https://doi.org/10.1016/j.trc.2021.103089

Differentiating and Modeling the Installation and the Usage of Autonomous Vehicle Technologies: A System Dynamics Approach for Policy Impact Studies

Jiangbo Gabriel Yu [a]; Anthony Chen [b*]
[a] AECOM
[b] Department of Civil and Environmental Engineering
Hong Kong Polytechnic University
[*] Corresponding Author: anthony.chen@polyu.edu.hk

Abstract

Existing research that forecasts market penetration of installed connected and autonomous vehicle (AV) technologies is often confused with the traffic composition in roadway networks. Users may override AV mode due to arrival time pressure, facility constraint (e.g., "I will have to make a U-turn a mile away if I do not cross the solid double-yellow lines here"), drug and alcohol influence, pleasure, envy (e.g., "why the front car can surpass that slow truck but I can't?"), insufficient law enforcement, driving culture, media and public sentiment, etc. Therefore, the installation and the usage of AV technologies should not be instantaneously assumed ignorable in planning and policy studies. This paper is dedicated to clarifying this confusion by demonstrating that ignoring the difference between the installation and the usage of AV technologies might lead to systematic bias in evaluating policy and investment decisions. Through a system dynamics (SD) model, the complex interactions of relevant factors are captured so that the mixed traffic condition influences traffic law enforcement adjustment effort and system investment decisions, which, in turn, influence the AV technology usage share and the system performance. The case study applies to the greater Washington, D.C. area for demonstrating the feasibility and advantages of the proposed model and for studying policy implications. This paper does not attempt to forecast; instead, we propose a modeling framework for studying the conditions under which differentiating the installation and the usage of AV technologies might be critical in forecasting the traffic composition trend and system performance for public policy and investment decisions.

Keywords: *Autonomous Vehicles Usage; FIFO Violation; Law Enforcement; Public Sentiment; Delay; System Dynamics*

1. Introduction

Many (e.g., Litman, 2017; Bansal and Kockelman, 2017) argue, and we agree, that mixed traffic of vehicles with and without connected and autonomous vehicle (CAV or AV for short) technologies will last for at least decades even in extreme market growth conditions. Existing research has been focused on forecasting market penetration (MP) of the installation of AV technologies (e.g., Bansal and Kockelman, 2017; Nieuwenhuijsen et al., 2018). However, this MP is unlikely to equal the traffic composition of the active fleet (i.e., the *usage* share of AV technologies).

Since most potential AV stockholders adopt a progressive strategy (Deloitte, 2020), AV users, for a relatively long time, may still have the freedom to drive in manual mode temporarily or continuously in various circumstances for comparative advantages over the AV mode. In mixed traffic, vehicles manually driven by humans might be more favorable in insufficient AV infrastructure, inadequate or biased law enforcement, ill-natured collective driving culture, pressure for punctual arrival, envies, etc. Some people might prefer ridesharing services using manually driven vehicles (HVs) because AVs would drive "too conservatively," would not surpass the front vehicle when "it supposes to," or would not across the double-yellow lines even though it looks "perfectly safe" to do so. Privately-owned freight operators might have similar incentives for such a temporary switchover to reduce delivery time and cost if human drivers are in the trucks or can override the AV mode remotely. Although lacking empirical evidence, one can still derive from general principles that human drivers tend to follow traffic rules less strictly than AVs in terms of whether and how to surprising a front vehicle, speed limit, solid double yellow lines, no U-turns, no right-turn-on-red, no use of emergency lanes, illegal parking, etc. Moreover, the peer AVs being perceived to be more predictable to yield may further contribute to manual driving, which might lead to a vicious cycle.

"First-In-First-Out (FIFO) violation." in traffic studies, are often used to describe scenarios where flows that enter a link (without intermediate exit) at a later time propagate faster than flows that entered a link earlier (e.g., Carey, 1992; Boyce et al., 2001; Jin et al., 2005; Jin, 2007). Specific to this paper are the two types of driving behaviors that tend to incentivize people to (at least temporarily) override the AV mode. The first type includes driving behaviors that violating traffic rules such as making a U-Turn at an intersection that prohibits such movement and using the lane of opposing traffic when a "double-yellow line" is present. These behaviors are likely considered unallowable by the AV algorithms. The second type is the unrecommended risky driving behaviors (that may not violate traffic rules). For example, a driver might lose patient and tries to surpass the slow heavy-duty truck in front of it even though the AV technology does not recommend it. For another example, it is considered a (HV) FIFO violation when an HV surpasses a leading vehicle (while an AV would not); this behavior may cause the destabilization of the fleet and motivate surrounding vehicles to switchover to manual mode for performing similar behaviors. The common feature of these two types of FIFO violations is the intention and the associated driving behaviors of reaching a destination faster than what it would have been than using AV technologies. This paper uses the term FIFO violation as an all-encompassing term for any manual driving behaviors that fail to strictly follow traffic rules and courtesies, negatively impact the reliability of the system (relative to that of the AV technologies), and induce envies and behavioral mimicry.

FIFO violations may have both short-term and long-term impact on the system performance, law enforcement effort, network planning and policy decisions, etc., which, in turn, may influence the utility of FIFO violations. In a society where a switchover from AV mode to manual mode becomes common, people would be less incentivized to pay extra for vehicles to equip with advanced AV technologies. In this paper, we focus our study of implications of FIFO violation on network asset management, capital investment, public sentiment, and traffic law enforcement, to gain better understanding of three main policy leverages: (1) the effort of traffic law enforcement, (2) the overall fleet size and AV technology equipment regulations, and (3) the overall network facility investment and those that tend to encourage FIFO violations.

We hypothesize that differentiating AV technology installation and usage may be relevant to transportation systems planning and policy decision making. AV usage may be influenced by the overall AV installation rate, law enforcement, network performance (e.g., average networkwide speed), and traffic composition (HV-AV ratio). On the other hand, network performance is influenced by traffic composition and network investment. We will explore the conditions, under which the network stakeholders might make seeming rational policy and investment decisions that lead to rewarding HV's rule-violating behaviors and penalizing the attractiveness of AV technologies. For a concrete example, when adding solid double yellow lines to prohibit the left turns to driveways on the opposite side of a roadway might be a paradoxically contributing factor for more users to switch to the manual mode. The realization of such situations can serve as a reminder for analysts, urban system engineers, and policymakers to think in broader scope and to take more realistic human behaviors into account (including their owns); a system dynamics (SD) approach is well-suited for studying such complex system interactions.

An SD approach is particularly suitable for capturing complex systems that contain interacting stock and flow variables and is characterized by positive and negative feedback loops often mistakenly thought as exogeneous by analysts and decision-makers (Richardson, 2011). Abbas and Bell (1994) list multiple advantages of SD for strategic policy analysis and decision supporting tool development. Shepherd (2014) summarized some common applications of SD in transportation. Particularly relevant to this study are the applications in adoption forecast of the uptake of vehicles with alternative fuel and automation technologies (e.g., Struben and Sternman, 2008; Kwon, 2012; Gruel and Stanford, 2016; Niuwenhuijsen et al., 2018), infrastructure maintenance funding allocation (e.g., Guevara et al., 2017), and strategic urban policy decisions (e.g., Pfaffenbichler, 2011). One main reason that we adopt the SD approach, rather than an agent-based approach, is that more unformal knowledge and insights (e.g., panel discussions in conferences and webinars) are available about the macroscopic (aggregate) phenomena than formal data in behavioral details of the present subject.

In a SD model, stock variables are commonly used to capturing the accumulation (or decumulation) process of materials and information. For example, the size of a vehicle fleet can be captured as a stock variable, where purchase (inflow) and shedding (outflow) change the fleet size over time. Similarly, a stock variable can be used to capture the total mileage of a roadway system, and the construction (inflow) and demolition/deterioration (outflow) change the state of the stock. Stock variables can also be used to capture abstract (non-material) entities and their "inertia" over time. For example, the traffic law enforcement resources available of the current year is partially influenced by the funding level determined in the last fiscal year, while the decision about the funding level in the last fiscal year might be partially influenced by the average rate of incidents in the past several years due to delayed data collection and the wait of a sufficient sample size. Loosely speaking, a stock variable at a given moment contains "memory" about its past self. Whether a variable can be seen as a stock variable or an auxiliary variable (no "memory" about its value in the past) depends on the study scope. In a simulation where the time step is in, say, minute, the traffic condition should be considered stock since whether the roadway system is congested in the past minute influences whether the system is congested in the present minute. However, in a simulation where the time step is, say, year, it might be reasonable to assume that the traffic pattern can be fully determined by the transportation network and the land use pattern of that year, and therefore, might be seen as a regular auxiliary variable to sufficiently serve the purpose of a study. On the other hand, this step size (year) might be considered short for the change of the transportation network and the land use pattern, which, therefore, should still be treated as stocks.

Next, we review some relevant studies in the process of developing of a high-level causal loop diagram (CLD). Then we specify the SD model structure and its five main components based on this CLD. We apply the model to the Washington D.C. Metropolitan Area to demonstrate the model and propose policy and investment implications obtained from the sensitivity analysis. We conclude by proposing the importance of taking FIFO violation and traffic law enforcement into account in forecasting AV usage, network performance, and policy and project evaluation. Throughout the study, we assume that, within the study horizon, regulations and vehicle manufacturers will not adopt extreme measures such as removing driving wheel to force people to use AV technologies. In other words, users always have the option to control their vehicles even in Level 5 automation.

2. Causal Loop Diagram (CLD) and Relevant Studies

This section describes a high-level causal loop diagram (CLD) that serves as the conceptual framework for the more concrete SD model to be specified in the next section. Such a conceptual framework can assist the readers in developing an intuition on how different model components interact. A CLD is often used to explain or hypothesize the causal relationship among variables within the system under a study. Figure 1 illustrates the overarching CLD associated with the analytical structure of the proposed SD model. Different factors are connected through directed links. Each arrow of a link is associated with a polarity mark ("+" or "-") showing the qualitative causal relationship. Positive (negative) marks suggest that the increase (decrease) of the cause variable raises (reduces) the effect variable than *what it would have been*. For example, the link from "FIFO Violation Intensity" to "Congestion in Favor of FIFO Violation" has a positive polarity ("+") to suggest that the former tends to increase the latter than *what it would have been*. Note that it is possible that the latter still decreases when the former increases due to the existence of other factors such as "Law Enforcement" and "HV Fleet Size". In this case, the latter would have decreased even more without the increase of the former. The double crosses indicate delays. Relevant studies are cited along the description; these studies are either the bases of individual causal relationships represented by individual arrows or individual feedback loops formed by a set of arrows.

<ins>Individual Causal Links.</ins> Due to lack of observation (since the AV-HV traffic condition has not started), the proposed CLD can only be developed based on relevant studies with generalization. In a sense, each element of the CLD is not new and has been used broadly and only needs minor generalization. For instance, although we have not observed the research and development (R&D) feedback loops in the era of significant AV market penetration (MP), we can still hypothesize its existence by observing and generalizing the past experiences in

the R&D feedback loops in the auto industry in terms of driver safety technologies, fuel efficiency improvement, alternative fuel technologies, etc.

The interplay between traffic law enforcement and traffic conditions have been studied for decades. de Waard and Rooijers (1994) conduct experiments to evaluate the effectiveness of various methods and intensities of law enforcement on driving-speed violation behaviors. Tay (2009) studies the effectiveness of automated and non-automated traffic enforcement on driving behaviors that violate traffic laws. Mehdizadeh and Shariat-Mohaymany (2020) study the attributes of individuals that tend to break the rules of congestion charging. Fakhrmoosavi et al. (2020) use a mesoscopic traffic simulator to study the network impact of mixed traffic of different vehicle types (including AVs) and heterogenous drivers, where the compositions and distributions of the mixed traffic are exogenous.

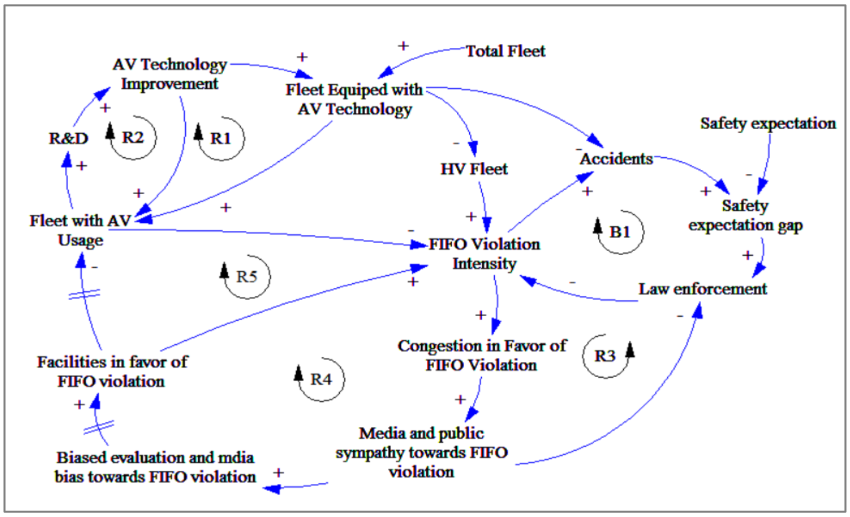


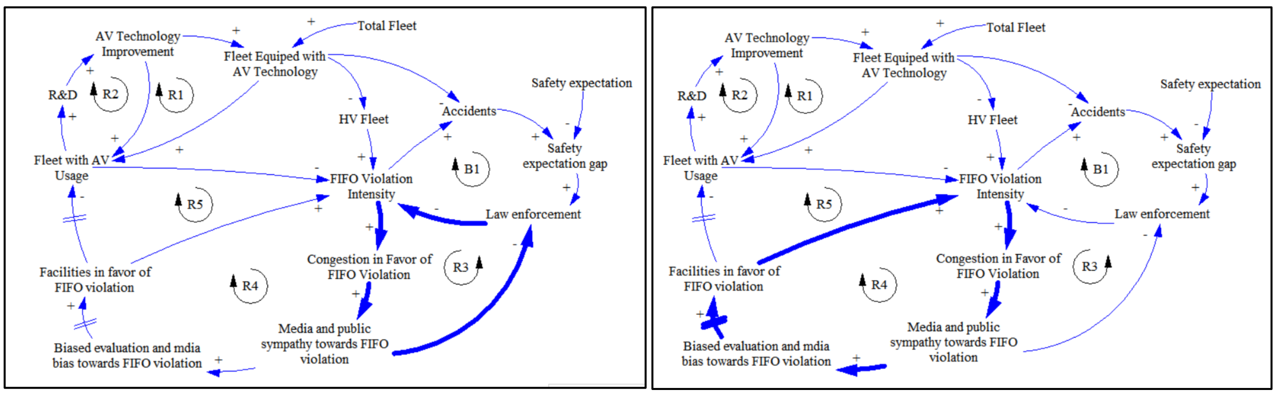
Figure 1. The causal loop diagram of the analytical structure of this paper.

It has been well-recognized that the transition period where HVs and AVs co-exist and share network capacity should not be ignored in evaluating policy, investment, and planning scenarios. Zhao et al. (2020) conduct field experiments about HVs' driving behaviors when following AVs. However, it has been rarely studied for the possibility of HV drivers surpassing AVs in risky manners (e.g., HV drivers making left-turn without yielding AVs from the opposing direction; AV users temporarily switching back to manual mode to surpass the speed limit or make a U-turn at a location that prohibits such movement). Quantifying FIFO violation intensity and its impact on network performance and policy evaluation is critical in studying the dynamics of such mixed traffic, which shall not be confused with the case where users are heterogeneous in terms of preferred driving speed, distance to leading vehicles, reaction time, etc. Jin and Jayakrishnan (2005) develop measurements of FIFO violation and show how it connects to the solution of a commodity-based kinematic wave model for vehicular traffic networks. Mesa-Arango and Ukkusuri (2014) model the asymmetrical car-truck interactions in dynamic traffic flows. Argarwar and Lämmel (2016) model the seepage behavior of smaller vehicles in mixed traffic via agent-based simulations. Ryu et al. (2016) consider the asymmetric interactions among different vehicle types (e.g., the impact of trucks on passenger vehicles tends to be higher than that of passenger vehicles on trucks) in a stochastic traffic equilibrium framework. Yu (2018) proposes a three-dimensional fundamental diagram that captures the relationship among density, speed, and MP of AVs, under different levels of technologies and regulation intensities. Zhong et al. (2020) study the impact on mixed microscopic traffic flow characteristics using high-resolution vehicle trajectory simulation data.

<u>Feedback loops.</u> The individual causal links justified above might form feedback loops, which might be either self-enforcing or self-balancing. Some key loops labeled in Figure 1 are important for developing an intuition of the proposed system dynamics. Loops R1 and R2 capture the R&D feedbacks where more AV technology installation tends to lead to more usage, data, and revenue for the researchers and engineers to further improve

the attractiveness of AV technologies. In Loop R3, more FIFO violation tends to lead to more public and media sympathy, which might cause the reduction of law enforcement and, hence, more FIFO violations than what it would have been. Loops R4 and R5 capture the similar reinforcement but through network project planning and construction with longer delays. Loop B1 captures the phenomena where more FIFO violation tends to cause more traffic accidents which lead to stricter law enforcement than what it would have been. We focus our discussion on Loops R3, R4, and R5 since this paper does not strictly differentiate different levels of automation.

Figure 2(a) and Figure 2(b) highlight the two key loops, R3 and R4, in the discussion of this paper. Other feedback loops are also important, but they have been discussed or at least identified by previous researchers. However, as far as we know, R3 and R4 have not been identified in any formal studies.



(a) R3 loop (b) R4 loop

Figure 2. The two highlighted feedback loops, R3 and R4, are the focus of the discussion in this paper.

The general idea of each loop is not novel and has been proposed and utilized in existing research. The R&D and network facility feedbacks are captured by Struben and Sterman (2008) in modelling the adoption of alternative fuel vehicles. The impact of social exposure and word-of-mouth on adopting new vehicle technologies have been incorporated by Struben and Sterman (2008) and by Shepherd et al. (2012). Causal loop diagram for cycling in London contains the “perceived safety-injury” loop, “bicycle facility investment-safety” loop, and “social normality” loop (Macmillan and Woodcock, 2017). The mutual impact of human biases towards the type of highway projects and the (delayed) resultant system performance have been captured by Guevera et al (2017). Friedman (2006) shows that the mental model underlying the policy of road maintenance to reduce accidents could in most cases *increase* accidents due to ignoring the induced risky driving behaviors (e.g., change of speed), change of traffic volume and composition, the changed polishing effect of the pavement, etc. Time-of-day of day and drug or alcohol use are also identified by Friedman (2006) as factors that might increase the accident rate and worth further studies.

<u>Exogeneity Assumption.</u> In this framework, variables such as total vehicle fleet size dynamics, trip length, and land-use patterns are considered exogeneous and given. The proposed model can be seen as a model built on top of existing ones that do not differentiate AV technology installation and usage, which leads to the proposed model to further incorporate the impact of system performance, safety, roadway planning and construction, and public sentiment into consideration. It is indeed possible that the impact of factors such as system performance and public sentiment influence not only AV technology usage conditioning on AV technology installation but also the AV technology installation itself. Such feedback can be further incorporated in the future study (if the impact is found significant) in a relatively straightforward manner.

Various approaches have been proposed to forecast household-owned vehicle fleets size and general selling prospects (e.g., Hülsmann et al., 2012; Wu et al., 2014). Some focus on the forecast of the share of AVs (e.g., Bansal and Kockelman, 2017; Litman, 2017). Lavieri et al. (2017) study the preference between owning and sharing autonomous vehicle technologies. Kim et al. (2020) examine the potential benefits and concerns associated with AV adoptions. Raj et al. (2020) summarize the barriers to the adoption of AVs based on previous literature and expert inputs. They then propose that industry standards and the absence of regulations and certifications can be addressed to improve customer acceptance.

Sensitivity and elasticity. Due to the considerable uncertainty of future prospects of almost any studies related to factors such as fleet size, technology, and emerging business models, researchers have proposed and utilized various scenario analysis approaches. Wang and Kockelman (2018) discuss three methods − local sensitivity analysis with interaction, Monte Carlo, and Bayesian Melding − for considering the uncertainty of transportation and land-use models. Hsieh et al. (2018) develop a Monte Carlo based approach to sample model parameters in projecting aspects of private car diffusion of the Chinese private car sales and stock. Milkovits et al. (2019) demonstrate and motivate the use of an exploratory modeling and analysis tool, TMIP-EMAT, for transportation systems modeling.

3. Model Specification

Derived from the CLD are the five interactive model components identified as particularly related to the present study. Other various (e.g., those related to the R&D loop) in the CLD have been at least considered in the past, and hence we consider them exogenous in this section to allow the focus on the research gap – the differentiation between the installation and the usage of AV technologies. However, these exogenous variables can be potential incorporated into the proposed model as explicit and endogenous by combining with existing methods (e.g., Struben and Sterman, 2008).

The first component captures the trend of the overall fleet size. The second component captures the trend of the AV technology installation as the portion of the fleet being equipped with AV technologies. The third component captures the planning, construction, and maintenance process of the transportation systems whose stakeholders might be influenced by recent network performance and public sentiment when making planning decisions. This component is also influenced by funding availability and delays from planning and construction. A stock-chaining and co-flow structure is adopted for simulating the cumulative effect of investment on the network and different types of projects in terms of the tendency to favor or disfavor manual driving. The fourth component captures the intensity of FIFO violation and its interplay with the lagged adjustment in traffic law enforcement and the general public sentiment towards manual driving. The fifth component captures the "*equilibrium*" between the travel demand and network capacity in each stimulation step to model the short-term interaction between the network facility and (vehicular) travel demand a macroscopic fundamental diagram augmented to be sensitive to the MP of the usage of AV technologies. Compared with the 45-year simulation horizon (2020 to 2065) with one year as the step size, the impact of day-to-day traffic oscillations during a given simulation year is considered small.

Component 1: Fleet Size Dynamics. The variable fleet size is modeled as a stock whose level at t is determined cumulatively by the purchase rate, ps_t, and the shedding rate, shr_t, from the initial time t_0 to a later time step T ($t_0 \leq t \leq T$). We assume that the changing rate of the overall fleet size is exogenous and given for each scenario. This assumption can be relaxed when combined with existing models whose fleet size is endogenous (but do not differentiate AV installation and AV usage). Following the vein of the formulations used by Struben and Sterman (2008) and by Nieuwenhuijsen et al. (2018), the accumulation of fleet size can be described as:

$$FS_T = FS_0 + \int_{t_0}^{T} (ps_t - shr_t) dt \tag{1}$$

where $ps_t = bfr_t \cdot ad\_bfr_t$ and $shr_t = FS_t \cdot shf$. bfr_t is the base change rate of the fleet size. ad_bfr_t is an exogeneous adjustment factor of bfr_t for considering factors such as the overall economy, innovative technology, and emerging business models for vehicle ownerships. FS_t is the fleet stock at t. shf_t is the percentage of FS_t shredded at t. Technically, shr_t, shedding rate, should be determined by an aging chain of stocks, where each stock represents vehicles with a particular year since usage. A coflow structure can be further added to capture different types of vehicles. In this paper, we choose to use a single stop to simplify the case since the purchasing rate and shedding rate are assumed time-invariant and exogenous.

The active fleet size on roads, veh_t^{or}, is converted from FS_t using three exogeneous variables. p^{or} captures the base percentage which can be estimated from historical data (e.g., total vehicles in usage through travel surveys and total registered vehicles from the departments of motor vehicles). veh_t^{or} is also influenced by its composition (e.g., vans and trucks), which is captured by traffic composition factor, f^{cs}, and the new service business models that have a substitution or a complementary effect, f^{nsbm}. Hence, the operational fleet size for the next time step can be calculated as

$$veh_t^{or} = FS_t \cdot p^{or} \cdot f^{cs} \cdot f^{nsbm} \tag{2}$$

Traffic composition is considered as an exogenous variable to simplify the discussion, though it can be extended as an endogenous variable as well. In a sense, the exogenous f^{cs} can be seen as the passenger car equivalent (PCE) factor for considering the heterogeneous traffic composition. The model does not explicitly model the case where veh_t^{or} is influenced by system performance (e.g., when roadways is highly congested, some people might plan to shift to other travel modes), though such influence with perception delays can be either incorporated by making these factors endogenous.

The following discussion introduces the variables in other components, and to simplify the notation, the subscript "t" are dropped unless confusion might arise.

<u>Component 2: AV Adoption Forecast.</u> We define the MP of AV technology installation to be MP^{AV} and the MP of the actual AV usage to be MP^{AVe}. Similarly, we define the MP of vehicles without AV technologies installed as MP^{HV} and the MP of the manually-driven vehicles (i.e., AVs that are driven manually and HVs) as MP^{HVe}. Based on the above definitions, we have $MP^{AV} + MP^{HV} = 1$ and $MP^{AVe} + MP^{HVe} = 1$, and $MP^{AVe} \leq MP^{AV}$ and $MP^{HVe} \geq MP^{HV}$. To simplify the case, we consider vehicles with Level 4 and Level 5 technologies described by Bansal and Kockelman (2017) as AVs.

MP^{HVe} is a function of MP^{HV} and J. HV FIFO violation intensity, $J \in [0,1]$, will be described in detail in the section on quantifying FIFO violation (Component 4). λ captures the sensitivity of the FIFO violation intensity to the HV usage share, $\lambda > 0$. Eqn. (3) implies that MP^{HVe} should always equal to or greater than MP^{HV}, which reflects the fact that a user with an AV can switch to the HV mode (i.e., choose to drive manually) while a user driving a HV has no option to switch to the AV mode.

$$MP^{HVe} = (MP^{HV})^{J^\lambda} \tag{3}$$

Eqn. (3) also implies that the difference between MP^{HV} and MP^{HVe} is most significant when the mixed degree is moderate. Indeed, when MP^{HV} approaches 0^+, MP^{HVe} approaches 0^+; when MP^{HV} approaches 1^-, MP^{HVe} approaches 1^-. We set br^{adopt}, the base growth rate of MP^{AV}, as exogenous, though it can be relaxed to become endogenous when integrating with existing models that mainly deal with MP^{HV} without differentiating AV installation and AV usage.

<u>Component 3: Planning & Construction of Network Projects.</u> The lifecycle of the infrastructure is captured by a structure of process chains. The changing rate of the generalized network capacity or service capability, $Capa$, measured by equivalent mileage, is proportional to the implementation rate of the network projects, whose changing rate, in turn, is positively influenced by the length of the project list (backlog), PL, and funding availability, AF. The decline of the system performance tends to lead to more scheduled projects added to the project list due to pressure from the public and possibly the additional funding (e.g., dedicated ballot). We referred to the formulations proposed by Guevara et al. (2017) for the formulation development of this component.

For modeling the overall dynamics of the network capacity, we consider two phases of implementation sequentially – planning and construction. In the planning phase,

$$\frac{dPL}{dt} = dPL = pln\big(v(\rho_{eq}), v_f^{mp}, AF, cpp\big) - cstr(PL,\ dlp) \tag{4}$$

where $pln\big(v(\rho_{eq}), AF, cpp\big) = AF \cdot cpp^{-1} \cdot \exp(-\beta_v^{dPL} \cdot (v(\rho_{eq}) - v^*))$ can be seen as the number of projects funded and scheduled (but not built) in a year. cpp the average cost per project, and β_{pln} captures the sensitivity of project planning activities to the system performance. $v(\rho_{eq})$ and v_f^{mp} capture the system performance and reference speed, respectively. More details about system performance are provided in the description in the subsection for Component 5. $cstr$, the construction rate, is assumed to follow a first-order material delay (i.e., $cstr(PL, dlp) = PL \cdot dlp^{-1}$), which is sensitive to the length of the project list -- the longer the project list, the higher the pressure of construction. dlp is the average delay during the planning phase. em is the average project equivalent mileage for capturing the average project contribution to the network capacity measured in mileage equivalent.

$$\frac{dCapa}{dt} = cstrd(PL, em, dlc) - dt(Capa, det) \tag{5}$$

where $Capa$ is the network capacity or service capability in terms of equivalent lane mileage. We assume a first-order material delay, and hence $cstrd(PL, em, dlp) = PL \cdot em \cdot dlp^{-1}$. dt is the outflow rate of the capacity due to deterioration or accidental damage, which is assumed first-order delay (i.e., $dt(Capa, det) = Capa \cdot det$). det is the base deterioration rate, and dlc is the average delay during the construction phase. $Capa_0$ is the length of the project list at the initial time step.

Different types of transportation facilities have different impacts on the FIFO violation tendency. Examples of facilities that are in favor of manual driving (and FIFO violation) are adding new lanes, more road restriction such as double yellow lines no U-turns, and lower speed limit. To consider the impact of different types of projects on the FIFO violation tendency, a co-flow structure is designated to capture the dynamic characteristics of the overall network capacity and that of the portion that tends to induce FIFO violations.

Conceptually, this structure runs in parallel, as coflows, with the overall investment structure just described above. The coflow-specific variables are denoted with the "∼" symbol. Similar to Eqn. (4), we consider the planning list that favors FIFO violations as a stock so that

$$\frac{d\widetilde{PL}}{dt} = \frac{dPL}{dt} \cdot w_{FT} - cstr(\widetilde{PL},\ dlp) \tag{6}$$

where $\widetilde{PL}$ is the project list for the type of projects and designs that tend to induce FIFO violation; $btHV$ is the public and media bias towards manual driving, the details of which can be found in the next subsection for Component 4; and $w_{FT} = \left(1 + \exp(\beta_{btHV}^{w_{FT}} \cdot btHV)\right)^{-1}$ can be seen as a logistic model where the percentage of a project has a tendency to induce FIFO violation at t. Note that when $btHV = 0$, the planning decisions do not favor either type of projects/designs. Similar to Eqn. 5, the counterpart of the systemwide capacity, we have

$$\frac{d\widetilde{Capa}}{dt} = cstrd\big(\widetilde{PL}, em, dlc\big) - dt(\widetilde{Capa}, det) \tag{7}$$

where we assume the same deterioration rate as that in Eqn. (5). The percentage of facilities that tend to incentivize FIFO violation, p_{FT}, is the portion of $\widetilde{Capa}$ in $Capa$. The combined usage of process chains and co-flow structure is illustrated as Component 4 in Figure 5 , where Eqn. (4) and (5) correspond to the process chain on the lower two stocks, while Eqn. (6) and (7) correspond to the process chain on the upper two stocks.

Some network policies and investments might instead favor AVs (e.g., AV only lanes), but we anchor on the overall condition as the reference point. That is, the facilities that are in favor of manual driving are determined by the *relative* impact to those facilities that are either neutral or in favor of automated driving, *not* the neutral impact on FIFO violation.

Component 4: FIFO Violation, Law Enforcement, and Public Sympathy. This component considers the interconnections among manual driving proportion, FIFO violation intensity, the adjustment of law enforcement, and public and media sympathy that might influence the planning bias towards facilities that further incentivize manual driving (and FIFO violation.) The FIFO violation intensity, $J \in [0,1]$, is the key variable of this component. As shown in Eqn. (8), we capture the magnitude of J using two main factors: traffic law enforcement intensity, le, and the percentage of facilities that tend to incentivize HV FIFO Violation, p_{btHV}. Note that J measures the FIFO violation intensity for an "average" manually driven vehicle in a network, not the intensity of the overall network's FIFO violation.

$$J = \left(1 + \exp\left(\beta_{p_{FT}}^{J} \cdot p_{FT} + \beta_{le}^{J} \cdot le\right)\right)^{-1} \tag{8}$$

When le increases and $\beta_{le}^{J} > 0$, J decreases than what it would have been. When p_{FT} increases and $\beta_{p_{FT}}^{J} < 0$, J increases than what it would have been. The intensity of traffic law enforcement depends on the available enforcement resources, LE, as shown in Eqn. (9).

$$le = \left(1 + \exp\left(-\beta_{LE}^{le} \cdot (LE - LE_{t_0})\right)\right)^{-1} \tag{9}$$

where $le \in [0,1)$ is an auxiliary variable of LE - a stock that represents the law enforcement resources to capture delays due to information feedback (e.g., accumulation of accident data and response to the public sentiment), police enforcement funding cycle, recruiting and laying-off process, etc. The sensitivity of enforcement intensity to the law enforcement resources, $\beta_{LE}^{le} \geq 0$, is a parameter to be estimated and LE_{t_0} is the baseline initial resources that can be obtained from the annual financial report from local or regional agencies responsible for traffic law enforcement.

We model the dynamics of the law enforcement resources, LE, through its adjustment to capture the delays from the funding allocation process, police officer recruiting, staff training, etc. Funding availability is not endogenously modeled in this paper, but such constraint can be conveniently incorporated when studying a specific area where the exact funding mechanism and interactions with other funding options are transparent. Three main influencers of LE are FIFO violation intensity itself, media and public sympathy, and the overall percentage of AV usage. However, when FIFO violation is common and the facilities tend to incentivize FIFO violation, the media and public sympathy towards such violation might increase and cause the gradual adjustment rate for traffic law enforcement effort. Traffic collision is inherently captured by J. Let

$$\frac{dLE}{dt} = dLE^{(r)} + \beta_{btHV}^{dLE} \cdot btHV + \beta_{J}^{dLE} \cdot J \tag{10}$$

and $btHV$ is expressed as

$$btHV = sym \cdot \exp\left(\beta_{J}^{btHV} \cdot J + \beta_{p_{FT}}^{btHW} \cdot p_{FT}\right) \tag{11}$$

where sym is exogenous and captures the general sentiments towards manual driving. This can be influenced by external forces and considered as a policy leverage. β_{J}^{btHV} is the coefficient that scales the FIFO violation intensity to capture the influence of traffic condition (travel experience) on the general sentiment. $\beta_{p_{FT}}^{btHW}$ is the coefficient that captures the influence of facilities. In general, the higher percentage of facilities that tend to induce or even encourage FIFO violations in their effect, the more likely that the public tend to show sympathy on such behaviors. On the other hand, people might also turn negative towards manual driving when FIFO violation intensity deems too high. Note that $btHV$ is defined based on its impact on the planning and traffic enforcement, *not* the positive or negative sentiment the public and media have. Indeed, both public revulsion and penchant towards manual driving could lead to biased funding towards changing a two-way left turn lane (TWLTL) into a double-yellow line and induce FIFO violation.

Note that whether a proposed project would accommodate/incentivize FIFO violations is assumed to be only influenced by the media and public opinions and the design conventions. For example, a newly proposed arterial does not require the installation of median barriers except some portions near intersections based on certain design guideline or convention. But due to the potential influence from the public, more physical median barriers might be added or removed. For another example, planners might propose an AV-only lane (assuming such lane has the technologies to detect whether the vehicle is indeed driven autonomously,) but whether such a lane should be constructed as a new lane or converted from a general-purpose or a high-occupancy toll (HOT) lane might be influenced by the public and media sentiment in the public outreach phase. Such a decision might also be influenced by the funding availability, but the funding availability tends to be greatly influenced by the public and median sentiment.

Component 3 and 4 interact through two main mechanisms. First, the composite effect of FIFO violation and percentage of facilities that tend to incentive FIFO violation influence media and public sentiment, which, in turn, influences both the network planning process and law enforcement effort. Second, the FIFO violation influence how common to manually drive vehicles installed with AV technologies, which then influences traffic systemwide performance and the network planning decisions. The next sub-section on Component 5 describes how the model captures the influence of HV and AV usage share on systemwide performance.

<u>Component 5: Equilibrium Density and System Performance.</u> We adopt the generalized network speed, v, as the proxy for system performance, in which travel time reliability and safety are factored. To capture the "*equilibrium*" between the travel demand and network capacity in each stimulation step, a macroscopic fundamental diagram augmented to be sensitive to the MP of the usage of AV technologies is proposed. The formulations of this augmented macroscopic fundamental diagram are mainly a quantitative substantiation of the qualitative relationships proposed by Yu (2018). For the original concept of the macroscopic fundamental diagram, please refer to Godfrey (1969), Vickery (1991), and Geroliminis and Daganzo (2007). System performance is a function of vehicle-based travel demand and the network capacity. We first use veh_{total}^{or} to obtain the generalized density, with two exogenous adjustment factors: adj_{temp} for the uneven temporal distribution of the vehicles on the network, and adj_{spat} for the uneven spatial distribution. The generalized density is further normalized using the generalized maximum density ρ_{max} to obtain the normalized equivalent density ρ_{eq}. The calculation is shown in Eqn. (12).

$$\rho_{eq} = \frac{veh_{total}^{or}}{Capa} \cdot adj_{temp} \cdot adj_{spat} \cdot \frac{1}{\rho_{max}} \tag{12}$$

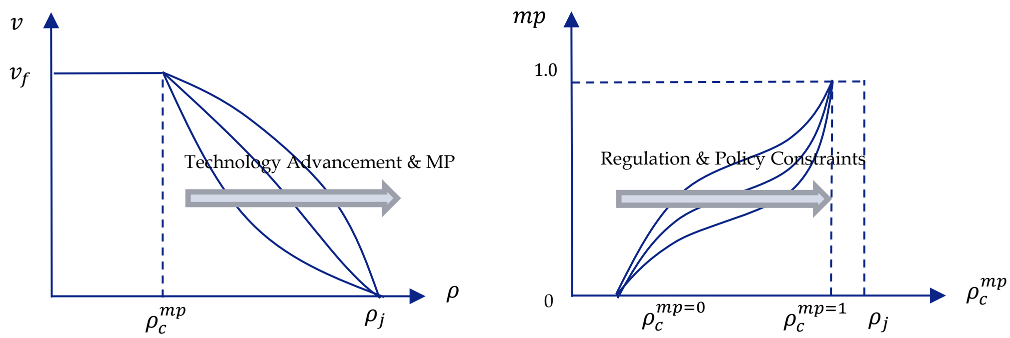
where the equivalent network capacity, $Capa$, is a function of present network capacity and newly added capacity, as described in the subsection for Component 3.

The three-dimensional fundamental diagram (Yu, 2018) is utilized to capture the macroscopic network performance for the AV-HV mixed traffic so that the network performance is sensitive to the MP of AV technology usage, MP^{AVe}, and the (equivalent) traffic density, ρ_{eq}. Eqn. (13) formulates the subspace formed by the density and the generalized network speed given a specific MP, denoted as, mp.

$$v(\rho_{eq}) = \begin{cases} v_f^{mp}, \ \rho_{eq} < \rho_c^{mp} \\ v_f^{mp} \cdot \left(\frac{\rho_j^{mp} - \rho_{eq}}{\rho_j^{mp} - \rho_c^{mp}}\right)^{\alpha}, \ \rho_{eq} \geq \rho_c^{mp} \end{cases} \tag{13}$$

where $mp = 1 - MP^{HVe} = MP^{AVe}$ in the specific context of this paper, v_f^{mp} is the reference system speed at mp, and α is for capturing the technology advancement. When v_f^{mp} is not sensitive to mp, we simplify it as v_f. Equivalent density and the HV usage share (i.e., MP^{HVe}) are used as the two main factors that the system decision makers consider in project programming. ρ_j^{mp} is the jam density given mp. When the vehicle length is not correlated with whether a vehicle is AV or HV, $\rho_j^{mp} = \rho_j, \forall mp \in [0,1]$. Eqn. (13) is illustrated in Figure 3(a), as postulated by Yu (2018).

Next, we consider the subspace formed by the normalized density, $\breve{\rho}_c^{mp}$, and mp, as a function of AV usage share (and hence of FFIO violation intensity.) Yu (2018) postulates a "*S*" shaped nonlinear relationship as illustrated in Figure 3(b). Let the function that describes such a relationship be $\breve{\rho}_c^{mp}(mp)$, where $\breve{\rho}_c^{mp} = \frac{\rho_c^{mp} - \rho_c^{mp=0}}{\rho_c^{mp=1}}$. $\breve{\rho}_c^{mp=0}$ and $\breve{\rho}_c^{mp=1}$ are the critical density when $mp = 0$ and $mp = 1$, respectively. $\gamma > 0$ allows further adjustment on the "*S*" shaped curvature to capture the impact of different regulation and policy constraints so that the critical density threshold, $\breve{\rho}_c^{mp}$, to be used for calculating system performance is $\left(\breve{\rho}_c^{mp}(mp)\right)^{\gamma}$. In the numerical simulation, this relationship can also be expressed using a lookup table, whose curvature can be manually adjusted to reflect the regulation and policy impact. Note that $\breve{\rho}_c^{mp}$ in Figure 3 (b) as a function of mp, is the input for obtaining v in Figure 3(a). Without loss of generality, $\breve{\rho}_{jam}$ is set as 1.0 to be the normalized jam density.



(a) (b)

Figure 3 (a) The hypothetical impact of technology advancement and MP on the relationship between the equivalent system speed and equivalent system density; (b) the hypothetical impact of regulation and policy constraints on the relationship between the MP of AV usage and the equivalent traffic density.

The overall three-dimensional fundamental diagram (density-speed) is illustrated in Figure 4. ρ_j is fixed by assuming that the average vehicle dimension is unchanged. The assumption $v_{max}^{mp=0} = v_{max}^{mp=1}$, in this paper, can be conveniently relaxed by making $v_{max}^{mp=p}$ as a function of AV usage share.

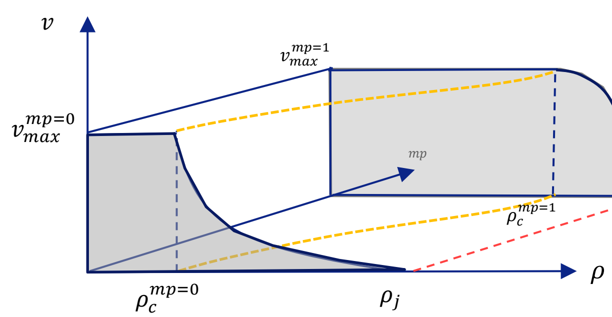


Figure 4 Generalized MP-v-ρ relationship proposed by Yu (2018).

This component is mainly used to capture how network stakeholders make decisions. It is arguable that planning activities are usually based on the "projected" or "expected" situations. However, we use the "present" system performance for three main reasons. First, it is still common for public agencies to make reactive decisions instead of proactive ones (even though many claim they are proactive) due to lack of funding and high pressure from the voters to deliver immediate results or before the next election. Second, how the forecast is done is typically influenced and framed by the present conditions. Indeed, many forecasts are used to prove that the stakeholder's intuition is right instead of guiding the decisions (Sternman, 2000). Third, we put additional effort to keeping the model as simple as possible to show the most important dynamics related to the core subject of this paper, and it is relatively straightforward to further incorporate the "projection" and "expectation" into this component when needed.

<u>System Dynamics of the Five Components.</u> The system dynamics model is illustrated in the diagram shown in Figure 5. The text of each variable is color-coded to reflect which one of the five model components it belongs to.

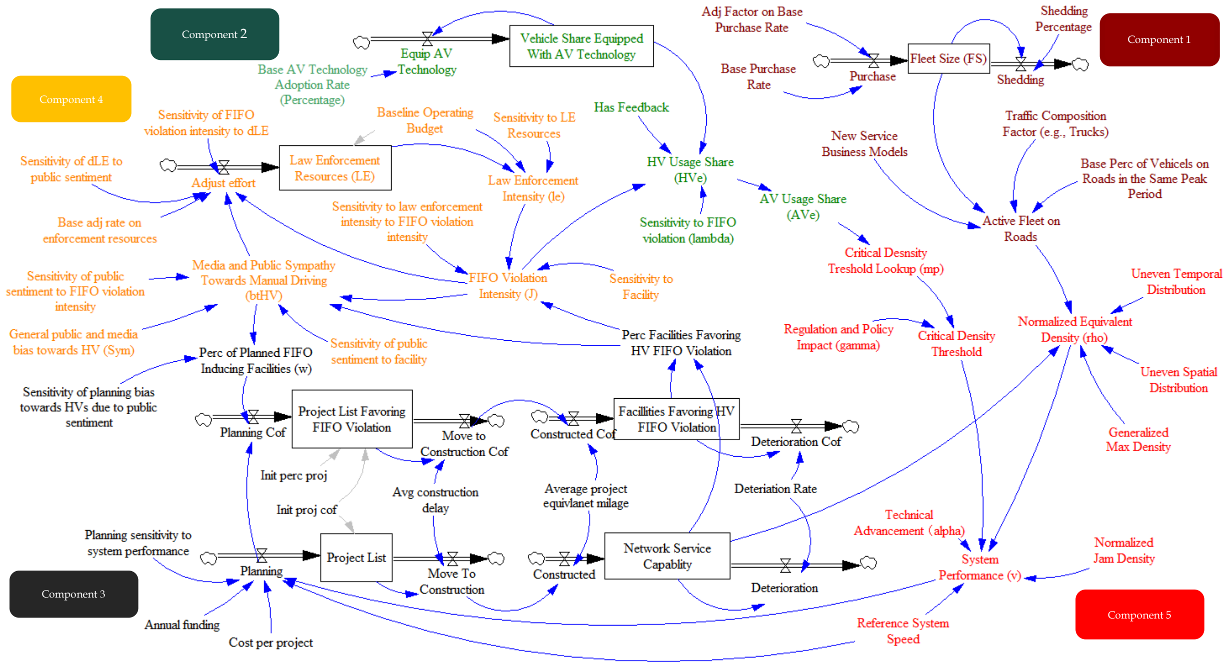


Figure 5 System dynamics diagram of the interaction of the five main components

4. Case Study & Scenario Analysis

<u>Background & Baseline Development.</u> The greater Washington, D.C. area is chosen for a case study to demonstrate the model and provide general policy implications. The base rate fleet size is based on a simple projection of the county-level household estimates of the 2012-2016 5-year estimates from the Census Transportation Planning Products (CTPP) data, a special tabulation of the American Community Survey (ACS) data. Figure 6 shows the selected county-level geographies that form the study area. Table 1 shows the estimates and their 90-percentile margin of error (MOE) of the available household vehicles in each geography. This scope is generally consistent with that of the National Capital Region Transportation Planning Board (TPB)[1] model.

[1] The National Capital Region Transportation Board (TPB) is the federally designated metropolitan planning organization (MPO), housed at the Metropolitan Washington Council of Governments (MWCOG).

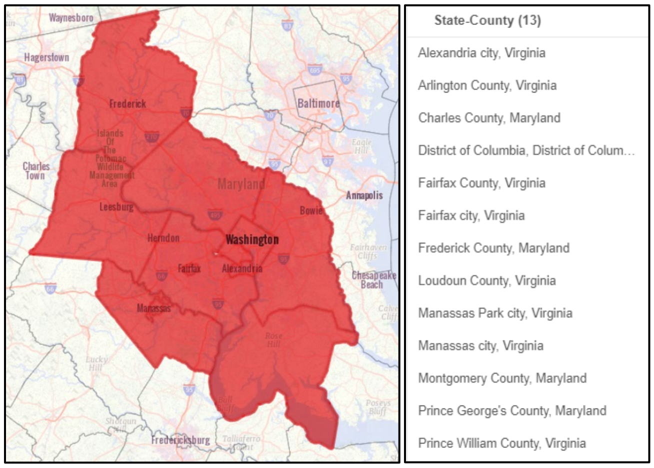


Figure 6 The selected county-level geography based on the definition of the MWCOG/TPB Region

Table 1 The 2012-2016 ACS/CTPP estimates of household vehicles available and the margin-of-errors.

Geography	Vehicles Available	MOE (90%)
District of Columbia	276,545	±1,295
Montgomery County, MD	367,765	±1,241
Prince George's County, MD	306,710	±1,132
Arlington County, VA	100,705	±915
Fairfax County, VA	393,360	±1,138
Alexandria city, VA	68,065	±684
Fairfax city, VA	8,475	±135
Falls Church city, VA	5,300	±130
Total	*1,526,925*	*±2,671*

*source: 2012-2016 Estimates from the Census Transportation Planning Product (CTPP), a special tabulation of the American Community Survey (ACS) data.

The vehicle trend is mainly developed based on the household forecast made by the TPB model baseline scenario. The effect of innovative business models (e.g., car-sharing services and virtual shopping) on the reducing needs of owning vehicles is considered exogenous. The base and the sensitivity analysis range of the AV technology adoption rate, br^{adopt}, is developed based on the study by Bansal and Kockelman (2016). To simplify the demonstration, we use Level 4 (inclusive) as the threshold where we refer to a vehicle to be equipped with AV technologies. The project size is developed based on the Regional Transportation Improvement Program (RTIP) of the MWCOG, released on March 18, 2020 (The Metropolitan Washington Council of Government, 2020). Table 2 to Table 6 list the variables and parameters used in the model, along with the variable type, units, and baseline values for each model component.

The model is further tuned using the data between 2000 and 2019 from the city-wide and regional traffic fatality report (Metropolitan Police Department, 2020), police budget and full-time equivalent (Office of the Chief Financial Officer, 2019), active vehicle registrations (Department of Motor Vehicles), ABS installation and other safety measures (National Highway Traffic Safety Administration, 2020), recorded traffic violations and the enaction of the automated traffic enforcement program (District Department of Transportation, 2020), etc. It is challenging to completely isolate the budget portion for traffic enforcement (e.g., general training on administration and law enforcement skillsets can improve law enforcement effectiveness), so we used the overall budget of the Police Department budget (with Year 2019 as the baseline) to approximate. We assume that v_{max} is not sensitive to the MP to simplify the analysis, though this assumption can be relaxed by modifying Eqn. (13). The original starting year is 2020 and horizon year is 2065. We set modeling time horizon to be from 2022 to 2067, where 2022 is assumed to be the official market entrance time of the AV technologies. Since the same generalized network density can be achieved from infinite possible combinations of network capacity, fleet size, and PCE factor, we simply fixed f^{cs} in the baseline as 1.05 to adjust others without loss of generality. This can be easily modified when a specific PCE is known for a region.

A set of tests for assessment specified by Sternman (2000) is used to adjust the model to increase the confidence that the proposed model can be useful for the study objective. We deem the model boundary adequate for the differentiation between AV usage and AV installation to be mainly interacted with the traffic condition, network facility planning decisions, traffic law enforcement, and general public sentiment. We also leave sufficient space for further extension to consider more factors endogenously. For example, the installation of AV technologies may be influenced by how transportation facilities are "AV-friendly." The model structure, though significantly simplified from the real-world situations, is considered sufficiently homeomorphic to the actual systems. The stocks such as fleet size, project list, and network lane miles are verified to be conserved, and the generalized network lane miles for facilities that tend to induce FIFO violations are always smaller than the total network lane miles. The four main types of system "players" – travelers, law enforcement forces, network planners, and the general public are verified to be bounded by the actual information available to them. The units are confirmed to be consistent within the model.

As this paper is being written, no observation is available about official market entrance by the AV technologies, so it is impossible to estimate the parameters using real-world data directly. However, we utilize the existing driver assistant technologies (e.g., anti-lock braking system, lane-departure warning, parking functions) as a reference to assess the reasonableness of the parameters and the model behaviors. For example, it is not uncommon for drivers to turn off the lane-departure warnings (Forbes, 2016), and therefore, it is logical that the installation percentage of AV technologies should always be equal to or higher than the actual usage percentage of AV technologies. This type of test can also be counted as a "family member" test, where the proposed model is shown to be able to generate reasonable behaviors in similar system dynamics. Extreme conditions are further tested to the boundaries of the theoretical ranges of each parameter specified in Table 2 to Table 6, where we used large numbers to represent negative and positive infinite values. We cut the time step in half and do not observe significant changes in the model behavior. The key long-term loop is removed for the behavior anomaly test, as shown in the comparison between the "Baseline" and the "NoDiff" scenario in the latter part of the present section. Surprise behavior tests and sensitive tests are also conducted, parts of which are shown and discussed in the latter part of the present section and the next section. Based on these tests, we fail to falsify the proposed model for the specific application context, and we are convinced that the model may provide useful insights to policy and public investment decision-making.

Table 2 (Component 1) Model variable, description, type, unit, theoretical range, and their baseline value.

Variable	Description	Type	Unit	Baseline Value
FS	Fleet size	Stock	Households Vehicles	Init: 1.6E+6
ps	Purchasing rate	Inflow (Auxiliary)	Vehicles Per Year	--
bfr	Base purchase rate	Exogenous	Vehicles Per Year	40,000
ad_bfr	Adjustment factor on the base fleet change rate	Exogenous	Diml	1.0
shr	Shedding rate	Outflow (Auxiliary)	Vehicles Per Year	--
shf	Shedding Percentage	Exogenous	Percent Per Year	1.1
veh^{or}	Active Fleet on Roads	Auxiliary	Vehicles	--
p^{or}	Percentage of active/operational vehicles	Exogenous	Percent	39
f^{cs}	Traffic composition factor to convert to passenger car equivalent (PCE)	Exogeneous	Diml	1.05
f^{nsbm}	Substitution or complementary effect of new business models and technologies	Exogeneous	Diml	0.226

* "Diml" in this column refers to dimensionless units.
* Auxiliary variables are determined by other variables, and therefore, the baseline values show as "--".

Table 3 (Component 2) Model variable, description, type, unit, theoretical range, and their baseline value.

Variable	Description	Type	Unit	Baseline Value
MP^{AV}	MP of vehicles equipped with AV technologies	Stock	Percent	Init: 0
MP^{HV}	Market penetration of vehicles without AV technologies	Auxiliary	Percent	--
MP^{AVe}	AV usage share (equivalent MP of vehicles using AV technologies to drive)	Auxiliary	Percent	--
MP^{HVe}	HV usage share (equivalent MP of vehicles not using AV technologies to drive)	Auxiliary	Percent	--
λ	Sensitivity to FIFO violation	Exogenous	Diml	0.235
br^{adopt}	Base (reference) rate of MP^{AV}	Exogenous	Percent (as decimal) Per Year	2.1

* "Diml" in this column refers to dimensionless units.
* Auxiliary variables are determined by other variables, and therefore, the baseline values show as "--".

Table 4 (Component 3) Model variable, description, type, unit, theoretical boundary, and their baseline value.

Variable	Description	Type	Unit	Baseline Value
PL	List of scheduled network projects	Stock	Projects	Init: 25.0
$\widetilde{PL}$	List of scheduled network projects that tend to induce FIFO violation	Stock	Projects	Init: 7.5
$Capa$	Generalized network capacity	Stock	Lane Miles	Init: 2600.0
$\widetilde{Capa}$	Generalized network capacity for facilities that tend to induce FIFO violation, as a subset of the overall network capacity	Stock	Lane Miles	Init: 500.0
AF	Available funding	Exogenous	US Dollars	5.0E+07
cpp	Average cost per project	Exogenous	US Dollars Per Project	4.7E+05
β_v^{dPL}	Planning sensitivity to system performance	Exogenous	Diml	-0.05
p_{FT}	Percent of facilities that tend to induce HV FIFO violation	Auxiliary	Percent	--
w_{FT}	Planning judgement on the appropriate portion of facilities that tend to induce HV FIFO.	Auxiliary	Percent	--
$\beta_{btHV}^{w_{FT}}$	Sensitivity of planning bias towards HVs due to public sentiment	Exogeneous	Diml	-4.5
dlp	Delay in the planning phase	Exogenous	Years	2.79
dlc	Delay in the construction phase	Exogenous	Years	3.24
det	Deterioration rate as percentage	Exogenous	Diml	0.001
em	Average project contribution to network capacity (mileage equivalent)	Exogenous	Mileage Per Project	4.125

* "Diml" in this column refers to dimensionless units.
* Auxiliary variables are determined by other variables, and therefore, the baseline values show as "--".

Table 5 (Component 4) Model variable, description, type, unit, theoretical boundary, and their baseline value.

Variable	Description	Type	Unit	Baseline Value
J	FIFO violation (including traffic law violation) intensity	Auxiliary	Diml	--
LE	Law enforcement resources	Stock	Million Dollars	Init: 553.54
le	Law enforcement intensity	Auxiliary	Diml	
$btHV$	Media and public sympathy/bias towards HVs	Auxiliary	Dmnl	--
Sym	General public's sentiment towards HV driving	Exogenous	Diml	1.0
$dLE^{(r)}$	Reference annual adjustment rate on traffic law enforcement resources	Exogeneous	Million Dollars	9.0
β_{LE}^{le}	Sensitivity of law enforcement intensity to the resources of law enforcement.	Exogenous	Diml	0.01
$\beta_{p_{FT}}^{J}$	Sensitivity of FIFO violation intensity to the percent of facilities that tend to induce FIFO violation.	Exogeneous	Diml	-3.1
β_{le}^{J}	Sensitivity FIFO violation intensity to of law enforcement intensity	Exogenous	Diml	3.0
β_{btHV}^{dLE}	Sensitivity of law enforcement resource adjustment to public sentiment	Exogenous	Diml	-0.1
β_{J}^{btHV}	Sensitivity of public sentiment to FIFO violation intesntiy	Exogenous	Diml	1.5
β_{btHV}^{dLE}	Sensitivity of the adjustment of law enforcement resources to public sentiment.	Exogenous	Diml	20.0
$\beta_{p_{FT}}^{btHW}$	Sensitivity of the percent of facilities that intend to induce FIFO violation to public sentiment	Exogeneous	Diml	2.5

* "Diml" in this column refers to dimensionless units.
* Auxiliary variables are determined by other variables, and therefore, the baseline values show as "--".

Table 6 (Component 5) Model variable, description, type, unit, theoretical boundary, and their baseline value.

Variable	Description	Type	Unit	Baseline Value
ρ_{eq}	(Normalized) equivalent traffic density for the macroscopic fundamental diagram	Auxiliary	Vehicles Per Mile	--
adj_{temp}	Adjustment factor for considering temporally uneven distribution of traffic	Exogenous	Diml	0.0913
adj_{spat}	Adjustment factor for considering spatially uneven distribution of traffic	Exogenous	Diml	0.0526
α	Technical advancement	Exogenous	Diml	0.89
γ	Regulation & policy impact factor	Exogenous	Diml	0.88
v_f	Reference systemwide speed	Exogenous	Miles/Hour	55
ρ_{max}	Generalized maximum density	Exogenous	Vehicles Per Mile	170
$\breve{\rho}_j$	Generalized and normalized Jam density	Exogenous	Vehicles Per Mile	1.0
$\breve{\rho}_c^{mp}$	Generalized and normalized critical density as a lookup function of mp (with linear interpolation), where $mp \in [0,1]$	Auxiliary	Vehicles Per Mile	(0.0, 0.1), (0.153, 0.123), (0.297, 0.175), (0.422, 0.276), (0.529, 0.417), (0.587, 0.570), (0.661, 0.728), (0.740, 0.846), (0.844, 0.912), (1.0, 0.943)

* "Diml" in this column refers to dimensionless units.
* Auxiliary variables are determined by other variables, and therefore, the baseline values show as "--".

Figure 7 compares the baseline scenario ("Baseline") with the "No Differentiation (NoDiff)" scenario. By "NoDiff," we refer to the scenario where the network stakeholders and decision-makers do not differentiate the installation and the usage in their evaluation and forecast of the system performance. Specific to the proposed model, "NoDiff" is reflected through setting the AV usage rate equals to the AV installation rate. This way, the network stakeholders tend to overestimate the AV usage when planning for network projects. The comparison serves two main purposes. First, it quantifies the impact of differentiating the AV installation and AV usage *in the planning process* with all else equal. Second, it serves as a validation and stress testing where the only difference is whether the planners and decision makers would accommodate the impact of manually driven AVs on the network performance by adjusting the overall network capacity.

The six selected endogenous variables are: AV usage share, system performance, project list, network service capacity, percentage of facilities that tend to induce FIFIO violation, and FIFO violation intensity. Not surprisingly, the AV usage share is always lower than the AV installation rate since users of vehicles equipped with AV technologies can choose to drive manually. The system performance tends to be overestimated in the "NoDiff" scenario compared with the "Baseline," especially when the share of vehicles with AV technologies installed is similar to that of their counterpart. In the "NoDiff" scenario, there is a tendency to under-invest the network, compared to the "Baseline" due to the overoptimism. Among those under-invested projects, the "NoDiff" scenario tends to plan more projects that may incentivize more FIFO violations. There exists minor "overshoot" for the project list or backlog when AV technologies start to enter the travel market. Which can be interpreted as that the decision-makers tend to overreact to the need for network projects due to the information feedback delay about the effect of planning and construction. The non-zero project list after around 2052 for the "Baseline" (and 2060 for the "NoDiff) is mainly due to the accommodation effort of the increase of the active fleet size and the network deterioration. For FIFO violation intensity, the small difference between the two scenarios is mainly because FIFO violation exists regardless of whether network stakeholders take it into account in their planning process.

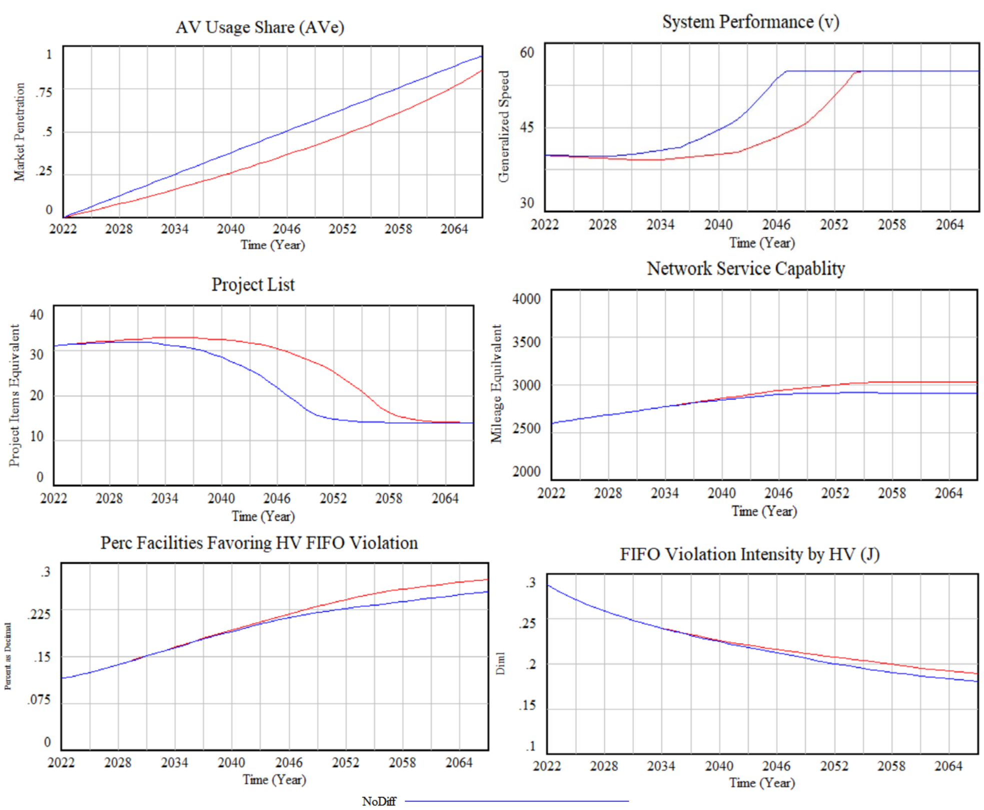


Figure 7 The dynamics of the six selected endogenous variables in the "Baseline" scenario and the "NoDiff" scenario.

<u>Sensitivity Analysis.</u> The robustness and sensitivity of all exogenous variables are tested, which are anchored at the values in the "Baseline." We present in detail three main exogenous variables that may have particularly

useful policy implications: “Base Adjustment Rate on Enforcement Resources,” “Base AV Technology Adoption Percentage Change,” and “Public and Media Bias Towards Manual Driving.” These three variables, selected from the domains of policy leverages (law enforcement, auto management/administration, and media regulations), are considered critical to understanding the model and providing policy implications. Each leverage is first studied while holding all other exogenous variables as those in the “Baseline.” As an exploratory analysis, the variable(s) are assumed to be uniformly distributed within the test range. Then, random perturbations of all three variables are used to study their combinatory effect.

The variable “Base Adjustment Rate on Enforcement Resources (Million Dollars)” is tested in the range between -6.0 and 12.0 (uniform distribution). When positive, the higher (lower) this variable is, the faster (slower) the resources related to the traffic law enforcement grow with all else equal. Similarly, when negative, the lower (higher) this variable is, the faster (slower) the resources decrease with all else equal. The testing results of four selected variables are shown in Figure 8.

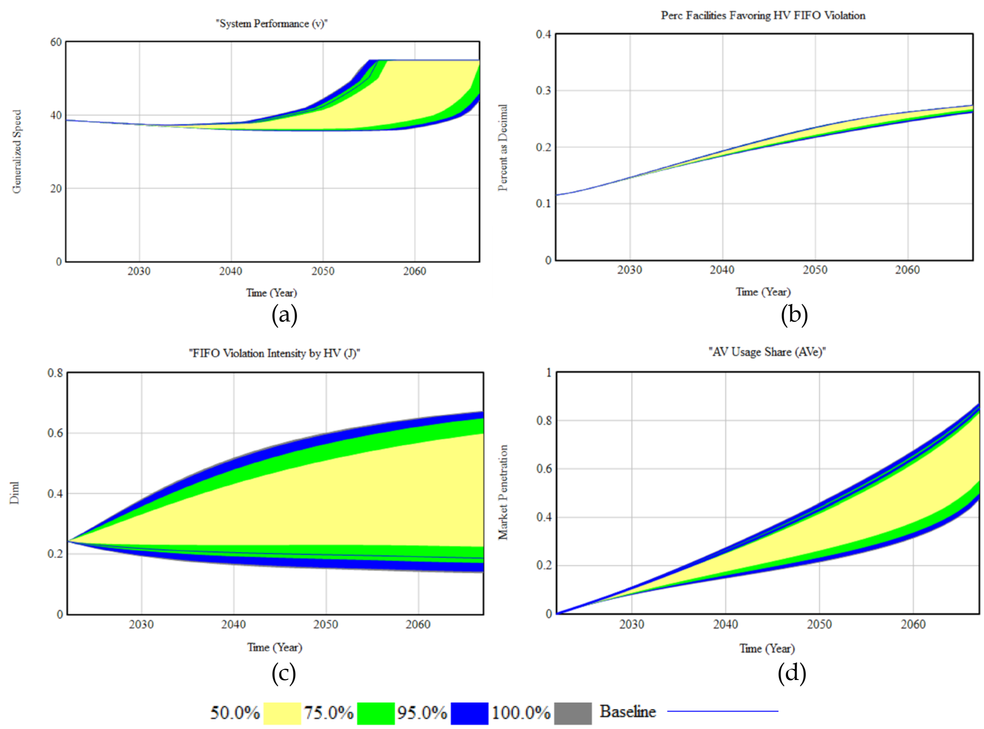


Figure 8 Univariate sensitivity analysis on “Base Adjustment Rate on Enforcement Resources”

From Figure 8(a), each percentile range has a clear trend of improvement until the maximum allowable speed (threshold) is reached. This threshold is determined jointly by technological and policy constraints. The pattern of the percentiles shows that the variable influence when the threshold similarly until reaching the 95 percentile, which suggests a “small” tail of the outcome distribution. Figure 8(b) shows the percentage of facilities that tend to induce FIFO violation. Due to the planning and construction delay, the effect of system capacity responds to the change of system performance later. Also note that the outcome range is smoother than the system performance since it was “buffered” (or “smoothed”) by the four stocks in Component 3. The outcome range of this variable is relatively small, suggesting that the law enforcement tends to have a smaller impact on it. Figure 8(c) shows the influence of the tested variable on the range of the FIFO violation intensity. Figure 8(d) shows that the MP of AV usage has a similar trend to that of the AV installation but always lower than the MP of the AV installation.

The variable “Base AV Technology Adoption Percentage Change” is tested in the range of 0.005 and 0.031. This variable reflects the change of circumstances external to the modeling scope on the base AV technology installation rate. The smaller (greater) the variable is, the slower (faster) the MP of vehicles installed with AV technologies. The testing results of four selected variables are shown in Figure 9.

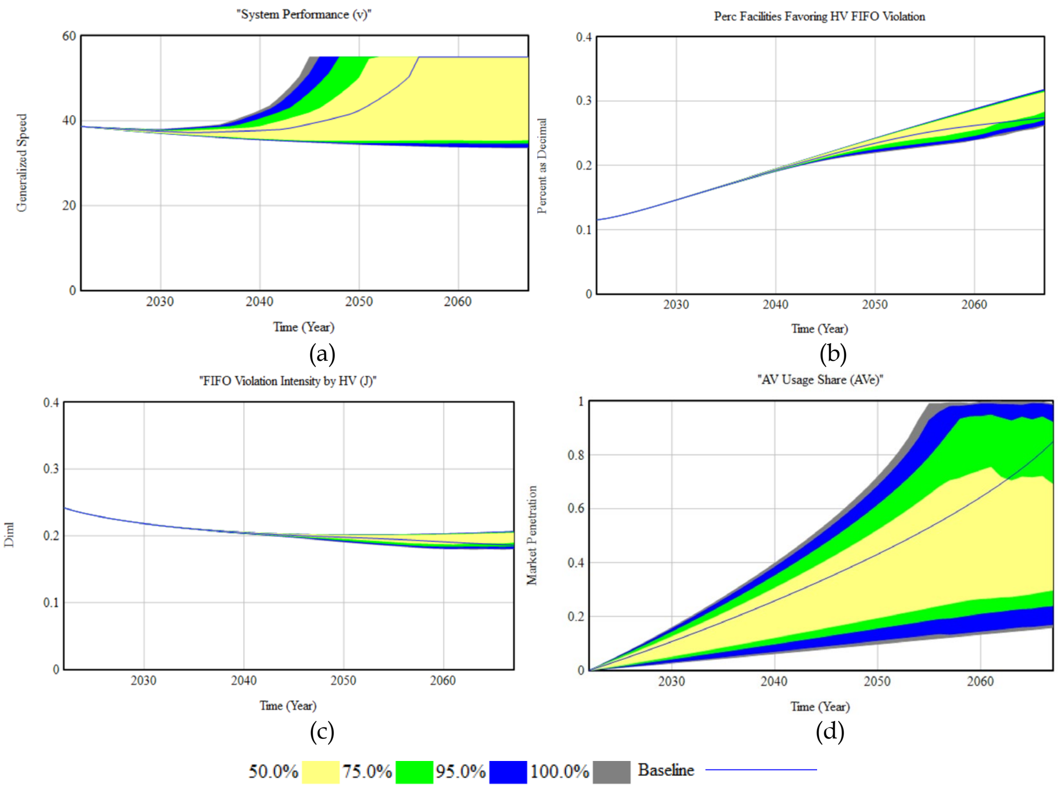


Figure 9 Univariate sensitivity analysis on "Base AV Technology Adoption Percentage Change."

Figure 9(a) shows that each percentile range has a clear trend of improvement until the maximum allowable speed (threshold) is reached. This threshold is determined jointly by technological and policy constraints. The pattern of the percentiles shows high sensitivity, which suggests a "small" tail of the outcome distribution. Also, it seems possible for the system performance to be worsened when the baseline funding constraints cannot keep up with the increasing demand of the network capacity without the traffic mitigation effect of the AV usage. Figure 9(b) shows the percentage of facilities that tend to induce FIFO violation. Due to the planning and construction delay, the effect of system capacity responds to the change of system performance in the later phase of the simulation. The range of the 50 percentile is relatively wide suggesting a high sensitivity of the variable around the baseline value. The overall outcome range of this variable is relatively small, suggesting that the law enforcement tends to have a smaller impact on it. Figure 9(c) shows the influence of the tested variable on the range of the FIFO violation intensity. In addition to the MP of AV usage having a similar trend to that of the AV installation but always lower than the MP of the AV installation, Figure 9(d) also suggests that the AV usage share starts having minor irregular oscillation when the MP of AV usage reaches 100%. This is probably because when most vehicles are automated, it becomes safer to drive manually since other vehicles tend to do what they "can" to avoid collisions with manually driven vehicles and are significantly better at it than human drivers, which gives human incentives to switch back to manual mode.

The variable "Public and media bias towards HV" is tested in the range of 0.1 and 0.9. This variable captures the public's general sentiments towards manual driving. This general sentiment can be influenced by both external factors (e.g., political entities promote the freedom of choice between manual and automated driving through new networks and social media) as well as usable policy leverage (e.g., regulations on requiring media to proportionally report AV-involved incidents and HV-involved incidents) with and without delays as a potential extension of the model. When this variable is large (small), the public sentiment tends to be positive (negative). Such as positive (negative) sentiment might further cause the increase (reduction) of the bias in favor of manual driving and might increase (decrease) the planning and design bias towards manual driving. The testing results of four selected variables are shown in Figure 10.

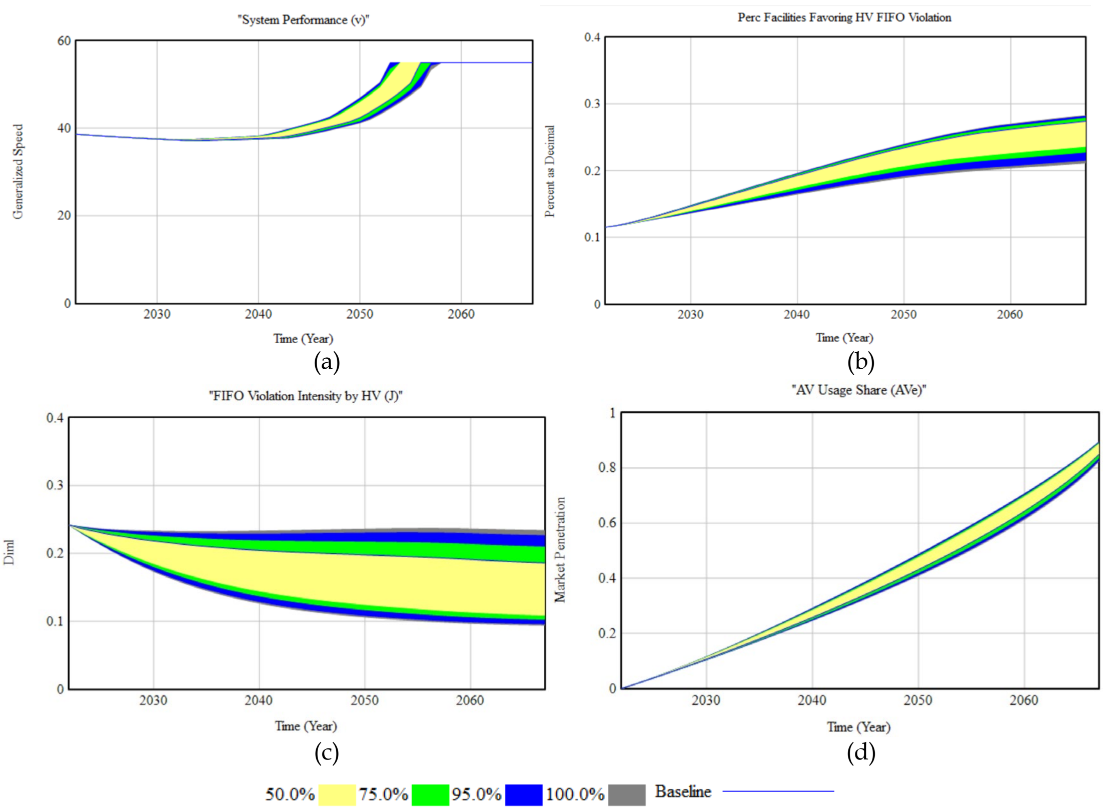


Figure 10 Univariate sensitivity analysis on "Public and Media Bias Towards Manual Driving."

Figure 10(a) shows that each outcome percentile range for system performance, though narrow, has a clear trend of improvement until the maximum allowable speed (threshold) is reached. Figure 10(b) shows the percentage of facilities that tend to induce FIFO violation. The outcome range of this variable is relatively small, suggesting that the law enforcement tends to have a smaller impact on it. Figure 10(c) shows the influence of the tested variable on the range of the FIFO violation intensity. The relatively wide range is mainly due to the high sensitivity of law enforcement effort in the short term and the high sensitivity of the facility types in the long term. It is also noticeable that the pattern tends to stabilize between 2040 and 2050 before it further trends down. Figure 10(d) shows that the MP of AV usage has a similar (and strictly lower) trend to that of the MP of the vehicles with AV technologies installed.

The sensitivity analysis of simultaneously perturbating the three selected variables is shown in Figure 11. In addition to understand the general boundaries, these simultaneous perturbations reveal potential nonlinear interaction of these variables, which might not be additive and intuitive. Comparing them with their counterparts in Figure 8 to Figure 10, the major (50 and 75 percentile) portion of the output distributions for (a) tends to be dominated by the second test variable, "Base AV Technology Adoption Percentage Change." Although the overall trend of (b) tends to be more influenced by the first testing variable, "Base Adjustment Rate on Enforcement Resources," and the output range is more similar to that of the third testing variable, "Public and Media Bias Towards HV". The output boundary and the distribution of (c) is mainly determined by the first testing variable, while that of (d) is mainly influenced by the second testing variable.

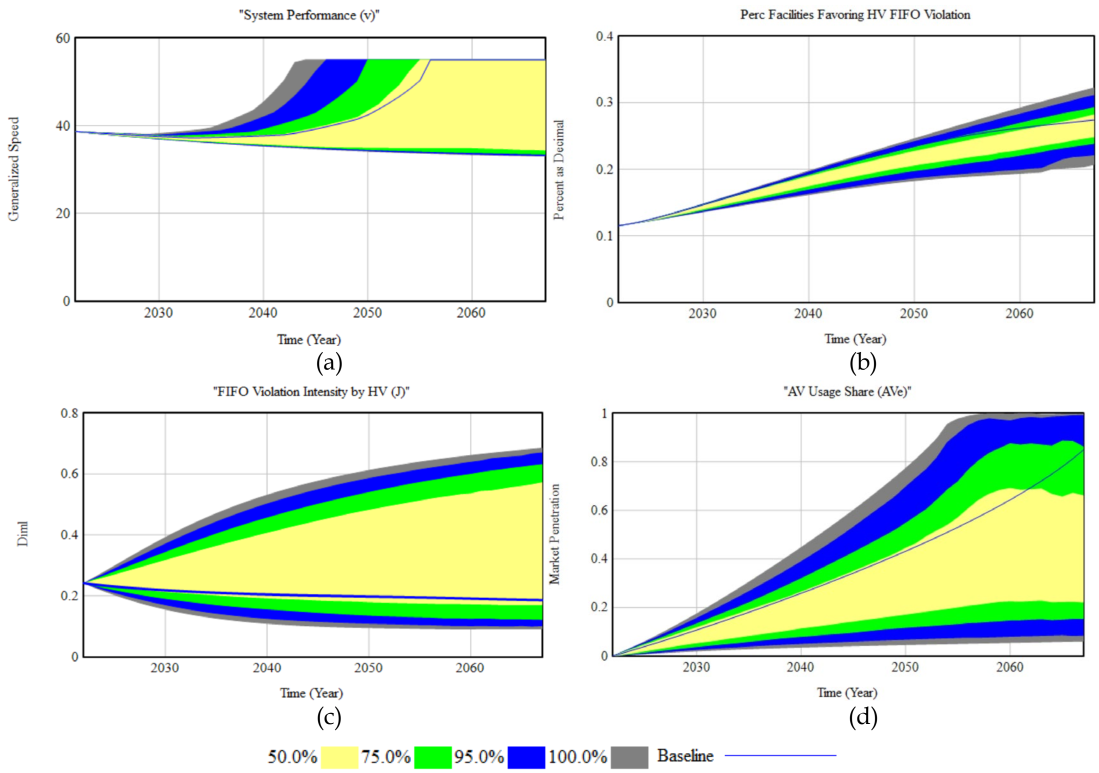


Figure 11 Sensitivity analysis on simultaneously perturbating the three selected variables.

5. Discussion

Results Interpretation. The case study supports the hypothesis that differentiating AV technology installation and usage could be relevant for network planning and policy decision-making in the age of autonomous and connected vehicles when the AV usage share has non-negligible impact on system performances. There is no evidence of a significant difference of the overall evolving pattern and trend of the AV usage compared with that of the installation studied by Bansal et al. (2017) and Nieuwenhuijsen et al. (2018). Nonetheless, there would be a pronounced systematic forecast bias that worsen with the increase of study time horizon. As shown in Figure 7, the improvement of system performance "plateaued" around the year 2047 in the "NoDiff" scenario where the planners do not differentiate the installation and the usage, while the "Baseline" scenario shows that such situation would not occur until the year 2054 or so. This bias is mostly caused by the self-induction of multiple positive feedback loops. The seemingly small deviation of the type of projects funded and scheduled due to the influence of driving behaviors encourages more instances to manually drive vehicles installed with AV technologies, which in term further increase the FIFO violation intensity. FIFO violation intensity might then influence public attitude and sentiment towards FIFO violation, which in term, further influence the type of project being scheduled and the adjustment on the effort of traffic law enforcement.

As expected, the share of the AV usage is equal or lower than that of the AV installation in any scenarios. When the share of the AV installation is near 0% or 100%, so is the share of the AV usage. Because of the characteristics of the three-dimensional macroscopic fundamental diagram (density-speed), it is observed that system performance experiences a diminishing improvement that eventually "plateaus." This is consistent with the claims by Yu (2018) as the AVs tend to serve as a "buffer" for the traffic oscillation and potential risky driving behaviors, and the marginal "buffer effect" is strong even with a small MP and gradually decreases to zero.

Because the increase MP of AV usage tends to lead to the higher critical density, not differentiating AV installation and AV usage tends to be overly optimistic about future system performance. Such overoptimism might lead to fewer network capacity improvement projects and hence lower network capacity in a long run. In the planning phase, since how the percent of facilities that tend to induce FIFO violation is influence heavily by FIFO violation intensity and public opinion, how the overall network-wide percentage of such facilities evolve depending on the initial percentage – low (high) initial percentage tend to have low (high) planning bias, and hence, low (high) percentage in the future.

From the sensitivity analysis, FIFO violation intensity might increase over time if the law enforcement is significantly insufficient. Indeed, if there is no concern of getting penalized at all, people might have the tendency to take advantage of the "kindness" or "politeness" of AVs frequently, which causes significant increase of FIFO violation and trend behaviors. The general sentiments of media report to the public about new AV technology advancement and AV-involved accidents) tend to only have a "local" impact without affecting the overall trend of the system performance, planning and construction of different project types (with delay), and the actual usage of AV technology.

Jointly looking at Figure 8-10, two feedback loops may seem to have considerable impacts on the trend of AV usage and could lead to biased forecasts if ignored. The first one is that FIFO violation might affect the adjustment of law enforcement effort both directly (e.g., perceived violation intensity and documented collision records) and indirectly (e.g., public and media sentiment) with delay. The second one is that AV technology equipment rate and AV usage influence the system performance in terms of average system speed, which, in turn, influence the planning and construction of facility amount and facility types. Different facility types have different impacts on FIFO violation tendency, which influences the difference between the installation and the usage. The latter feedback loop is longer in delay and harder to reverse. Both feedback loops contain policy leverages that allow policymakers to influence the trend of the AV usage and the corresponding system performance, though decisions have to be made under incomplete information since both loops contain exogenous variables with high uncertainty.

Public and media bias and their influences towards funding allocation and law enforcement intensity for FIFO violation also have a strong impact. But due to the delay in the planning phase and construction phase, their effects on FIFO violation might be hard for policymakers to incorporate in their mental models and their decisions. Clearly, more research is needed to study the sensitivity of different policy leverages to network topology, specific features, and area types to quantify this systematic bias with higher precision.

Somewhat surprising is the minor "overshoot" pattern observed in some of the parameter combinations (as shown in Figure 10). Specifically, it seems there might be a tendency to observe system performance initially when the AVs start to get implemented. However, if the facilities could not "catch up" and the newly proposed facilities tend to systematically incentivize manual-drive AVs, there might be a "drop" in system performance. Another somewhat surprising finding is that the project list tends to be longer in the "Baseline" scenario compared to the "NoDiff" scenario, which could be mainly due to decision-makers' overreaction to the insufficient system performance than expected. A longer project list also imposes risk of longer delays and higher risks in project implementation and construction. If the decision-makers keep assuming that the installation forecast is the same as the usage forecast, they will keep overestimating the benefits of the MP of the vehicles installed with AV technologies to the system performance, and, hence, keep falling short than expected and keep acting reactively. In the scenario where the type of projects added to the network might further incentivize HVs and manually driven AVs, the forecast gap between the AV installation and AV usage might grow even larger.

Adding to the complexity of these feedback loops is their variable delays (e.g., it sometimes takes multiple years for a programed project to be funded and eventually constructed) and the potential path-dependency (e.g., hard to moderate once the system tends to encourage FIFO violation). Since different regions have different planning procedures, driving cultures, law enforcement intensities, and pressure levels from the general public, the system dynamics may exhibit different patterns. Indeed, Figure 11 shows that the FIFO violation intensity might not necessarily go down with the increase of MP for vehicles installed with AV technologies.

Policy Implications. The policy implications from the SD framework and the case study have two folds. First, the manual driving behaviors and the improvement (or non-deterioration) of the system performance are not necessarily going in the opposite direction. As shown in Figure 7, when network stakeholders ignore the differentiation between the installation and the usage, the FIFO violation intensity tends to decrease especially when the traffic enforcement effort is insufficient. The network stakeholders, therefore, seem to face a trade-off between future system performance and FIFO violation. When the FIFO violation dominates, this situation might be worsened if the decision-makers do not (sufficiently) consider the planning, funding, and construction delays. Hence, decision-makers should plan in a way that allows the network capacity improvement dominates through an appropriate amount of "pessimism" about future system performance. On the other hand, the initial encouragement of AV without the sufficient infrastructure to "catch up" might cause increase (than what it would have been) or even minor "overshoot" of the manual driving behaviors.

Second, how the public and network stakeholders respond to the FIFO violation through the law enforcement effort and through the response to the public sentiment in the network planning and operation should be considered. Due to the positive feedbacks, the seemingly small influence can have a major impact on the overall evolution of the characteristics of the system (in terms of whether the system is more AV friendly or HV friendly) and the performance. This feedback can be further strengthened due to people's tendency to compare with others -- (relatively speaking) encouraging FIFO violation is essentially (relatively speaking) penalizing people who are strictly following the traffic rules and being kind to others. Such feedback might have a long-term effect in terms of the overall driving culture in a region. These two main implications boil down to whether decision-makers should react to the trend of AV technology installation trend for its own sake or encouraging people to use AV technologies to improve safety and system performance.

The two folds of implications suggest that a balanced mix of reactive and proactive approaches might be superior to purely proactive or reactive. In other words, overly emphasizing one of the types may be inappropriate. Certain policies that react to FIFO violation might be intuitively reasonable but might cause a more frequent violation. For a concrete example, lowering the speed limit as a response to intensive speeding might further cause the increased disutility of AV usage and the utility of manual driving behaviors. Due to the nonlinearity and interaction of various inputs, it is suggested to use system-wide key performance indicators (e.g., injuries and fatalities, cumulative travel time delay of the entire simulation scope, construction cost) to evaluate alternative policies and investment portfolios. The high uncertainty of the model parameters and exogenous variables further backs up the claim of the balance.

Delays are a key factor in both folds of implications. The first type of delay is mainly due to project planning, financing, and construction, which has been discussed thoroughly by Guevara et al. (2017) in the context of interstate highway systems and only requires minor generalization by differentiating the projects based on whether the projects tend to induce manual driving (and hence FIFO violation). Decision-makers may find it helpful to consider a "reverse reasoning" – if the positive sentiments towards manual driving increases, it might be because of more existing facilities that tend to induce manual driving, not because of such facilities being insufficient. At its least, the interpretation should be made on a case-by-case basis, not just go along with the public sentiment due to the upcoming performance review or election pressure of the decision-makers. Both types of delay might cause overreaction and waste of resources, which should be considered by policymakers who are under pressure for "delivering the results."

The lack of observations is the biggest challenge for proposing an SD model of this topic. At the moment this paper was produced, high-level AV technologies has not officially entered the mass travel market, except some assistant technologies such as crash avoidance warning, lane warning, ABS, etc. Adding more challenges is the possibility to make HV illegal or impossible (e.g., Level 5 automation without human control options), and the potential technologies can detect whether a vehicle has AV technologies equipped and whether the vehicles are driven by AV technologies or manually. This paper assumes that the public agencies and auto manufacturers cannot force people to use AV technologies (e.g., Level 5 automation), but a future extension is needed if these scenarios become increasingly likely. Some exogeneous variables might also need to be incorporated into the model as endogenous for broader policy strategies. As shown in Figure 9, the high sensitivity of "Base AV Technology Adoption Percentage Change" has a high impact on the system performance over time, implying a strong impact of the specific policy that influences the base adoption rate (especially in the later phase of a simulation). For instance, whether public agencies adopt an effective rebate program or a set of mandatory adoption regulations about the installation of AV technologies might have a strong impact on the trend of the system performance and the AV usage share. However, the feasibility of such a rebate program partially depends on how the cash flow and the fund balance (as a stock variable) evolves, which, therefore, should be also incorporated in the model.

One may also question whether policymakers might directly influence the structure of the CLD. Is it possible for public agencies to influence how media reporting the crashes to help people think more rationally when there is a report of an AV-involved collision? Should it be required that every time an AV-related crashes occur, the media report whether the AV was actually driven by a human and make sure to present the latest statistics about how risky such crashes occur? Is it possible that new technologies might emerge and create additional incentive to just let the vehicles drive themselves? Many might agree that emerging smartphone technologies make the spare time and increase the relative utility of using AV technology usage. With the development of Virtual Reality (VR) and more convenient entertainment and working systems, the relative utility of pursuing other activities than driving manually might further increase. These questions bring the necessity of more extensive studies of broader scenario specifications.

It should be noted that public agencies of different countries and regions have different decision styles. For example, a jurisdiction might tend to directly respond to the public's sentiment surge due to biased media reporting without considering the long-term negative impact because, say, the "mess" will be taken care of by the next mayor. The intensity of the FIFO violation issue might also be different among regions with different driving cultures. For example, some regions might have a more friendly and fair-minded driving culture and AV-friendly facility design, so HV and FIFO violations might be perceived less attractive and less common.

6. Conclusions

The paper starts with no presumption that AV installation and AV usage is negligible in public policy and investment decision making. We consider various feedback loops under a framework of SD to investigate the potential impact of differentiating AV technology installation and AV technology usage in forecasting future system performance and, in turn, forecasting AV usage trend itself. In the proposed model, AV users may choose to manually drive their vehicles for bypassing the front vehicle, driving above the speed limit, speeding up when a signal light turns yellow, making a left turn at solid double yellow lines, illegal parking, etc. These behaviors along with insufficient AV-friendly infrastructure might further induce other travelers to do so.

The SD-based approach facilitates the study of various complex feedbacks that that relevant to such a differentiation. We identify the factors, such as public sentiment, facility composition, traffic law enforcement, network planning decisions, and construction delay that may influence the FIFO violation intensity. Such violations might have complex interactions with the MP of AV usage and, in the long run, affect the public and media sentiment, which, in turn, influences the intensity of traffic law enforcement and the types of network projects to be programed and implemented. Additionally, some exogenous variables such as the market entrance pattern of the vehicles equipped with AV technologies can become endogenous by incorporating existing models that are particularly targeted at modeling these variables. The framework can also serve as a qualitative mental model to help transportation planners and traffic operators to make more productive and constructive discussion and realize implicit assumptions.

A case study is used to demonstrate the SD-based model and provides preliminary analysis and potential policy implications. The sensitivity analysis shows that a mix of reactive and proactive approaches might be superior to purely proactive or reactive due to the high uncertainty of the exogenous variables in different model components and the potential feedback loops not captured by the model. In other words, overly emphasizing one of the types may be inappropriate. It is also important to realize that certain policies that seem reasonable might worsen the situation. Adding more complexity is the high uncertainty of the variable. Therefore, one major policy implication is a mix of policies that balance reactive and proactive strategies.

Other variables and feedbacks that this paper does not consider are worth further investigation in future studies. The model does not explicitly consider the impact of system performance on/from the change of trip length, trip purpose, occupancy (including shared ride), departure time, destination choice, business innovation, land-use change, and climate. It might be interesting to extend the model to consider user heterogeneity on characteristics such as fatigue levels, DUI, perceptions about the severity of a traffic accident, cognitive capacities, personalities, and age pyramid (e.g., younger people might drive in a riskier manner and tend to override the AV mode more frequently). Equity implications might emerge if early AV adopters tend to be high in income – since people tend to unintentionally violate traffic rules more in MVs, low-income populations might, on average, bear more share of injuries and traffic violation penalty that would further exacerbate their financial situations. The total fleet trend might have interactions with the MP of the AV technology installation. Embedding the proposed model into a broader model that considers other modes of transport, land use evolvement, and socioeconomic activities can also be considered. However, even without considering these factors, it seems clear that the difference between the vehicles equipped with AV technologies and the vehicles driven automatically on roads should not be instantaneously assumed ignorable in planning and policy decision-making.

Acknowledgements

The work described in this paper was partly supported by the Research Grants Council of the Hong Kong Special Administrative Region (Project No. 15212217), the Research Institute for Sustainable Urban Development at the Hong Kong Polytechnic University (1-BBWF), and the National Natural Science Foundation of China (72071174). Their support is gratefully acknowledged. We also appreciate the valuable comments and suggestions from the anonymous reviewers.

References

Abbas, K. A., Bell, M. G., 1994. System dynamics applicability to transportation modeling. *Transportation Research Part A: Policy and Practice 28*(5), 373-390.

Agarwal, A., Lämmel, G., 2016. Modeling seepage behavior of smaller vehicles in mixed traffic conditions using an agent-based simulation. *Transportation in Developing Economies 2*(2), 8.

Bansal, P., Kockelman, K. M., 2017. Forecasting Americans' long-term adoption of connected and autonomous vehicle technologies. *Transportation Research Part A: Policy and Practice 95*, 49-63.

Boyce, D., Lee, D. H., Ran, B., 2001. Analytical models of the dynamic traffic assignment problem. *Networks and Spatial Economics*, *1*(3), 377-390.

Carey, M., 1992. Nonconvexity of the dynamic traffic assignment problem. *Transportation Research Part B: Methodological 26*(2), 127-133.

DC StreetSafe: Automated Traffic Enforcement, 2020, Accessed on 21 August 2020 at https://ddot.dc.gov/page/dc-streetsafe-automated-traffic-enforcement

de Waard, D., Rooijers, T., 1994. An experimental study to evaluate the effectiveness of different methods and intensities of law enforcement on driving speed on motorways. *Accident Analysis & Prevention 26*(6), 751-765.

Deloitte., 2020. *Autonomous Driving: Moonshot Project with Quantum Leap from Hardware to Software & AI Focus.* *https://www2.deloitte.com/content/dam/Deloitte/de/Documents/consumer-industrial-products/POV_Autonomous-Driving_Deloitte.pdf*

Department of Motor Vehicles, 2019, Assessed on 21 August 2020 athttps://dmv.dc.gov/sites/default/files/dc/sites/dmv/page_content/attachments/Historical%20Registration%20Data%20FY2010-FY2019.pdf

Fakhrmoosavi, F., Saedi, R., Zockaie, A., Talebpour, A., 2020. Impacts of connected and autonomous vehicles on traffic flow with heterogeneous drivers spatially distributed over large-scale networks. *Transportation Research Record*. 0361198120940997.

Forbes, 2016. Drivers Are Tuning Out, Turning Off An 'Annoying' High-Tech Auto Safety Feature, Jim Gorzelany. https://www.forbes.com/sites/jimgorzelany/2016/04/13/drivers-are-tuning-out-turning-off-an-annoying-auto-safety-feature/?sh=3fdc71462069

Friedman, S., 2006. Is counter ‐ productive policy creating serious consequences? The case of highway maintenance. *System Dynamics Review: The Journal of the System Dynamics Society 22*(4), 371-394.

Geroliminis, N., Daganzo, C. F., 2008. Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transportation Research Part B: Methodological 42*(9), 759-770

Godfrey, J. W., 1969. The mechanism of a road network. *Traffic Engineering & Control 8*(8).

Guevara, J., Garvin, M. J., Ghaffarzadegan, N., 2017. Capability trap of the US highway system: Policy and management implications. *Journal of Management in Engineering 33*(4), 04017004.

Gruel, W., Stanford, J. M., 2016. Assessing the long-term effects of autonomous vehicles: a speculative approach. *Transportation research procedia 13*, 18-29.

Hsieh, I. Y. L., Kishimoto, P. N., Green, W.H., 2018. Incorporating multiple uncertainties into projections of Chinese private car sales and stock. *Transportation Research Record 2672*(47), 182-193.

Hülsmann, M., Borscheid, D., Friedrich, C. M., Reith, D., 2012. General sales forecast models for automobile markets and their analysis. *Trans. MLDM 5*(2), 65-86.

Jin, W. L., Jayakrishnan, R., 2005. First in, first out properties of a commodity-based kinematic wave simulation model. *Transportation research record 1934*(1), 197-207.

Jin, W. L., 2007. A dynamical system model of the traffic assignment problem. *Transportation Research Part B: Methodological 41*(1), 32-48.

Kim, M., Park, J., Oh, J., Lee, W., Chung, D., 2019. Identifying and prioritizing the benefits and concerns of CAV: A comparison of individual and expert perceptions. *Research in Transportation Business and Management* 32, 100438.

Kwon, T. H., 2012. Strategic niche management of alternative fuel vehicles: A system dynamics model of the policy effect. *Technological Forecasting and Social Change 79*(9), 1672-1680.

Lavieri, P. S., Garikapati, V. M., Bhat, C. R., Pendyala, R. M., Astroza, S., Dias, F. F., 2017. Modeling individual preferences for ownership and sharing of autonomous vehicle technologies. *Transportation research record 2665*(1), 1-10.

Litman, T., 2017. *Autonomous Vehicle Implementation Predictions*. Victoria Transport Policy Institute, Victoria, Canada.
Macmillan, A., Connor, J., Witten, K., Kearns, R., Rees, D., Woodward, A., 2014. The societal costs and benefits of commuter bicycling: simulating the effects of specific policies using system dynamics modeling. *Environmental Health Perspectives 122*(4), 335-344.
Macmillan, A., Woodcock, J., 2017. Understanding bicycling in cities using system dynamics modelling. *Journal of Transport & Health 7*, 269-279.
Mehdizadeh, M., Shariat-Mohaymany, A., 2020. Who are more likely to break the rule of congestion charging? Evidence from an active scheme with no referendum voting. *Transportation Research Part A: Policy and Practice 135*, 63-79.
Mesa-Arango, R., Ukkusuri, S. V., 2014. Modeling the car-truck interaction in a system-optimal dynamic traffic assignment model. *Journal of Intelligent Transportation Systems 18*(4), 327-338.
Metropolitan Police Department, 2019. Traffic Fatality. Accessed on 21 August 2020 at https://mpdc.dc.gov/node/208742
Milkovits, M., Copperman, R., Newman, J., Lemp, J., Rossi, T., Sun, S., 2019. Exploratory modeling and analysis for transportation: An approach and support tool-TMIP-EMAT. *Transportation Research Record 2673*(9), 407-418.
National Highway Traffic Safety Administration, 2020. Accessed on 21 August 2020 at https://one.nhtsa.gov/Research/Light-Vehicle-Antilock-Brake-Systems-(ABS)-Research-Program
Nieuwenhuijsen, J., de Almeida Correia, G. H., Milakis, D., van Arem, B., van Daalen, E., 2018. Towards a quantitative method to analyze the long-term innovation diffusion of automated vehicles technology using system dynamics. *Transportation Research Part C: Emerging Technologies 86*, 300-327.
Office of the Chief Financial Officer, 2020. Budget and full-time equivalent. Accessed on 21 August 2020 at https://cfo.dc.gov/sites/default/files/dc/sites/ocfo/publication/attachments/fa_mpd_chapter_2019m.pdf
Pfaffenbichler, P., 2011. Modelling with systems dynamics as a method to bridge the gap between politics, planning and science? Lessons learnt from the development of the land use and transport model MARS. *Transport Reviews 31*(2), 267-289.
Raj, A., Kumar, J. A., Bansal, P., 2020. A multicriteria decision making approach to study barriers to the adoption of autonomous vehicles. *Transportation research Part A: Policy and Practice 133*, 122-137.
Richardson, G. P., 2011. Reflections on the foundations of system dynamics. *System Dynamics Review 27*(3), 219-243.
Ryu, S., Chen, A., Choi, K., 2016. Solving the stochastic multi-class traffic assignment problem with asymmetric interactions, route overlapping, and vehicle restrictions. *Journal of Advanced Transportation 50*(2), 255-270.
Shepherd, S., Bonsall, P., Harrison, G., 2012. Factors affecting future demand for electric vehicles: A model based study. *Transport Policy 20*, 62-74.
Shepherd, S. P., 2014. A review of system dynamics models applied in transportation. *Transportmetrica B: Transport Dynamics 2*(2), 83-105.
Struben, J., Sterman, J. D., 2008. Transition challenges for alternative fuel vehicle and transportation systems. *Environment and Planning B: Planning and Design 35*(6), 1070-1097.
Tay, R., 2009. The effectiveness of automated and manned traffic enforcement. *International Journal of Sustainable Transportation 3*(3), 178-186.
The Metropolitan Washington Council of Government, 2020. *FY 2021-2024 Transportation Improvement Program*. Mar. 18, 2020.
Vickery, W. S., 1991. Congestion in midtown Manhattan in relation to marginal cost pricing. Technical Report. Columbia University
Wang, G., Kockelman, K. M., 2018. Three methods for anticipating and understanding uncertainty of outputs from transportation and land use models. *Transportation Research Record 2672*(49), 12-22.
Wu, T., Zhao, H., Ou, X., 2014. Vehicle ownership analysis based on GDP per capita in China: 1963–2050. *Sustainability 6*(8), 4877-4899.
Yu, J., 2018. *Paradigms of Identifying and Quantifying Uncertainty and Information in Constructing a Cognition-Modeling Framework of Human-Machine Transportation Systems*. Doctoral dissertation, University of California, Irvine, USA.
Zhao, X., Wang, Z., Xu, Z., Wang, Y., Li, X., Qu, X., 2020. Field experiments on longitudinal characteristics of human driver behavior following an autonomous vehicle. *Transportation Research Part C: Emerging Technologies 114*, 205-224.
Zhong, Z., Lee, E.E., Nejad, M., Lee, J., 2020. Influence of CAV clustering strategies on mixed traffic flow characteristics: An analysis of vehicle trajectory data. *Transportation Research Part C: Emerging Technologies 115*, 102611.