the catchment area of all important bus routes originating from the rail station.

We use the official annual household travel diary surveys from 2005 to 2014 to realize the current travel demand, including 2709 respondents (1.7% of the population) and 1639 AM trips in the CSA during the morning peak (6:30 am to 9:30 am).

Seven modes are observed in the travel diaries from the CSA and accounted in the mode choice mode: walk, bike, car, taxi, bus, rail and park/kiss + ride (P/K + R). The multimodal trips are classified based on their distance-based main mode, except for rail and P/K + R. Rail is defined as all bus + rail trips that involves a rail leg and P/K + R is listed separately to target first-mile travelers with car access. Table 3 outlines the current mode share. "Trips to Downtown" represents trips that have origins in CSA and destinations in the city downtown area (account for 11% of all trips). "Intrazonal Trip" represents trips that have both origins and destination in the CSA (account for 54% of all trips). Note that "All Trips" consists of trips other than"Intrazonal Trips" and "Trips to Downtown".

It is worth noticing in Table 3 that the CSA is relatively car-dependent with 58% of all trips and 57% of intrazonal trips using car. Trips to downtown are dominated by the rail service because of the fast rail connection to CBD. 74% of trips have rail as the main mode, and 16% rely on car as the first-mile access to rail. The mode share of bus is very small due to the inefficient service. Walk is popular for short intrazonal trips. Bike and taxi trips are very rare for all trips.

We denote the total trip dataset by **T** as an input to the mode choice subproblem MODECHOICE.

Table 3
Current mode share in CSA.

Mode	All Trips	Trips to Downtown	Intrazonal Trips
Walk	20%	0%	32%
Bike	1%	0%	1%
Car	58%	10%	57%
Taxi	0%	0%	0%
Bus	8%	0%	10%
Rail	11%	74%	0%
P/K + R	2%	16%	0%

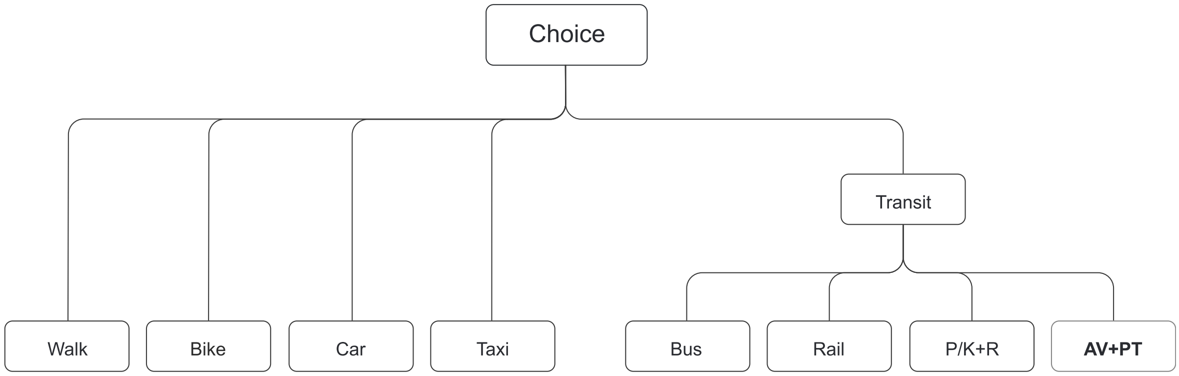


Fig. 2. The nested logit structure for the model choice model.

4.4. Mode choice model

To estimate the demand for the AV + PT service, we use a classic discrete choice model and predict the mode shares from a (constant) total travel volume. We assume no induced or latent demand in this research, and assume that individual travelers choose their travel modes based on the attributes of the existing modes and the simulated AV mode.

The mode choice subproblem $\mathbf{D} = \text{MODECHOICE}(\mathbf{T}, \mathbf{L}, \mathbf{V}_d, \mathbf{S}_d)$ is built using a nested logit model. MODECHOICE is driven by the total travel demand data $\mathbf{T}$ and responsive to the level of service $\mathbf{L}$ of AV + PT. The subproblem is for the purpose of understanding the sensitivity of demand to modal attributes in CSA and allowing the demand estimation to be responsive to supply factors.

After testing various specifications with different variables both with and without nesting structures, we chose to apply a nested logit (NL) structure for MODECHOICE shown in Fig. 2. This structure was selected based on the model's performance when considering goodness of fit, likelihood ratio test, and significance of coefficients and nest parameters.

We first estimate the NL model for the status quo mode split, and then apply the NL model to calculate the probability of shifting from the existing modes to the newly constructed AV + PT mode. The status quo mode choice model is reported in Appendix A. In this section, we present the assumptions, decisions variables, and level-of-service variables of the new AV + PT mode.

As for the assumptions $\mathbf{S}_d$, when AV + PT mode becomes available in CSA, it would serve both first-mile trips to the rail station and intrazonal trips that replace car and bus. As such, one typical AV + PT trip may have both AV and PT legs and we consider the synthetic AV + PT mode in the transit nest of the status quo nested logit model. The coefficient for AV travel time ($\beta_{T,av}$) is assumed to be the same as bus and the coefficient for PT travel time ($\beta_{T,pt}$) is assumed to be the same as rail. The coefficients for cost (β_C) and number of transfers (β_X) also apply to AV + PT.

The alternative Specific Constant (ASC) represents the intrinsic preference for the proposed AV + PT service after all independent variables such as travel time and travel cost are controlled. Due to the lack of existing services and the uncertainty in system design, the knowledge of the intrinsic preference for the AV + PT service is limited. Since the AV + PT mode inherently takes characteristics of both car and transit modes, we test a range of ASC in the simulation in this paper: −3.58 (lower bound benchmarked by P/K + R), −2.35 (upper bound benchmarked by car) and −3.00 (the midpoint case, average of bus and rail) and will discuss its impact in Section 5.3.

One decision variable $\mathbf{V}_d$ in the AV + PT utility function is fare. We adopt a three-part fare structure: base fare, per-unit-distance, and per-unit-time:

$$C_{av} = c_{base} + c_{time} T_{av} + c_{dist} D_{av} \tag{2}$$

C_{av} is the cost of the AV leg. T_{av} and D_{av} are travel time and distance. c_{base}, c_{time} and c_{dist} are per-trip-base, per-unit-distance, and per-unit-time parameters respectively. In our simulation we benchmark the proposed AV + PT product with current market prices for similar services as follows: Zipcar currently rents cars to Uber drivers at the fare of \$3.658/hour or \$0.061/min and \$0.289/km (Zipcar, 2017). These values can be used as proxies for the capital cost and the operating (fuel, insurance, depreciation, etc.) cost of the AV's. However, if the AV service is operated by a third party contractor, the operational cost of the vehicle dispatch platform has to be taken into account too. For this purpose we use the Uber revenue model. Uber has a fare of \$3.325/trip, \$0.199/min and \$1.033/km while taking 25% revenue (Cheer, 2016; The Telegraph, 2015). This makes Uber's revenue portion \$0.831/trip, \$0.050/min and \$0.258/km. Using the sum of these two prices (Zipcar fare and 25% of Uber fare), we can estimate the cost of providing the AV service based on travel time and distance. In this case, we have:

$$c_{base} = \text{"\$"}0.83, \quad c_{time} = \text{"\$"}0.11/\text{min}, \quad c_{dist} = \text{"\$"}0.55/\text{km}. \tag{3}$$

As assumed in the service design section, trips will be paired whenever they make sense and travelers cannot reject sharing. But we calculate the fare based on the shortest-path travel time and travel distance instead of the actual ones, and we offer a sharing discount DC_{share} to compensate travelers for the detour. In this case T_{av} and D_{av} are replaced by ST_{av} and SD_{av}, the shortest-path travel time and distance:

$$C_{av} = (c_{base} + c_{time}ST_{av} + c_{dist}SD_{av})(1 - DC_{share}) \quad (4)$$

A minimum trip price MP_{av} is also added to discourage very short trips. As well, discount $DC_{transfer}$ is given if transfers are involved. We assume no surge pricing. The total fare for an AV + PT trip is therefore:

$$C_{av+pt} = \max(C_{av}, MP_{av}) + \max((C_{pt} - DC_{transfer}), 0) \quad (5)$$

C_{av+pt} is the total travel cost of the trip. C_{pt} is the cost of the PT leg.

For the level of service of the AV leg, the indicator vector **L** is determined using the agent-based simulation platform discussed in Section 4.2. Based on the indicators from SIMULATION, average wait time and detour factor are used to update the travel time of an AV leg. Since the service rate can not always be 100%, we introduce a penalty wait time for those who walk away and propose the adjusted wait time:

$$AWT_{av} = WT_{av} \times SR_{av} + PWT_{av} \times (1 - SR_{av}) \quad (6)$$

and

$$T_{av} = AWT_{av} + ST_{av} \times DF_{av} \quad (7)$$

AWT_{av} denotes the adjusted wait time, calculated as a weighted average based on the average wait time WT for those served, the service rate SR_{av}, and the penalty wait time PWT_{av} for those rejected. DF_{av} is the detour factor.

The AV + PT utility of any specific OD pair is:

$$U_{av+pt} = ASC_{av+pt} + \beta_{T,av}T_{av} + \beta_{T,pt}T_{pt} + \beta_C C_{av+pt} + \beta_X X_{av+pt} + \epsilon \quad (8)$$

T_{pt} is the travel time for PT leg. C_{av+pt} and X_{av+pt} are cost and number of transfers respectively. Based on U_{av+pt} for each of the OD pairs, the mode choice model is used to predict the demand for the AV + PT mode and the current modes, denoted as **D**.

To start the interaction and solve the fixed-point problem, we initiate service rate, wait time and detour factor in **L** with initial values. MODECHOICE takes **L** as input and outputs **D**. SIMULATION then takes **D** to calculate **L** and **P**. The demand-supply interaction continues iteratively until reaching the equilibrium, i.e., the MSA condition is satisfied.

5. Results and analysis

5.1. System setting and simulation scenarios

Table 4 lists the values set for the assumption variables in the simulation, which remain constant throughout the paper.

For any scenario, the system setting variables for the simulation include the fleet size, vehicle capacity, hailing policy, fare policy and the ASC for the AV + PT service. The possible combinations of all levels of these variable are vast. It is beyond the scope of this paper to systematically explore all possible scenarios. For illustration purposes, we select only a subset of scenarios shown in Table 5 for our simulation experiments.

Service rate, wait time and detour factor are initiated to be 100%, 300 s and 1.00 for the demand-supply interaction. We define that the equilibrium is reached when the relative change of **D** is less than 0.5% of the total volume. The sharing discount is set as 25%. The minimum price of each ride is \$1.73, same as the bus base fare in the demand model. The transfer discount is set as \$1.33 for the rail leg.

5.2. Impact of service design

5.2.1. Fleet sizing

Starting with the initial values of service rate, wait time and detour factor, the demand-supply interaction usually reaches equilibrium in less than ten iterations. The computational times for each iteration directly depends on the size of network, demand levels and fleet size. On average, one hour of service simulation takes about one hour of computation using a regular desktop computer.

Table 4
Values for assumption variables.

Assumption Variable	Value
AV maximum wait time (MWT_{av})	10 min
AV penalty wait time (PWT_{av})	20 min
AV maximum detour factor (MDF_{av})	1.5
Period of simulation (T)	7200 s
Period of study (T_s)	3600 s
Period of warm-up (T_w)	1800 s
Period of cool-down (T_c)	1800 s
Interval of assignment (T_a)	30 s
Interval of rebalancing (T_r)	150 s

Table 5
System Settings and simulation scenarios.

Simulation	Parameters/Variable					Scenarios for Target of Interest
	Fleet Size	Vehicle Capacity	Fare Policy	Hailing Policy	ASC_{av+pt}	
Fleet Sizing (5.2.1)	Target	4	As in 5.1	As in 5.2.4	−3.00	200, 220, 240, 260, 280
Vehicle Capacity (5.2.2)	Dependent	Target	As in 5.1	As in 5.2.4	−3.00	Capacity = 1: fleet = 520, 560, 600, 640, 680 Capacity = 2: fleet = 280, 300, 320, 340, 360 Capacity = 3: fleet = 220, 240, 260, 280, 300 Capacity = 4: fleet = 200, 220, 240, 260, 280
Fare Policy (5.2.3)	220	4	Target	As in 5.2.4	−3.00	Original fare as in 5.1 vs. no minimum price/no transfer discount
Hailing Policy (5.2.4)	220	4	As in 5.1	Target	−3.00	On-demand requests vs. in-advance requests
Preference to AV (5.3)	Dependent	4	As in 5.1	As in 5.2.4	Target	ASC = −3.58: fleet = 100, 120, 140, 160, 180 ASC = −3.00: fleet = 200, 220, 240, 260, 280 ASC = −2.35: fleet = 350, 400, 450, 500, 550

A parameter/variable being “target” indicates it’s isolated and tested with a range of scenarios.
A parameter/variable being “dependent” indicates the choice of its value is dependent on that of the target.

Fig. 3a shows that demand volume converges around the 8th iteration. A larger fleet leads to higher service rates and shorter wait times shown in Fig. 3b, and consequently higher demand. In contrast, the fleet size has little impact on detour factor, which stabilizes at a low level of 1.15, shown in Fig. 3d. Average travel time is also steady (around 520s) regardless of the changing fleet size.

From the operator’s point of view, having more vehicles will inevitably increase the operational costs and decrease the fleet efficiency. As Fig. 3c shows, the total vehicle distance traveled outgrows the total demand volume, indicating higher variable cost of serving one trip as fleet becomes larger. This outgrowth of vehicle distance traveled results from the use of online rebalancing: the off-service AVs do not just idle; they are anticipatorily rebalanced to ideal locations in order to provide better service for future travelers. In fact, the average vehicle rebalancing distance (the upper bar in Fig. 3c) increases from less than 1 km/hour to more than 6 km/hour, the percentage increasing from 3% to 20%. In addition, the larger fleet size also results in the decline of the average load from 1.30 to 1.16 (the line with round marker in Fig. 3d).

The nature of the fleet sizing problem is therefore the trade-off between the benefits to the travelers and the cost to the operators. The wait time and service rate start to flatten after the fleet size reaches around 230, which we choose as a good point of balance. At this fleet size, we have about 1950 trips per hour and the re-balancing distance accounts for less than 10% of the total distance traveled. Again the choice of 230 is arbitrary and based on an intuitive trade-off. Future research can examine the optimization of the fleet size based on a formal cost benefit analysis.

5.2.2. Vehicle capacity and sharing

We only consider AVs with the maximum capacity of four travelers. Allowing trips to be shared affects travel experience: each traveler has longer travel time, higher uncertainty in travel time, and less space and less privacy in the vehicle. However, sharing has the advantage of bringing down the trip cost, and from the system-level perspective, it also helps reduce the number of vehicles on the road. To illustrate the power of sharing, we examine the scenarios where the AV operator limits the capacity to 1 (non-sharing), sharing of 2 or 3 or 4.

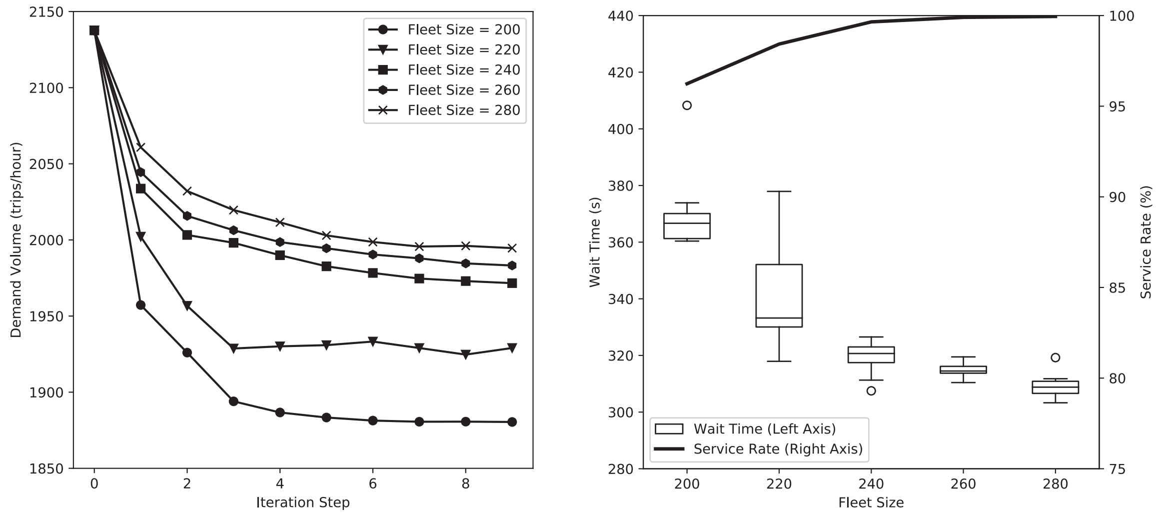
Our demand model takes into consideration the price discount, waiting time change, and travel time change due to the detour; but not the uncertainty in travel time or the privacy and social interaction preference. The solid curve in Fig. 4a shows that to guarantee the 99% service rate, the operator need around 560 vehicles if sharing is not available. If vehicles can be shared by at most 2, 3 and 4 travelers, the required fleet sizes are 310, 260 and 230 respectively. Sharing has a significant impact on the system performance: the number of vehicles on road can be reduced by more than half when a maximum of 4 passengers can be shared.

Consider the hypothetical extreme scenario (shared trips have the same origin, destination and departure time), where we only need 1/2, 1/3, and 1/4 of the fleet size respectively when two, three, four passengers can be shared (dotted curve in Fig. 4a). The gaps between the actual fleet sizes needed and those in the hypothetical scenarios enlarge with the number of shared passengers in the vehicle, and therefore, the benefit of moving to larger vehicles diminishes.

The power of sharing is also evidenced by the increase in the average vehicle load factor from 0.40 with no sharing (more than half of the time vehicles are traveling empty) to 1.18 when sharing capacity increases to 4, resulting in much higher efficiency and lower operational cost. Sharing also increases the detour (but to a much lesser degree than the load factor), with the detour factor ranging between 1.10 and 1.15, i.e., a 10% and 15% increase in passenger in-vehicle travel times.

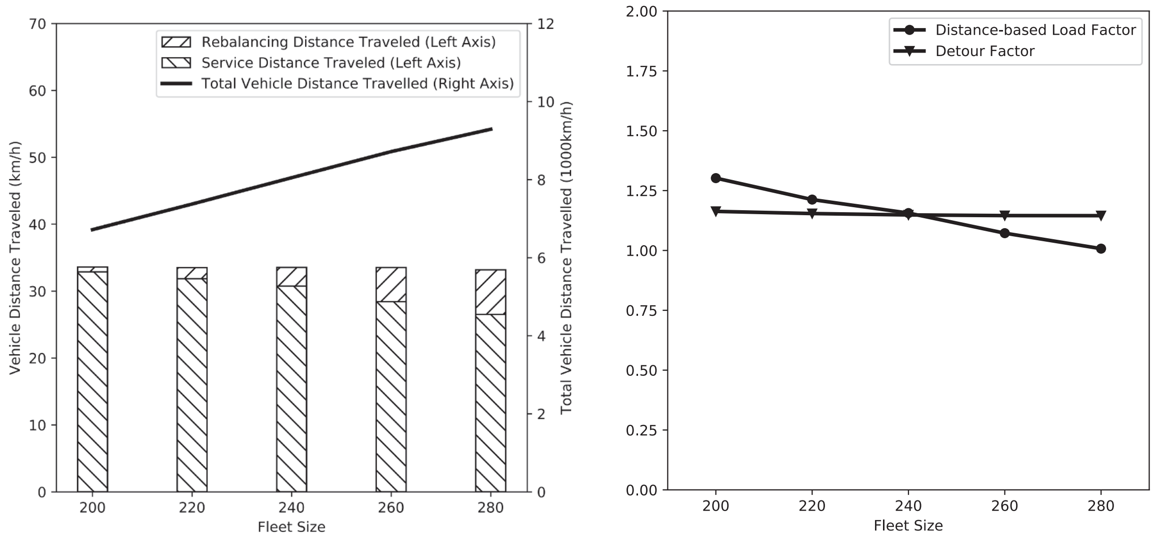
5.2.3. Fare policy

We compare the initial fare policy: i.e., minimum trip price ($1.73), sharing discount (25%), and transfer discount ($1.33) with an alternative policy: no discount for transfers ($DC_{transfer}$ = 0.00). Table 6 shows the comparison of the modal shift to the AV + PT mode in the three fare policy scenarios. The modal shift is defined as the percentage of trips leaving the existing mode for the AV + PT mode over the total trips by the existing mode before the introduction of the AV + PT mode. When the discount for transfers is removed, there is a significant decrease in modal shift in the downtown trips, since they often involve first-mile access and transfers.



(a) Demand volume converges to different levels when fleet size varies.

(b) A larger fleet leads to higher service rates and shorter wait times.



(c) Providing more vehicles results higher total vehicle service distance traveled.

(d) Load decreases when fleet size increases; detour factor remains steady.

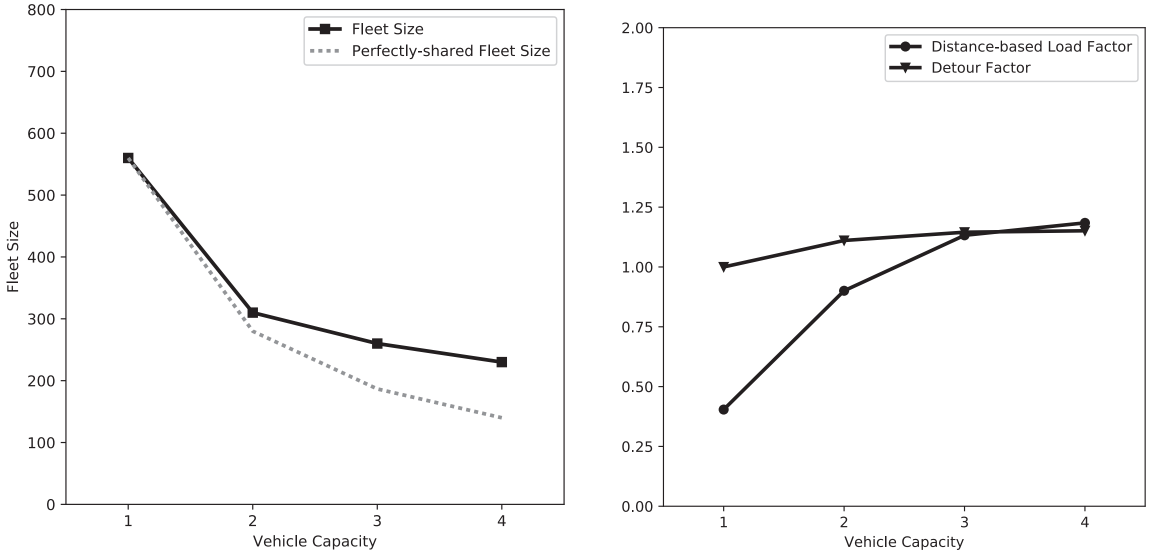
Fig. 3. Impact of fleet sizes on travelers and the operator (case ASC = -3.00).

5.2.4. Hailing policy

Fig. 5a shows the origins of the AV + PT trips scattered in the CSA. The service rate varies across the geographical location and some travelers may walk away due to long wait times. Fig. 5b shows the origin-specific probability of one walking away: The darker the marker is, the less likely a traveler from there will get a vehicle.

We notice that the satellite town in the southwest corner of the CSA is particularly under-served because it's distant from the main concentration of demand. Only 34.7% of trips from there are served there. In contrast, the service rate of the rest of CSA is 99.2%. To respond to the low availability of service, we propose that travelers from the satellite town are able to request in advance and assume that half of them choose to do so. We rerun the simulation and find that the service rate in the satellite town is substantially improved to 85.5% as shown in Fig. 5, with the service rate in the rest of CSA experiencing a relatively minor decrease to 98.5%. This could be a result of planning for the in-advance requests and moving vehicles from main area to southwest.

There are many options that the operator can explore: allowing different proportions of in-advanced booking, offering different levels of priority to the in-advanced bookings, and charging differently for in-advance requests. In the rest of the paper, we use this simple in-advanced booking policy (50% booking rate in the satellite town and 0% from the rest of the CSA) and maintain the same fare structure as before, i.e., no price discrimination.



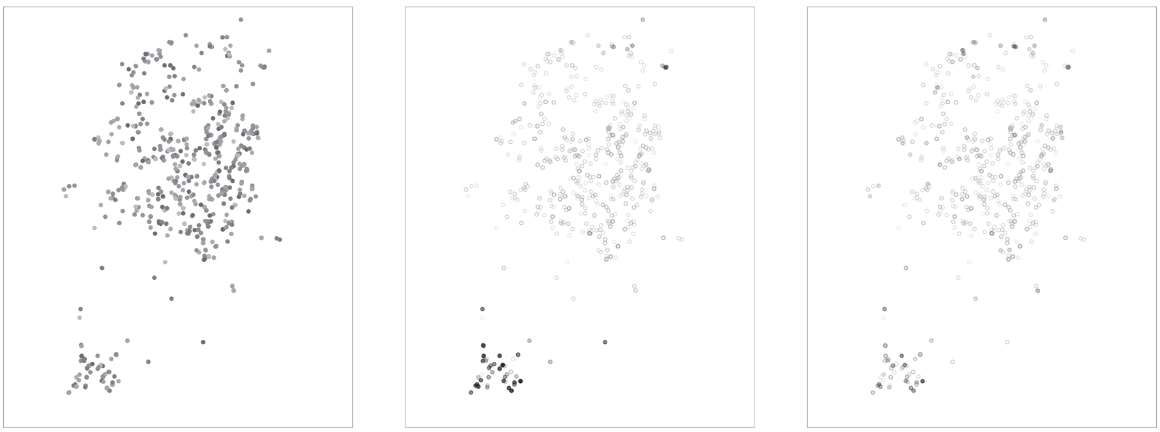
(a) The vehicle capacity-fleet size curve is decreasing and above the ideal curve.

(b) Both load factor and detour factor increases as vehicle capacity becomes larger.

Fig. 4. Impact of vehicle capacities on travelers and the operator (case ASC = -3.00).

Table 6
Modal Shift to AV + PT service under different fare policies.

Trip Type	Mode	Original Fare	$DC_{transfer} = 0$
Downtown	Car	43%	36%
	Rail	39%	31%
	P/K + R	43%	36%
Intrazonal	Car	10%	10%
	Bus	14%	14%



(a) Scatter of the origins of the predicted trips. The satellite town is the agglomeration in the southwest.

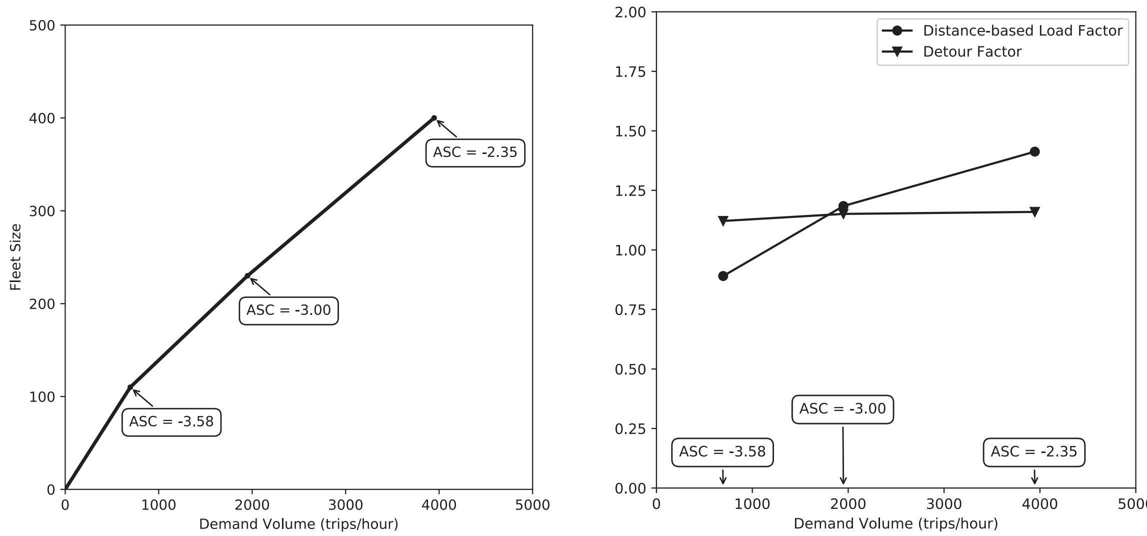
(b) Likelihood of walking away assuming on-demand requests only. Dark marker indicates high chance.

(c) Likelihood of walking away if in-advance requests are allowed in satellite town.

Fig. 5. Impact of hailing policy on service availability.

5.3. Preference to AV

We test three values of ASCs to reflect the uncertainty in travelers' intrinsic preference for the AV + PT service: -3.58 (lower bound), -3.00 (midpoint) and -2.35 (upper bound). Fig. 6a shows that as ASC becomes less negative, the utility of traveling by



(a) The relation between demand and fleet size indicates the economies of scale.

(b) A higher volume lead to higher load. Detour factor remains steady.

Fig. 6. Impact of preference (ASC) on system scale and performance.

AV + PT increases and the demand grows significantly (from around 700 to 1950 to 3950). To maintain the same level of service (99% service rate guarantee), the operator needs to deploy a larger fleet from 110 to 230 to 400 respectively. The average load factor also increases as in Fig. 6b. In contrast, the detour factor remains at the level of around 1.15. Given the high sensitivity to the ASCs, one important future research is the deeper understanding of the intrinsic preference for the AV technology.

5.4. Before-and-after mode shares

Table 7 shows the mode shares before and after the launch of AV + PT service. "Shift to AV + PT" shows the number and percentage of trips shifted from car, bus, rail and P/K + R to the AV + PT service. The mode share of the direct car trips to the downtown drops from 10% to 5%. The long-haul portion of the downtown trips is more dominated by rail but the access to the stations changes the mode mixture significantly. Walking trips to the stations decreases by 39%, and P/K + R trips reduces by 43%. For intra-zonal trips, the AV + PT service captures 11% of the mode share; car trips drops by 10% and bus trips decreases by 14%.

Fig. 7 shows the mixture of AV + PT trips switched from each of the existing modes. For intra-zonal trips, 82% of the AV + PT trips are from private cars and 18% from the buses. For downtown trips, 10% of the AV + PT trips are from private cars, 17% from P/K + R, and the majority (73%) from rail-they continue using rail but use AV as the access mode to the stations.

Table 7
AV + PT service ridership and mode share analysis.

Trip Type	Mode	Before		After		Shift to AV + PT	
		Trips	Share	Trips	Share	Trips	% Shift
Downtown	Car	775	10%	441	5%	334	43%
	Rail	5968	74%	3661	46%	2307	39%
	P/K + R	1277	16%	723	9%	554	43%
	AV + PT	0	0%	3195	40%	N/A	N/A
	Total	8020	100%	8020	100%	3195	40%
Intrazonal	Car	22715	86%	20373	77%	2342	10%
	Bus	3629	14%	3109	12%	520	14%
	AV + PT	0	0%	2862	11%	N/A	N/A
	Total	26344	100%	26344	100%	2862	11%

All of the downtown AV + PT trips use AV for rail connection. Only 10% of the intrazonal AV + PT trips involves rail and the rest of them are AV-only.

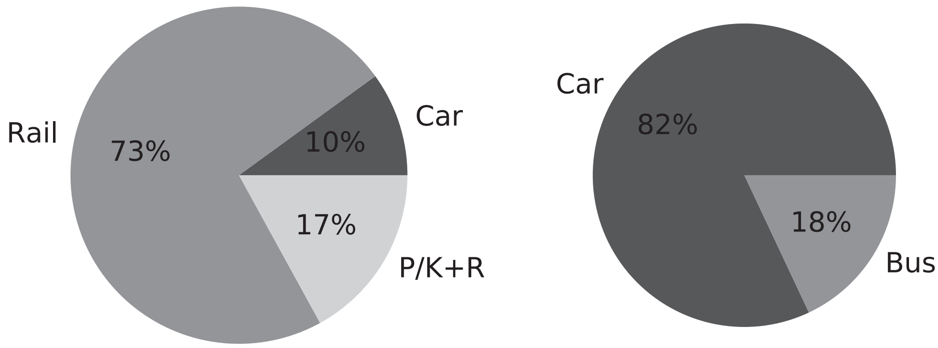


Fig. 7. Percentage of AV + PT trips from each existing mode: left: downtown trips; right: intrazonal trips.

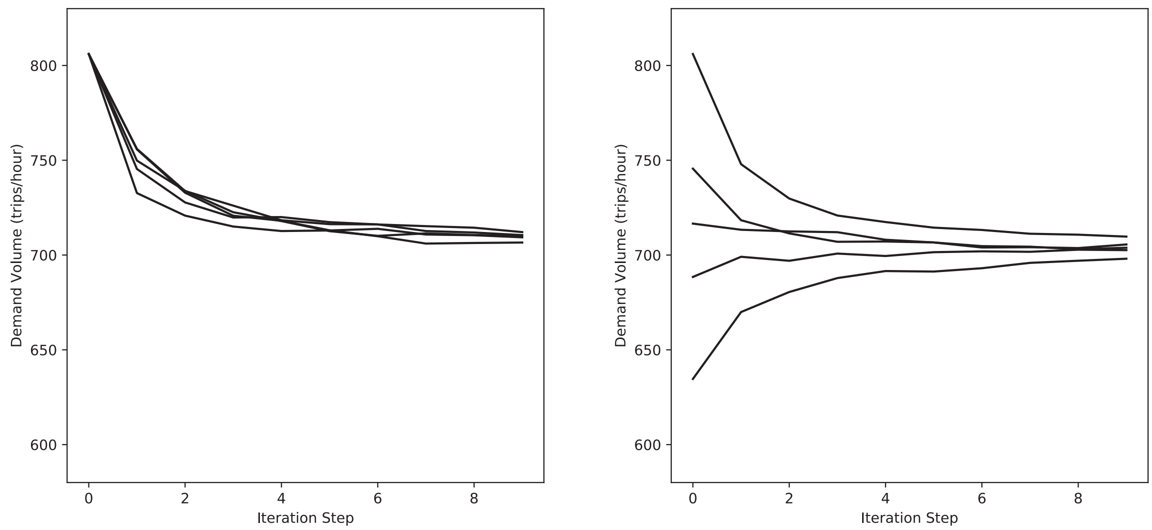
6. Robustness and convergence

The proposed agent-based simulation model is based on stochastic trip requests for the AV + PT service, generated according to pre-defined demand volumes for every origin-destination pair. The order of requests and their sending times change from one simulation run to another. The initial locations of the vehicles are also randomly generated at the beginning of the simulation time period. Therefore, two simulation runs may not have same results, even if all the decision variables (e.g. fleet size, fare scheme, ASC, etc.) and the demand volumes are identical. In this section, we report an experiment designed to evaluate the robustness and convergence of the simulation model.

To evaluate the impact of request stochasticity and test the robustness of simulation, we simulate a single scenario for five times and compared the results after convergence. In this scenario, the ASC for the AV + PT mode is set to −3.58, the fleet size is 120, and vehicle capacity is 4. Initial values for service rate, wait time and detour factor are 100%, 5 min and 1.0 respectively. The penalty wait time is 20 min. Fig. 8a shows the estimated demand for the five runs reported over 10 MSA iterations. All simulations start with the initially estimated total demand of 806, and they all converge to an estimate of about 709. This experiment shows that the simulation solution is robust and the stochasticity of requests does not affect the equilibrium results significantly.

In order to test the sensitivity to the initial level of service values, we simulate the scenario with five different sets of initial values, in which service rate ranges between 90% and 100%, wait time from 5 to 8 min, and detour factors from 1.0 to 1.3. Although these simulation results initiate at different demand levels, ranging between 634 and 809 (Fig. 8b), they all converge to an estimated demand level of about 705 after ten iterations.

These two experiments indicate that despite various stochastic features in the simulation process, the proposed method generates robust estimation of travel demand for AV + PT service.



(a) The repetitive runs of a single scenario show good robustness despite the stochasticity.

(b) The choices of initial values has limited impact on the convergence after iterations.

Fig. 8. Simulation results for robustness and convergence.

7. Conclusion and future research

This paper offers a systematic approach to the design, simulation and evaluation of AV + PT systems and demonstrates specific AV + PT service designs. The system design and modeling framework reflect the transit-oriented considerations that are important for high service availability, seamless connections, and equity. Using demand-supply interaction, we represent the choices of both travelers and operators. Results show the trade off between improving level of service and traveler experience, and the cost of larger fleet size and lower occupancy, and our model provides the starting point of identifying the optimal balance. Encouraging ride-sharing, allowing in-advance requests, and combining fare with transit are tools to enable service integration and sustainable travel.

We point the future research needs in four broad areas:

The first is to examine how the integrated AV + PT system will impact active modes. The convenience of the door-to-door service may deprive people of the opportunity to walk and cycle. Scheltes and de Almeida Correia (2017) show that AV service for last mile is competitive with walk, but less so with bike in Delft, Netherlands. As the next step, simulations, stated preference surveys, and field experiments could help empirically assess the impact of AVs on active modes in our case study city. Besides attracting trips from active modes, the AV service can also induce latent demand. Harb (2018) shows an over 80% vehicle mile traveled (VMT) increase in a naturalistic experiment with chauffeur drivers. This paper makes a strong assumption of fixed total travel demand, but it is critical for the future research to examine the latent demand.

The second is to investigate the AV fleet sizing and management strategies, as well as its impact on the repurposing of the bus services. In our simulated scenarios, the intra-zonal bus demand dropped by 14%. It is important to examine how bus service needs to be reoptimized so as to align itself with integrated AV + PT services. More broadly, the simulation model should be expanded to include bus service so that we can consider the AV fleet and bus fleet management jointly.

The third is to explore a range of AV + PT fare products and their impact on the system. The success of an AV + PT service depends on how much it costs and how attractive it is to the travelers. Based on the demand-supply interaction mechanism, we can study how different fare products such as monthly pass, early-bird discounts, and concession fares will change consumer behavior. Dynamic pricing (surges) may also be a tool for generating extra revenue and managing demand. This would help design and price the services based on the objectives of different stakeholders, though they may conflict each other.

The fourth is to make the simulation model congestion-aware. This paper uses static travel times rather than congestion-dependent ones and chooses a low-density suburb where it is reasonable to assume that traffic congestion is low and the AMoD vehicles do not change the underlying traffic conditions significantly. But for applications in congested networks, it is critical to capture the congestion effect and we should employ a proper traffic model in such applications. Such models need to be able to represent the current traffic flow and the travel time of each road segment, to reflect the flow-speed relationship including the AV driving behavior if they behave differently from human-driven vehicles, to model the stochastic features of the traffic and congestion, and to model the changes to the bus operation because of the traffic condition change and the ridership change.

Other future works include: implementing advanced algorithms to reoptimize the request-vehicle assignments; incorporating service reliability as a measure of performance, and developing a distribution for waiting times and in-vehicle times instead of using the average values; improving the accuracy of demand-supply interaction by repeating each iteration of the MSA multiple times and using the average; and better understanding the cost of operating and maintaining the AV fleet.

Acknowledgment

The authors thank our research sponsor for providing financial support and data. We also thank Nate Bailey, Jintai Li, Saeid Saidi, Yu Shen, Hongmou Zhang, and members from MIT Mobility Lab (mobility.mit.edu) and MIT Transit Lab for their comments. The authors would also like to thank two anonymous reviewers for their constructive comments.

Appendix A. Choice model of the status quo mode split

The existing choice set $\mathscr{C} = \{w, k, c, t, b, r, p\}$ contains seven modes: walk, bike, car, taxi, bus, rail and P/K + R respectively. U_i is the utility of mode i from $\mathscr{C}$. $\mathscr{C}_m$ is a nested mode set with utility U'_m, $m = 1, \ldots, M$ if M nested sets are available. Based on the utilities, the probability of choosing mode $i \in \mathscr{C}_m$ is as below:

$$\Pr(i|\mathscr{C}) = \frac{\exp(\mu_m U_i)}{\sum_{j\in\mathscr{C}_m} \exp(\mu_m U_j)} \frac{\exp(\mu U'_m)}{\sum_{k=1,\ldots,M} \exp(\mu U'_k)} \tag{A.1}$$

where

$$U'_m = \frac{1}{\mu_m} \ln \sum_{j\in\mathscr{C}_m} \exp(\mu_m U_j) \tag{A.2}$$

The utility equations are defined as follows:

$$U_w = ASC_w + \beta_{T,w} T_w \tag{A.3}$$

$$U_k = ASC_k + \beta_{T,k} T_k \tag{A.4}$$

$$U_c = ASC_c + \beta_{T,c}T_c + \beta_C C_w + \beta_N N_c \tag{A.5}$$

$$U_t = ASC_t + \beta_{T,t}T_t + \beta_C C_t + \beta_{D,t}D_t \tag{A.6}$$

$$U_b = ASC_b + \beta_{T,b}T_b + \beta_W W_b + \beta_C C_b + \beta_X X_b \tag{A.7}$$

$$U_r = ASC_r + \beta_{T,r}T_r + \beta_W W_r + \beta_C C_r + \beta_X X_r \tag{A.8}$$

$$U_p = ASC_p + \beta_{T,p}T_p + \beta_W W_p + \beta_C C_p + \beta_X X_p \tag{A.9}$$

ASC_i represents the alternative specific constant of mode i;T_i is the total travel time for all non-transit-related modes in 10-min increments; for bus, rail, and P/K + R, the walking time (for access and egress) is separated from the total travel time and represented by W_i, also measured in 10-min increments. C_i is the equivalent cost in U.S. Dollars ($); X_i is the number of transfers; D_i is the total travel distance in kilometers and N_i is the number of cars that the household owns. The mode-specific coefficient for travel time is denoted as $\beta_{T,i}$. The coefficient for walking time in transit modes is listed as β_W. Similarly, $\beta_{D,t}$ is the coefficient for taxi travel distance. β_C, β_X and β_N are the generic coefficients for cost, transfer and car ownership respectively. Also, $\mu_{transit}$ is denoted as the scale parameter of the transit nest (bus, rail, P/K + R and in the future AV + PT). Walking time and travel time for walk, bike, car, bus and rail are generated based on the fastest path during the morning peak using Google Maps API. Taxi and P/K + R times are estimated using car and transit respectively with an additional 3-min wait for taxi mode.

Other attributes including travel costs are determined under the following assumptions.

- The scale parameter μ in Eq. (A.1) is set to 1.
- Walk: As the reference mode, ASC is set to 0.
- Car: Parking cost reflects the morning, commuting, on-street parking cost. Congestion charging is $15.30 if a trip has origin or destination in the zone. Gas cost is $6.96/gallon with an average car fuel economy of 76.17 km/gallon.
- Taxi: Cost is estimated based on workday, day-time price.
- Bus/Rail: Bus trips are defined as "bus-only" and the cost is $1.73 per trip. Non bus-only trip alternatives are dominated by rail and as such classified as rail. Rail trips have distance-based fare based on a predefined fare scheme.
- P/K + R: Cost is assumed to be equivalent to transit travel cost.

The models are estimated using PythonBiogeme (Bierlaire, 2016). With 1639 observations, the model has an adjusted rho-squared of 0.613. Grouping transit alternatives in the nesting structure produces a robust model and matches our intuition that transit modes should be correlated. The results are shown in Table A.8.

The ASC of walk is 0, whereas for all other modes ASCs are negative. This indicates that in the absence of significant difference in other travel attributes (which is probably only true for short distance trips), walking is likely to be the first choice. The ASCs of all transit modes (bus −2.87, rail −3.14, P/K + R −3.58) and car (−2.35) are almost similar and less negative than others. The

Table A.8
Estimation results of mode choice model.

Coefficient	Value	Standard Error	t-test	p-value
ASC_k	−5.01	0.42	−11.81	0.00
ASC_c	−2.35	0.20	−11.87	0.00
ASC_t	−5.81	1.12	−5.18	0.00
ASC_b	−2.87	0.23	−12.30	0.00
ASC_r	−3.14	0.30	−10.47	0.00
ASC_p	−3.58	0.49	−7.27	0.00
$\beta_{T,w}$	−1.20	0.08	−15.21	0.00
$\beta_{T,k}$	−1.07	0.23	−4.66	0.00
$\beta_{T,c}$	−1.04	0.11	−9.55	0.00
$\beta_{T,t}$	−1.83	0.90	−2.03	0.04
$\beta_{T,b}$	−0.37	0.07	−5.56	0.00
$\beta_{T,r}$	−0.36	0.08	−4.31	0.00
$\beta_{T,p}$	−0.28	0.08	−3.43	0.00
β_W	−0.54	0.11	−5.06	0.00
β_C	−0.11	0.03	−4.16	0.00
β_X	−0.45	0.17	−2.64	0.01
$\beta_{D,t}$	0.44	0.11	4.14	0.00
β_N	0.75	0.08	9.03	0.00
$\mu_{transit}$	3.23	1.19	2.71	0.01

automobile network infrastructure and transit service provision have probably made these four modes dominant travel alternatives. Taxi and bike have very negative ASCs (−5.74 and −4.99) likely due to limited availability for taxi and necessary physical requirements for bike.

The coefficients for travel time are almost similar for car (−1.04), bike (−1.07) and walk (−1.20). This is probably because in either of these modes the traveler has to be actively engaged and cannot make use of that time to engage in other activities such as reading, writing, etc. In contrast, the coefficients for transit related travel time are about a third of the previous group. In these modes travelers are free to use that time for other activities. The transit parameter value for walking is slightly higher than travel time for other aspects of transit travel. This is likely due to walking being viewed as more onerous. Furthermore, the value of time of transit modes is about $20/hour. P/K + R has a slightly lower value of time at about $15/hour and walking time in transit has a slightly higher value of time at roughly $30/hour.

For the remaining three variables used in the mode choice model, it is shown that there is a negative correlation between number of transfers and utility. Roughly one transfer is equal to 12–16 min in transit, which is consistent with findings from the literature (Nassir et al., 2018). The positive value of car ownership correctly reflects the fact that people are more likely to choose car if it's available. Taxi distance is positive since the model already considers cost and travel time, showing that taxi with higher speed is preferred.

Appendix B. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.trc.2018.10.018.

References

Alonso-Mora, J., Samaranayake, S., Wallar, A., Frazzoli, E., Rus, D., 2017. On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment. Proc. Nat. Acad. Sci. 201611675.
Anderson, J.M., Nidhi, K., Stanley, K.D., Sorensen, P., Samaras, C., Oluwatola, O.A., 2014. Autonomous Vehicle Technology: A Guide for Policymakers. Rand Corporation.
Bierlaire, M., 2016. PythonBiogeme: a short introduction. Technical Report.
Cheer, L., 2016. Uber to hike its cut of fares to 25 per cent - and drivers warn the 'greedy' money-grab will lead to fewer cars on the road, in worse shape, and higher prices. Http://www.dailymail.co.uk/news/article-3538933/Uber-looks-hike-drivers-fee-25-cent-lead-price-surge.html.
Chen, T.D., Kockelman, K.M., 2016. Management of a shared autonomous electric vehicle fleet: Implications of pricing schemes. Transport. Res. Rec.: J. Transport. Res. Board 37–46.
Childress, S., Nichols, B., Charlton, B., Coe, S., 2015. Using an activity-based model to explore the potential impacts of automated vehicles. Transport. Res. Rec.: J. Transport. Res. Board 99–106.
Deng, Y., Cardin, M.A., 2018. Integrating operational decisions into the planning of one-way vehicle-sharing systems under uncertainty. Transport. Res. Part C: Emerg. Technol. 86, 407–424.
Fagnant, D.J., Kockelman, K.M., 2014. The travel and environmental implications of shared autonomous vehicles, using agent-based model scenarios. Transport. Res. Part C: Emerg. Technol. 40, 1–13.
Fagnant, D.J., Kockelman, K.M., 2015. Dynamic ride-sharing and optimal fleet sizing for a system of shared autonomous vehicles. In: Transportation Research Board 94th Annual Meeting.
Freemark, Yonah, Hudson, Anne, Zhao, Jinhua, 2018. Are cities prepared for autonomous vehicles? Planning for technological change by U.S. local governments. J. Am. Plann. Assoc (submitted for publication).
Gruel, W., Stanford, J.M., 2016. Assessing the long-term effects of autonomous vehicles: a speculative approach. Transport. Res. Procedia 13, 18–29.
Hall, J.D., Palsson, C., Price, J., et al., 2017. Is Uber a substitute or complement for public transit? Technical Report.
Harb, Xiao, C.M.W., 2018. Projecting travelers into a world of self-driving vehicles: Estimating travel behavior implications via a naturalistic experiment. Transportation Research Board 97th Annual Meeting.
Jorge, D., Correia, G.H., Barnhart, C., 2014. Comparing optimal relocation operations with simulated relocation policies in one-way carsharing systems. IEEE Trans. Intell. Transport. Syst. 15, 1667–1675.
Jorge, D., Molnar, G., de Almeida Correia, G.H., 2015. Trip pricing of one-way station-based carsharing networks with zone and time of day price variations. Transport. Res. Part B: Methodol. 81, 461–482.
Jung, J., Jayakrishnan, R., Park, J.Y., 2016. Dynamic shared-taxi dispatch algorithm with hybrid-simulated annealing. Comput.-Aided Civil Infrastruct. Eng. 31, 275–291.
Kloostra, B., Roorda, M.J., 2017. Fully autonomous vehicles: analyzing transportation network performance and operating scenarios in the greater Toronto Area, Canada. Technical Report.
Krueger, R., Rashidi, T.H., Rose, J.M., 2016. Preferences for shared autonomous vehicles. Transport. Res. Part C: Emerg. Technol. 69, 343–355.
Lenz, B., Fraedrich, E., 2016. New mobility concepts and autonomous driving: the potential for change. In: Autonomous Driving. Springer, pp. 173–191.
Levin, M.W., 2017. Congestion-aware system optimal route choice for shared autonomous vehicles. Transport. Res. Part C: Emerg. Technol. 82, 229–247.
Levin, M.W., Boyles, S.D., 2015. Effects of autonomous vehicle ownership on trip, mode, and route choice. Transport. Res. Rec.: J. Transport. Res. Board 29–38.
Levin, M.W., Kockelman, K.M., Boyles, S.D., Li, T., 2017. A general framework for modeling shared autonomous vehicles with dynamic network-loading and dynamic ride-sharing application. Comput. Environ. Urban Syst. 64, 373–383.
Liang, X., de Almeida Correia, G.H., van Arem, B., 2018. Applying a model for trip assignment and dynamic routing of automated taxis with congestion: system performance in the city of delft, the Netherlands. Transport. Res. Rec.: J. Transport. Res. Board.
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. Transport. Res. Part E: Logist. Transport. Rev. 93, 115–129.
Ma, J., Li, X., Zhou, F., Hao, W., 2017. Designing optimal autonomous vehicle sharing and reservation systems: A linear programming approach. Transport. Res. Part C: Emerg. Technol. 84, 124–141.
Marczuk, K.A., Hong, H.S.S., Azevedo, C.M.L., Adnan, M., Pendleton, S.D., Frazzoli, E., et al., 2015. Autonomous mobility on demand in simmobility: Case study of the central business district in singapore. In: Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM), 2015 IEEE 7th International Conference on. IEEE, pp. 167–172.
Marczuk, K.A., Soh, H.S., Azevedo, C.M., Lee, D.H., Frazzoli, E., 2016. Simulation framework for rebalancing of autonomous mobility on demand systems. In: MATEC Web of Conferences. EDP Sciences, pp. 01005.
Martinez, L.M., Correia, G.H., Viegas, J.M., 2015. An agent-based simulation model to assess the impacts of introducing a shared-taxi system: an application to lisbon

(portugal). J. Adv. Transport. 49, 475–495.
Martínez, L.M., Correia, G.H.d.A., Moura, F., Mendes Lopes, M., 2017. Insights into carsharing demand dynamics: outputs of an agent-based model application to lisbon, portugal. Int. J. Sustain. Transport. 11, 148–159.
Martinez, L.M., Viegas, J.M., 2017. Assessing the impacts of deploying a shared self-driving urban mobility system: An agent-based model applied to the city of lisbon, portugal. Int. J. Transport. Sci. nd Technol. 6, 13–27.
Nassir, N., Hickman, M., Ma, Z.L., 2018. A strategy-based recursive path choice model for public transit smart card data. Transport. Res. Part B: Methodol.
Panagiotopoulos, I., Dimitrakopoulos, G., 2018. An empirical investigation on consumers intentions towards autonomous driving. Transport. Res. Part C: Emerg. Technol. 95, 773–784.
Qiu, Han, Li, Ruimin, Zhao, Jinhua, 2018. Dynamic pricing in shared mobility on demand service. IEEE Trans. ITS arXiv:1802.03559.
Rayle, L., Dai, D., Chan, N., Cervero, R., Shaheen, S., 2016. Just a better taxi? a survey-based comparison of taxis, transit, and ridesourcing services in san francisco. Transp. Policy 45, 168–178.
Rossi, F., Zhang, R., Hindy, Y., Pavone, M., 2017. Routing autonomous vehicles in congested transportation networks: structural properties and coordination algorithms. Autonom. Rob. 1–16.
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. Int. J. Transport. Sci. Technol. 6, 28–41.
Sheffi, Y., Powell, W.B., 1982. An algorithm for the equilibrium assignment problem with random link times. Networks 12, 191–207.
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. Transport. Res. Part A: Policy Practice 113, 125–136.
Smith, B.W., 2012. Managing autonomous transportation demand. Santa Clara L. Rev. 52, 1401.
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.
Stuntz, A., Attanucci, J., Salvucci, F.P., 2017. A process for transit fare structure and fare level analysis: case study at the Massachusetts bay transportation authority. Technical Report.
The Telegraph, 2015. Proof that uber costs less than black cabs. Http://www.telegraph.co.uk/technology/uber/11902613/Proof-that-Uber-costs-less-than-black-cabs.html.
Vakayil, A., Gruel, W., Samaranayake, S., 2017. Integrating shared-vehicle mobility-on-demand systems with public transit. Technical Report.
Vasconcelos, A.S., Martinez, L.M., Correia, G.H., Guimarães, D.C., Farias, T.L., 2017. Environmental and financial impacts of adopting alternative vehicle technologies and relocation strategies in station-based one-way carsharing: An application in the city of Lisbon, Portugal. Transport. Res. Part D: Transport Environ. 57, 350–362.
Wen, J., Zhao, J., Jaillet, P., 2017. Rebalancing shared mobility-on-demand systems: A reinforcement learning approach. In: 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC). IEEE, pp. 220–225.
Xu, M., Meng, Q., Liu, Z., 2018. Electric vehicle fleet size and trip pricing for one-way carsharing services considering vehicle relocation and personnel assignment. Transport. Res. Part B: Methodol. 111, 60–82.
Yan, X., Levine, J., Zhao, X., 2018. Integrating ridesourcing services with public transit: An evaluation of traveler responses combining revealed and stated preference data. Transport. Res. Part C: Emerg. Technol.
Yap, M.D., Correia, G., Van Arem, B., 2016. Preferences of travellers for using automated vehicles as last mile public transport of multimodal train trips. Transport. Res. Part A: Policy Practice 94, 1–16.
Zhang, R., Pavone, M., 2016. Control of robotic mobility-on-demand systems: a queueing-theoretical perspective. Int. J. Robot. Res. 35, 186–203.
Zhu, S., Kornhauser, A.L., 2017. The interplay between fleet size, level-of-service and empty vehicle repositioning strategies in large-scale, shared-ride autonomous taxi mobility-on-demand scenarios. Technical Report.
Zipcar, 2017. Zipcar. Https://drive.uber.com/ukmarketplace/partner/zipcar/.