The Journal of Transport and Land Use jtlu.org
Vol. 17 No. 1 [2024] pp. 321–349

The value of scenario discovery in land-use modeling: An automated vehicle test case

Daniel Engelberg
Northeastern University
d.engelberg@northeastern.edu

Abstract: Long-range planning is an uncertain endeavor. This is especially true for urban regions, small ships in a global urban storm that are too small to influence macro policies and without the land-use powers of local governments. Exploratory scenarios, the established practice for planning under deep uncertainty, have inspired stakeholders to consider multiple futures but have fallen short of identifying robust and contingent policies. We need new tools to plan under conditions of deep uncertainty. Scenario discovery is a technique for using simulation models to explore the performance of policy options across uncertain scenarios. This paper presents an application of scenario discovery in land-use modeling and asks what this computationally intensive approach offers relative to a more circumscribed exploration of uncertainty space. The introduction of autonomous vehicles (AVs) and their associated impacts on land use provide a test case demonstrating this method, as well as a topic of substantive concern. This research concludes that scenario discovery is particularly valuable for identifying the conditions under which contingent policies are likely to succeed. In terms of AV policy, this research establishes that forward-thinking, transit-oriented-development strategies can mitigate spatial dispersion while also reducing overall housing costs. In addition, I find that AVs may blunt the impacts of some current policy tools if they extend the distance individuals are willing to travel to work.

Article history:
Received: June 13, 2023
Received in revised form: February 27, 2024
Accepted: March 11, 2024
Available online: May 9, 2024

1 Introduction

The analytic processes deployed in long-range planning have long aided in masking uncertainty. Land Use Transportation Interaction (LUTI) models, a tool utilized by Metropolitan Planning Organizations (MPOs) according to the predict and prepare paradigm that has dominated long-range planning approaches. Modelers and planners often rely on point forecasts whose basic assumption is that the future will be largely similar to the present—just with more people (Marsden & McDonald, 2019). Planning professionals are just beginning to develop analytic frameworks that acknowledge future uncertainty in the application of these models. I thus examine scenario discovery and robust decision-making, a technique for using simulation models to explore the performance of policy options across uncertain scenarios (Lempert et al., 2006). Scenario

https://doi.org/10.5198/jtlu.2024.2401

The *Journal of Transport and Land Use* is the official journal of the World Society for Transport and Land Use (WSTLUR) and is published and sponsored by the University of Minnesota Center for Transportation Studies.

discovery has limited deployment in the regional planning context, including a handful of transportation applications (Lempert et al., 2020; Milkovits et al., 2019), and even fewer in Land Use Transportation Integrated (LUTI) modeling (Swartz & Zegras, 2013). At a larger scale scenario discovery has been applied as part of an integrated assessment model that included a cellular automata model land use/land cover change model, a different approach to simulating land-use change than explored in this paper (Jafino et al., 2021; Jafino & Kwakkel, 2021)

The primary goal of this paper is to demonstrate the application of scenario discovery for land-use modeling to the urban planning community. In doing so, I assess the usefulness of this approach relative to more circumscribed approaches to incorporating uncertainty into modeling. I compare the information gained for a large number of futures to a more limited sampling of uncertainty space. In focusing on only the outer edges of the uncertainty space, the more limited sampling is designed to resemble the results of an exploratory scenario exercise which develops and examines scenarios at the edges of plausibility. I demonstrate scenario discovery by modeling the impacts of automated vehicles (AV) on land-use decisions. In doing so, I demonstrate how scenario discovery can allow us to draw policy insights regarding a deeply uncertain topic. I also draw conclusions regarding the potential land-use impacts of automated vehicles and policies to support desirable outcomes.

My research seeks to provide insight on the following questions:

- Q1: How well does scenario discovery within LUTI modeling perform in identifying robust and contingent planning strategies relative to exploratory scenarios?
- Q2: What are the potential land-use outcomes from autonomous vehicle adoption and what policies can be put in place to support desirable outcomes?

For the research questions, I hypothesize as follows:

- H1A: Scenario discovery within LUTI modeling will be more *precise* in identifying robust strategies than exploratory scenarios because additional scenarios will support a finer-grain measure of robustness. Scenario discovery within LUTI modeling will be more *accurate* in identifying robust strategies because the sampling technique ensures more even coverage of the uncertainty space.
- H1B: Scenario discovery will determine contingent policies, where more limited samples will be insufficient.
- H2A: Automated vehicles will contribute to more dispersed land-use patterns, as was the case with previous transportation technologies that increased mobility speed and comfort (Wegener & Fuerst, 2004). The dispersal encouraging effects of decreased value of accessibility in residential location choice will be stronger than the concentration effect of opening up newly developable central land on former parking lots.
- H2B: Policies that encourage more concentrated land use will be less impactful if automated vehicles change how people value accessibility in location choice.

In answering these questions, this work seeks to contribute to the literatures on the application of uncertainty analysis in planning, LUTI applications, and the impacts of new transportation technologies.

2 Uncertainty in planning

A primary tool for planning under uncertain conditions is "exploratory scenarios," which are defined by asking What can Happen (Börjeson et al., 2006)? In this type of scenario planning, conveners work with stakeholders to tell several dissimilar, plausible stories about the future to prepare for whatever comes (Schwartz, 1991). This form of scenario adapted approaches that Herman Kahn developed for Cold War strategy to the business world (Wack, 1985), but has rapidly gained acceptance within urban planning in the past two decades (Avin & Dembner, 2000; Chakraborty & McMillan, 2015; Zegras et al., 2004).

Deep uncertainty can be defined as a condition under which individuals know the potential outcomes but cannot define the distribution of key parameters (Kwakkel et al., 2010). Exploratory scenarios have been preferred for deep uncertainty because of their focus on the conditions under which policies should be advanced, rather than determining the most likely or preferable outcome. One objective of such exercises is to select policies that are robust to a variety of futures and identify other, contingent, policies that should be implemented in limited circumstances (Avin, 2007). Depending on the goals and models available to planners, scenarios may also be run through urban systems models to understand potential impacts (Knaap et al., 2020). For narrative intelligibility, scenario planners recommend 3–5 scenarios. However, from the modeling standpoint, this is insufficient to consider policy robustness (Lempert et al., 2006). This does not necessarily obviate the organizational learning and collaborative action potential of scenarios (Bootz, 2010; Wack, 1985; Xiang & Clarke, 2003), though those outcomes have been tested elsewhere (Zegras & Rayle, 2012).

Robust decision-making using scenario discovery could provide an analytics approach for considering deep uncertainties within LUTI models. In this approach, policies are modeled in an ensemble of hundreds or thousands of futures. Instead of asking which policies are likely to produce the highest expected value, scenario discovery asks what conditions under which policies perform well or poorly (Hall et al., 2012; Walker et al., 2013). McPhail et al. (2020) investigate the impact of different sampling strategies on robustness and the rank order of policy choices using the stylized Lake problem, providing the first systematic examination of their relative performance. Examining both small and large samples using diverse, targeted, and uniform sampling strategies, the authors find that measures of robustness vary significantly across samples, but rank order often remains similar. My research builds on this study by examining how an ensemble of land-use modeling runs relative to more limited exploratory scenario approaches. This paper then seeks to determine if their results remain valid for models that are significant for urban planning rather than water resources management.

3 Modeling automated vehicles and land use

As AVs approach the marketplace, there is great uncertainty regarding their impacts on household location choice and associated land-use patterns. AVs may exacerbate sprawl by making longer commutes more comfortable or facilitate infill by making near-to-destination parking obsolete. Though researchers have extensively modeled the travel demand impacts of AVs, few studies have utilized LUTI models or estimated second-order impacts on land use (Papa & Ferreira, 2018; Soteropoulos et al., 2019), even though previous changes in transportation technology profoundly altered large-scale urban form (Wegener & Fuerst, 2004).

Meyer et al. (2017) was one of the first studies to model the land-use impact of AVs. Using the Swiss national transport model, they found accessibility declines in urban areas associated with increased congestion but accessibility gains in suburban areas. The Swiss national transport model is a macroscopic travel demand model. This study used only the personal transport changes and no changes to freight. Because they only ran a transport model, the findings are gravity-based accessibility scores based on travel times on the network. They do not calculate resultant changes in land use.

Two other studies confirm that AVs could encourage population dispersal; finding that the inner urban population decreased between 1–4% while outer suburbs in nonurban and rural regions increased between 1–3% (Gelauff et al., 2017; Thakur et al., 2016). Gelauff et al. (2017) use the Dutch spatial equilibrium model (LUCA). LUCA microscopically models four types of agents: three different educational attainment consumer groups and land owners. The consumers choose the location and size of their dwelling, their job location, and commute mode by considering locational characteristics and commuting costs. The simulation experiments consider the impacts of AVs on lower perceived cost of travel and additional roadway capacity. Thakur et al. (2016) use a bespoke LUTI model for the Melbourne area. This model has thirty-one radiating zones and is integrated with the Victoria Integrated Transportation Model. The population is redistributed according to a discrete choice model in which accessibility to employment is a key variable in location choice. The scenarios they consider examine changes to real and perceived in-vehicle and out-of-vehicle time.

Though the aforementioned modeling indicated that AVs would encourage dispersal, a closer look into model results tells a more complex story. Distance to work could increase between 7–10% in Atlanta though retired households may move in closer (Zhang, 2017). The author develops an AV operations and dispatching model that integrates with UrbanSim's discrete choice residential and firm location model. The simulation experiments primarily consider behavioral adjustments associated with decreased disamenity of in-vehicle time and policies related to parking. Development may leapfrog the greenbelt in Seoul, South Korea, and become less clustered (Kim et al., 2015). Kim also uses an agent-based discrete choice framework. The authors only present one automated vehicle scenario. In this scenario, they assume increased accessibility of distant regions and decreased preferences for proximity to goods and amenities. Several authors noted the increasing importance of amenities in location choice (Meyer et al., 2017; Thakur et al., 2016). Specific results are difficult to compare because of the different contexts, models, and assumptions in each simulation.

Looking into more neighborhood-specific applications of AVs, Basu and Ferreira (2020) utilized the SimMobility long-term model to examine the deployment of automated mobility in association with a car-lite pilot in Singapore. SimMobility is a state-of-the-art transportation model with three modular components: long-term land-use decisions, mid-term travel demand, and short-term network simulation. The long-term model simulates daily behaviors in the housing market, including the decision to search for housing, bidding, and developer behavior. Utilizing accessibility and property values as variables, they determine that car-lite policies, in conjunction with automated vehicles, have the potential to increase the incomes of households moving into the study area. In a different study, SimMobility is also used to simulate the impacts of automated mobility on demand (AMoD) on vehicle ownership and residential choices in Singapore. They examine a partial automation scenario, in which AMoD is introduced into only a specific study area of central Singapore, and a full automation scenario, in which only AMoD and public transit are allowed to operate, while private cars are banned from the study area. In their full automation scenario, the already high-demand study area has increased demand relative to the baseline.

The SILO model, the microsimulation model used in this research, is used in two previous studies. Looking at two scenarios which decrease the value of time and increase vehicle occupancy in Austin, AVs decrease core population between 5.3% and increase growth outside the core by 5.6% (Wellik & Kockelman, 2020). In Munich, six scenarios examine reduced value of time, the decision to purchase an AV, and lower parking penalty in the core. The additional urban sprawl induced by less burdensome commuting is largely compensated by the increased attractiveness of the already popular urban core (Llorca et al., 2022).

Finally, the TRANSPACE model is used to examine the impact of roadway capacity and induced demand in the Bay of Santander (Cantabria, Spain). If AVs create new capacity without inducing demand, population growth increases by 2.1% outside the city, but if the capacity is consumed, growth could increase up to .7% in the central zone. Employment grew 1.6% in the core with increased core capacity, but decreased -.85% with the assumed behavioral change (Cordera et al., 2021)

Generally, these simulation experiments have estimated decentralizing behavior to have larger impacts on land-use outcomes than the reallocation of central land, however, that balance is not universal. Comparison across the current literature is difficult because the cases lack consistency in the selected models and variables. Even within individual experiments, the number of runs remains small and it's difficult to determine whether their results are truly robust beyond their specific parameterization, except for some sensitivity testing. None of these modeling efforts examined more than six scenarios or systematically explored the uncertainty space—something that this paper seeks to introduce.

4 Exploratory modeling in land use and transportation simulation

Exploratory modeling is a computational approach with a variety of techniques to assist reasoning regarding a system when there is uncertainty. When modelers cannot take system dynamics for granted because of these uncertainties, exploratory modeling approaches perform hundreds, thousands, or even more runs to rapidly test how those uncertainties impact model dynamics (Bankes, 1993). In the past two decades, exploratory modeling approaches have increased in variety and application, as led by decision-making under deep uncertainty scholars in Europe (Kwakkel et al., 2016) and at RAND, a private research organization long associated with strategic analysis (Groves & Lempert, 2007). This paper utilizes one of those approaches, scenario discovery for robust decision-making (Bryant & Lempert, 2010; Lempert et al., 2006).

Scenario discovery has found very limited application in LUTI modeling. Lempert et al. (2020) demonstrate robust decision-making using scenario discovery in travel demand modeling. Specifically, the authors work with the Sacramento Area Council of Government to determine the conditions under which their transportation plan simultaneously meets greenhouse gas emissions, local emissions, mobility, and equity goals. Using the PRIM algorithm, they determined that gas prices, fuel efficiency, employment growth, and vehicle miles traveled elasticity with respect to time were key drivers for meeting all the goals in 12% of the scenarios tested. They further determined that encouraging more rapid penetration of zero-emission vehicles would reduce the vulnerability of failing to meet one of their goals.

Milkovits et al. (2019) developed the Travel Model Improvement Program Exploratory Modeling and Analysis Tool (TMIP-EMAT), a travel demand model-oriented extension of the original Exploratory Modeling and Workbench (Kwakkel, 2017), and deployed TMIP-EMAT using the Greater Buffalo-Niagara Regional Transportation Council regional travel demand. Concerning land use, Swartz and Zegras (2013) provided a demonstration of concept of land-use modeling using UrbanSim to examine future growth in Lisbon, Portugal. This paper is just the second instance, to my knowledge, of exploratory modeling of land-use outcomes, and the first to compare scenario discovery results to more limited sampling approaches.

Two papers on water climate adaptation in the Mekong River Delta apply scenario discovery to an integrated assessment model that includes a cellular automata land use/land cover change component (Jafino et al., 2021; Jafino & Kwakkel, 2021). Cellular automata models of land use/land cover change simulate land cover changes on a raster according to locational and proximity characteristics (White et al., 2015). These models can incorporate the behavior of land owners, as is the case in the Mekong River Delta study (van Delden et al., 2011). Agent-based micro-simulation, presented in my research, differs in capturing the individual decisions of agents regarding where to move. This class of model also remains the standard within urban planning for ease of integration with existing transportation models and for realistic simulation of housing markets.

5 Research design and methods

Scenario discovery is a simulation research technique with two phases: sampling and data mining (Figure 1 provides details as described below). Given that an exploration of all future states is impossible, Latin hypercube sampling (LHS)—which ensures the maximum difference between runs—is preferable to intuition-based approaches, which might ignore regions of uncertainty (Groves & Lempert, 2007). LHS divides the sample space into proportional segments and draws one sample from each segment, thus ensuring that no region of the uncertainty space is insufficiently sampled as may occur in random sampling (Mckay et al., 2000). Scenario discovery then utilizes data mining to explore the broader uncertainty space for regions in which a policy performs particularly well or poorly. I employ the Patient Rule Induction Method (PRIM), an algorithm that searches for lower dimensional boxes of concentration within higher dimensional space (Friedman & Fisher, 1999). Because each box edge is defined by a single variable, PRIM is easier to interpret than comparable methodologies (Lempert & Groves, 2010). For my analysis, I utilize the *prim* library in R. Please see the appendix for more on how PRIM identifies boxes.

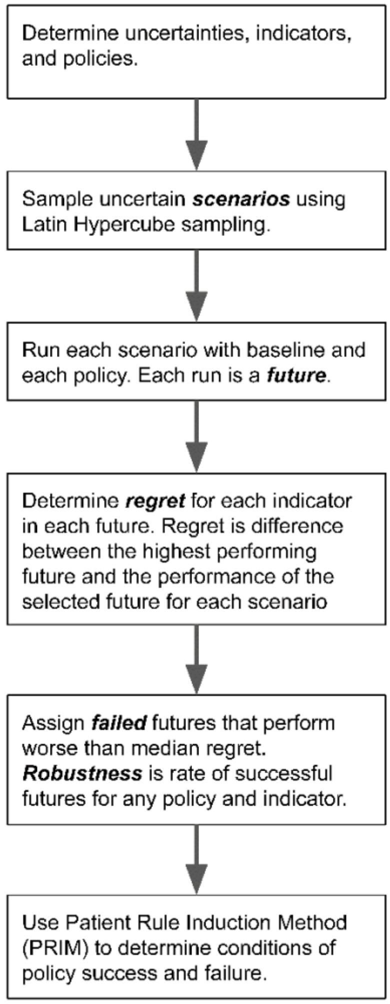

Figure 1. The scenario discovery process utilized for analysis of AV land-use futures

As all PRIM boxes are not equally informative. I selected to further investigate those boxes that performed above the median for density—the percentage of boxed futures that are failed futures - and coverage—the percentage of all failed futures covered by the box. I also excluded any boxes that were > 95% failures or successes because they provided insufficient information for determining clear regions of success and failure.

Each uncertainty parameterization is a scenario.1 Each scenario is modeled with baseline parameters and with each policy intervention. For each completed run, or future, the simulation model generates select indicators. The indicators of each future are translated into regret—i.e., the difference between the highest-performing future for each

[1] My use of the term "scenario" is in line with (Kwakkel et al., 2013) who uses scenario to describe the individual computational experiments. Other articles using similar approaches have used "scenario" to refer to the set of circumstances discovered through a PRIM box (Bryant & Lempert, 2010). I prefer the former definition as I find that it intuitively matches how scenarios are defined in the exploratory scenario as circumstances defined by unknown external forces. The latter definition conflates uncertainties and policy levers.

scenario and the performance of the selected future. Regret allows for comparison across unlike scenarios. A failed future refers to one with greater-than-median regret amongst policies with regret greater than 0.2 Let robustness be the percentage of scenarios that succeed for any given policy. For any indicator, robustness is calculated as $r = 1 - \frac{N_f}{N}$ where r is the robustness, where N is the total number of futures, and N_f is the number of failed futures. We can also call $\frac{N_f}{N}$ the failure rate. Finally, the PRIM algorithm is deployed to search for the conditions under which policies tend to outperform others (Gross, 2018).

A single PRIM run determines the conditions of policy success for one policy as measured against one indicator. Comparing multiple policies across several indicators requires generating PRIM boxes for each "policy/indicator" combination. PRIM iteratively separates regions of high regret; it generates a series of multi-dimensional boxes. Each of these solutions sits upon the Pareto optimum, trading off coverage—the rate at which the box captures the failed futures—for density—the proportion of futures in the box that meet failure criteria. From each PRIM solution set, I selected the box with the highest density of failure to maintain comparability between the 35 boxes for each combination of five policies and seven indicators that I introduce later in this section. These boxes also identify the more uncertainty conditions under which a policy might fail relative to boxes with greater coverage.

The experiment is run using the Simple Integrated Land Use Orchestrator (SILO), an agent-based land-use model microsimulation that stochastically simulates household location choice decisions, designed to integrate with existing transportation models (Moeckel, 2016). SILO utilizes discrete choice models to simulate each household, person, and dwelling unit in a modeling region. See the appendix for more details on the model weights and references to model details. The model is designed primarily to determine location choice decisions and associated land-use patterns and can be readily integrated with existing transportation models. Relying in part on behavioral heuristics, such as maintaining a relatively fixed distance to work distribution, it is simpler to set up and calibrate than other microscopic land-use models.

SILO is used to simulate the 2015–2030 timeframe in one-year increments. The model runs begin in the year 2015 because that is the year for which the model was initially built and validated. The observed effects of automated vehicles should not be much impacted by the start year.3 The simulation does not include the impacts of the Covid-19 pandemic on the housing market. The model consists of four modules: synthetic population generation, demographic changes, real estate development, and household relocation. Household relocation is determined by three logit models that determine whether to move or stay, which region to move into, and which dwelling to move into within that region. Location choice factors include accessibility to jobs, travel time to work for working household members, and housing costs. In the calibration of the application of SILO to Maryland and reflecting decision-making of actual households in the region, the modelers added racial segregation preferences, as well as measures of crime and school quality (Knaap et al., 2020). The weight of individual factors in the logit model depends on household size, income, and race.

[2] Other rules can be used for determining regret including policy responsive tools where such a threshold officially exists.

[3] This simulation did not include the Covid-19 or the more recent run up in inflation. This experiment was designed to isolate policies designed for managing the land-use impacts of AV uncertainty. However, future research could explore regional uncertainties more comprehensively.

SILO maintains a synthetic population of all individuals and dwellings as well as households. The demographic model is essential in determining realistic location choice behaviors. In every annual simulation each individual ages one year. Markov transition rates determine life events—such as marriage, parenthood, and death—for individuals of a given gender and age. This is crucial because the residential location choice model accounts for the home and work locations of both workers in married households and school quality for households with children. The development model increases home values where units are highly occupied and decreases home values where many units remain vacant. Developers respond to these signals by preferring to add units where prices are high. A development capacity layer acts as a hard cap on total units. This layer was developed from an analysis of zoned capacity in each Maryland jurisdiction and estimates based on projected growth locations in other states.

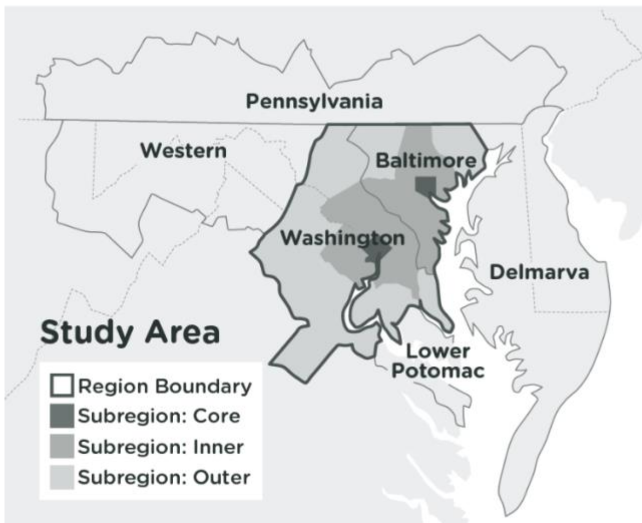

Figure 2. The SILO Modeling Region. Top map shows regional designation used for interpreting SILO results

A couple of the limitations of this experiment accompany my selection of the SILO model. First, the SILO model takes the future employment distribution as a given. Our instantiation used employment projections from the Maryland Statewide Transportation Model (MSTM). These projections include the metropolitan planning organization employment forecasts for the Baltimore and Washington region expanded with forecasts from state agencies in rural areas (Tadayon & Shemer, 2013). Second, I was unable to integrate SILO with a transportation model at this time, so this experiment does not directly model feedback between land use and transportation. Travel times on the network are thus treated as an uncertainty, as explained further below.

This instantiation simulates the Baltimore-Washington region (Figure 2), where SILO has already been exercised in exploratory scenarios and AV modeling (Knaap et al., 2020). The Baltimore-Washington region is an older US region with two downtowns and several major suburban job centers. Many central areas and inner suburbs possess limited capacity to absorb new development under current land-use regulations, something that SILO accounts for. To track regional growth, all counties are either assigned as core, inner, outer, or beyond the region. Core jurisdictions are Baltimore and Washington, DC. Inner suburbs are those adjacent to the core jurisdictions and Howard County, which is well suburbanized at this point. Outer jurisdictions include the remaining jurisdictions within the two metropolitan planning organizations. Beyond the region is largely rural but does include smaller population centers in Wilmington, DE, York and Lancaster PA, and Ocean City, MD. SILO does not model beach home development in Maryland and Delmarva.

The uncertain parameters (Table 1) within the SILO model align with those uncertainties elevated in the literature on AVs and previous modeling (Llorca et al., 2022; Soteropoulos et al., 2019; Sperling, 2018). My use of "uncertain parameters" reflects that what is now uncertain in the model are fixed or stochastic parameters, such as the value of access or distance to work constraints. It is also consistent with the previous literature (Groves & Lempert, 2007). Uncertainty parameters include the auto operating cost, increased infill capacity due to lower parking demand, travel times, and three parameters reflecting changing values of accessibility: the value of access in location decisions, zonal accessibility score, and distance to work constraints (Table 1). The ranges were determined from estimates in the literature and previous AV modeling (Litman, 2018; Soteropoulos et al., 2019).

Because this experiment was not integrated with a travel demand model, all the parameters are within the SILO location choice module. My uncertain parameter in this case then selects from two zone-to-zone travel time scenarios rather than modeling the full impact of vehicle automation and household relocation on travel times. In the first case, AVs use the road space as efficiently as human-driven cars—the 2030 baseline zone-to-zone travel times from the MSTM. In the second case, AVs use the road space more efficiently—2015 travel times are maintained throughout the simulation even as the population grows. This constitutes a 19.9% reduction of travel times relative to baseline, well within the results found in previous studies which found vehicle hours traveled changes from –41% to +25% according to a 2019 review of transportation modeling studies (Soteropoulos et al., 2019).

Table 1. Uncertain parameters

Uncertain Parameter	*Impact of AVs*	*Baseline value*	*Sample range*	*Sources for range*
Auto operating cost	Increase with new technology; decrease with increased platooning and efficient driving	8.4 cents/mile	2.1 – 12.3 cents per mile	Reductions in energy and use intensity up to between -75% and +30% (Brown et al., 2014; Stephens et al., 2016)
Infill capacity	Allows for redevelopment of existing parking	Set at zone level	0-50% increase in capacity for new units	Reduction in parking could be up to 90% in high adoption scenarios. (Zhang et al., 2015; Zhang & Wang, 2020)
Relative value of access in location choice	Decreases value of access due to in-vehicle comfort	Set by income group	0-25% decrease	Value of time reduction between 18% and 50% for personal automated vehicles (Andrei et al., 2022; Kolarova et al., 2019; Steck et al., 2018; Zhong et al., 2020)
Distance to work constraint	Willing to move further from work due to in-vehicle comfort	Travel times = $\Gamma(k = 2, \theta = 17.2)$	θ in [17.2, 34.4]	Value of time reduction between 18% and 50% for personal automated vehicles (Andrei et al., 2022; Kolarova et al., 2019; Steck et al., 2018; Zhong et al., 2020)
Zonal access to jobs beta – Hansen accessibility	Decrease the overall value of proximity	3	β in [1.5,3]	Value of time reduction between 18% and 50% for personal automated vehicles (Andrei et al., 2022; Kolarova et al., 2019; Steck et al., 2018; Zhong et al., 2020)
Zone-to-zone travel times	Decrease with AV efficient use of network	2030 baseline travel times from MSTM	Binary: {2030 baseline travel times; 2015 travel times (-19% average travel time)[4]}	Childress et al (2015) found between -41% and + 17% change in vehicle hours traveled.

Policy interventions (Table 2) include common approaches for encouraging concentrated land uses. Increasing *transit-oriented development capacity* is modeled via a 25% increase in capacity, measured in allowable new dwelling units, in zones with transit stations. Without integration with a travel model, I assume that increasing the *fuel tax by* 1 cent/mile decreases travel times over the network and household travel time to work by .5%. The 1 cent per mile increase is effectively a 1.2% increase on the baseline auto-operating costs in the SILO model, which includes fuel costs and other mileage-

[4] Estimated in previous modeling work; approximate for the purposes of this research

dependent costs. A review of several of the literature found long-run VMT demand elasticities with respect to fuel cost as high as -.4, including in a recent (Litman, 2022). One of the recent studies that finds an elasticity of -.4 uses a microeconomic model to determine the fuel use impacts of AVs (Taiebat et al., 2019). The FHWA reports a one-to-one relationship between VMT change and travel time changes nationwide (Brand, 2009). This would translate to an effective .48% reduction in travel times.

Each policy is examined twice: first starting at simulation year 0 and delaying each policy to start in year 6 (2021). In "delayed" policy runs, SILO is run with the baseline settings for year 0-5. This sets the groundwork for adaptive policy approaches (Walker et al., 2013). One hundred scenarios are sampled using the LHS, which is comparable with other scenario discovery experiments in terms of the density of the sample in the multi-dimensional space (Swartz & Zegras, 2013).[5]

Table 2. Policy alternatives modeled

Policy	***What is it?***	***When is it implemented?***	***How it is implemented in SILO?***
Baseline	No additional policy action	NA	NA
Transit-oriented development	Expanded residential development capacity at heavy rail, light rail, and commuter rail	2015 simulation year	50% increase in residential unit capacity in zones with heavy rail, light rail, and commuter rail
Delayed transit-oriented development (year 6)	Expanded residential development capacity at heavy rail, light rail, and commuter rail	2021 simulation year	50% increase in residential unit capacity in zones with heavy rail, light rail, and commuter rail
Gas price increase	Increase gas price by 1 cent per mile	2015 simulation year	.5% decrease in the zone-to-zone travel times .5% decrease in travel times to work preferences
Delayed gas price increase (year 6)	Increase gas price by 1 cent per mile	2021 simulation year	.5% decrease in the zone-to-zone travel times .5% decrease in travel times to work preferences

Finally, seven indicators (Table 3) capture additional points of comparison between the LHS sample and more limited scenario approaches. For instance, the results for a single indicator might indicate that the full LHS and the more limited sample produce a similar measure of policy robustness. However, comparing the robustness between LHS and a more limited sample across several indicators will help to determine whether the robustness measures are consistently similar or different.

[5] See methodological appendix for additional details on the selected number of scenarios

In providing multiple indicators for comparison, the results reflect different modules within the SILO model as well as different priorities for urban development. Core area households, inner suburban households, and high transit access households are all indicators associated with the concentration of households in already developed areas of the region. High transit-accessible households are those in zones that are 75th percentile or higher in access to employment via transit. This includes some zones that do not contain rail stops, such as zones near the core with high-frequency bus service, and excludes some zones with rail, such as outer suburban zones with infrequent commuter rail. Outer suburban households and households beyond the metro areas measure dispersion. Median housing cost track housing affordability. These indicators are also directly produced by the real estate development module rather than the household relocation module. Finally, households that are located in modeling zones that are higher than 75% targeted ecological area (TEA)[6] are a proxy for environmental impacts.

Table 3. Indicators

Indicator	***Description***	***Purpose***	***Desired direction***
Core households	Number of households located in Baltimore and Washington, DC	Measure relative regional concentration	Higher
Inner Suburban	Number of households located in inner suburban jurisdiction	Measure relative regional concentration	Higher
Outer Suburban	Number of households located in outer suburban jurisdiction	Measure relative regional dispersion	Lower
Beyond Region	Number of households located beyond the two regions as defined by MPO boundaries	Measure relative regional dispersion	Lower
High transit accessibility	Number of households located in zones that are 75th percental or higher in access to employment	Measure growth in areas with high regional access via transit	Higher
Households in > 75% TEA	Number of households that are located in zones that > 75% targeted ecological areas by land area	Measure environmental impacts of development patterns	Lower
Median housing prices ($)	Median housing price within the region	Measure regional housing cost impacts	Lower

[6] Targeted ecological areas are watersheds in the top decile for protection, as designated by the Maryland Department of the Environment.

To determine the information value of scenario discovery, I consider the sensitivity, the percentage of futures in the box that fail, and the precision, the percentage of all failing scenarios captured in the PRIM boxes. Additionally, I examine the results of the 100 futures against two scenario sets that approximate exploratory scenarios. The first case is the convex hull of the LHS sample (9 scenarios) and the second case selects the eight extreme points from three uncertainty dimensions (8 scenarios). To limit the dimensionality, I eliminated the per-mile cost of the automobility parameter and set the three accessibility parameters to vary together, i.e., they are together set to either their highest or lowest values. Both of these sampling techniques examine only futures on the outer edge of the sampling space, designed to resemble exploratory scenarios that focus on the edge of plausibility.

6 Results

6.1 Baseline automated vehicle futures

I begin with an examination of the 100 baseline futures. For every indicator, AV scenarios scored above and below the default, no-AV scenario (Table 4). On average, more households were located in the core (+1.4%), more households could access transit (1.3%), and fewer households were cost-burdened (-3.3%). More importantly, all the outputs differ considerably from scenario to scenario. They vary as much as 10.3% (outer suburban households) and as little as 5.7% (households in TEAs). This reiterates the importance of considering multiple scenarios. Though deep uncertainties cannot be validated against data, the range of simulations included growth outcomes similar to other modeling efforts (Soteropoulos et al., 2019). For instance, past modeling has found that core residential growth ranged from +.7% to -5.7% relative to the baseline. The same experiments found that residential growth outside the core increased between 1-5.7% (Gelauff et al., 2017; Thakur et al., 2016; Wellik & Kockelman, 2020). Our baseline futures also included outcomes beyond the previous finding, such as core household growth increasing by 6.1%. This is in part explained by the volume of futures examined in this experiment. Whereas the cited papers each examine less than ten scenarios, this experiment includes 100 baseline scenarios. We should expect some results beyond the bounds of previous experiments. Additionally, the Baltimore-Washington application case is in a different context from previous experiments.

Table 4. Baseline automated vehicles futures

Indicator	***No AV***	***Scenario Range (Relative to No-AV)***	***Mean (Relative to No-AV)***	***Standard Deviation***
Core households (thousands)	833.09	(810.9, 884.0) -2.7%, 6.1%	844.9 1.4%	20.0
Inner Suburban (thousands)	2,378.2	(2,324.9, 2,440.0) -2.2%, 2.6%	2,387.2 0.1%	28.0
Outer Suburban (thousands)	603.8	(571.8, 631.0) -5.3%, 4.5%	602.3 -0.2%	16.1
Beyond Region (thousands)	1,895.4	(1,828.0, 1,928.6) -3.6%, 1.8%	1,878.4 -0.9%	29.1
High transit accessibility households (thousands)	294.7	(285.9, 315.1) -3.0, 6.9%	298.6 1.3%	7.6
Households in > 75% tea (thousands)	358.1	(347.7, 367.7) -.3%, 2.7%	357.9 -0.0%	4.7
Median housing prices ($)	709	(669, 726) -5.7%, 2.4%	708.8 0.0%	11.8

6.2 Robustness of policies in automated futures

From the policy perspective, transit-oriented development (Figure 3: orange bars) is highly robust across the majority of indicators, with robustness greater than .9 for five of the indicators. It is never the worst-performing policy. Delaying transit-oriented development (Figure 3: gray bars), however, reduces robustness (or conversely increases regret) on nearly all indicators. Though increasing the price of gasoline (Figure 3: yellow bars) is the most robust with respect to core growth, it performs middling or poorly on all the other indicators. The low robustness with respect to inner household growth can partly be explained by the better performance within the core—the two regions often compete for residents. However, the lack of additional growth capacity also deflects a significant portion of growth to the outer suburban tier and beyond the region. Interestingly, the no-policy baseline often performed better than the studied interventions for several of the indicators. Similarly, all policies that perform well in encouraging inner suburban growth perform poorly in encouraging core growth.

Housing unit prices are the best measure of equity within this experiment. All else equal, lower housing prices throughout the region will reduce the relative cost of housing most substantially for lower-income households. Transit-oriented development performs particularly well in lowering housing costs. The results speak to costs throughout the region that are relieved due to increased supply, not specifically costs close to the transit stations. Many areas near transit are among the most well-developed in the region and close to buildout capacity.[7] They are also often in popular core and inner jurisdictions. Opening up capacity near these transit stations is valuable for creating more units and relieving prices where demand is high in almost all scenarios. This is further affirmed in the baseline runs—the infill capacity variable was by far the most important in lowering housing costs.

While TOD reduces housing costs relative to other policies, its performance is middling in advancing other TOD goals, such as encouraging growth in high transit accessibility locations or reducing the impact on ecological areas. Rail service in the Baltimore-Washington region is relatively suburban-oriented. Opening all TOD areas only means opening up land in many car-oriented areas while maintaining restrictions on urban core areas that have great bus but no rail service.

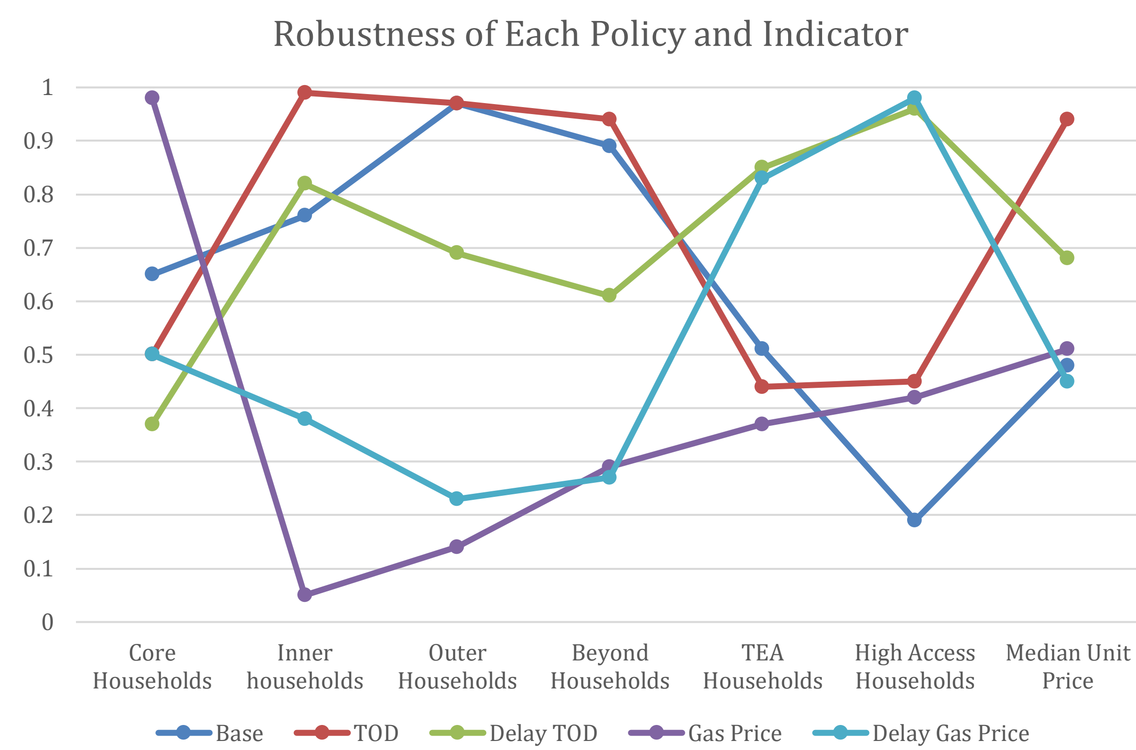

Figure 3. Policy robustness for each indicator

6.2.1 Conditions for policy success

The conditions for policy success are determined by the PRIM algorithm, which generates multidimensional boxes designed to capture high concentrations of failed

[7] Fully built out zones may be high or low density depending on their zoning. Increasing development capacity implies either relaxed zoning or regulations that are exempt from zoning

policy futures. Not all PRIM boxes performed equally well in identifying the conditions under which a policy performs well. PRIM generates several boxes for each policy/indicator combination along the density and coverage tradeoff. By increasingly further restricting the uncertain dimensions, PRIM changes the box containing failed futures, but any changes can either improve the density or coverage, usually at the expense of the other. These many boxes are often represented in the density/coverage tradeoff curve seen in many other studies such as Bryant and Lempert (2010). From these many boxes, I chose the box with the highest density because that box identifies the most uncertainties that could impact the success or failure of the outcome.

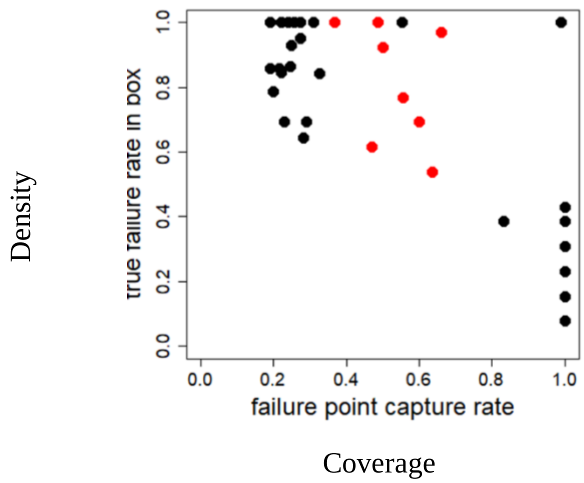

Figure 4. Density and coverage of PRIM boxes for each policy/indicator combination

Nonetheless, many of the thirty-five policy/indicator boxes still appeared to be uninformative because of low overall density or coverage. The lower coverage was expected given the selection of the highest-density boxes. Surprisingly, several of these high-density boxes still contained more successes than failures. Given that limitation, I selected eight PRIM boxes that met my criteria for performance: performed above the median for density (>.615) and coverage (>.347), a robustness score between .05 and .95 (Figure 2). These are additional thresholds not often applied in scenario discovery analysis but which I found useful when considering multiple indicators. The decision to only select those performing above the median was a professional judgment that reflects a commitment to only using the most informative PRIM results. Table 5 below presents the limiting dimensions for the eight selected PRIM boxes[8] (red points in Figure 4). The limiting dimension of those uncertainties that best characterize the region of high regret. In testing the performance of a policy for any given indicator, only some of the uncertainties are influential in determining this region. The first two columns then indicate the eight PRIM boxes that were identified in the above analysis. The next column indicates the limiting dimensions. The final two columns indicate the range covers the region of high regret. All the uncertainty ranges are standardized from 0 to 1.

[8] The PRIM procedure sometimes determines boxes whose limits extend beyond allowable parameter values—less than zero or greater than one. For interpretation these values are no different from the true limit value.

For instance, the baseline policies perform worse in promoting core households when distances to work are shorter, the value of access is higher, and travel times are briefer. Markings on the limiting dimensions indicate whether high regret areas resemble default value (* in Table 5), more radical changes († in Table 5), or are ambiguously in between (unmarked).

Table 5. Latin hyper-sample robustness contingent policy dimensions; markings on the limiting dimensions indicate whether high regret areas resemble default value (*), more radical changes (†), or are ambiguously in between (unmarked)

Policy	Indicator	Limiting Dimensions	Low Limit	High Limit
Base	Core households	Distance to work* Value of access* Travel times†	NA .28 NA	.45 NA .75
Base	Inner households	Infill capacity* Distance to work* Travel times*	.18 .08 .24	.71 .49 NA
Base	Beyond households	Infill capacity† Value of access	.35 .21	.92 .80
TOD	Core households	Distance to work* Value of access*	NA .26	.65 NA
Delay TOD	Inner households	Infill capacity* Distance to work*	.18 .07	.78 .47
Delay TOD	TEA development	Infill capacity† Zonal access to jobs beta* Travel times†	.20 .44 .06	.90 .97 .73
Gas Prices	Housing unit prices	Infill capacity† Distance to work* Value of access* Travel times*	.21 NA .38 .25	.95 .75 .94 NA
Delay Gas Prices	TEA development	Infill capacity Value of access†	.25 .13	.77 .74

The first thing that stands out is that there are often clear conditions under which the baseline policy is regretful. When the distance to work preferences and value of access preference are similar to the default values, the baseline does poorly in promoting core households. In such circumstances, higher gas prices are better at promoting core growth. While higher gas prices still encourage core growth when people are willing to live further from work and other daily activities, they are less effective relative to the baseline. When households are willing to live further out because AVs have reduced the value of proximity, higher gas prices encourage households to locate in more central locations in the inner suburbs, rather than locating in the core.

Similarly, when infill capacity, distance to work, and travel times are close to the defaults, the baseline scenario often fails to promote inner suburban growth relative to

TOD policies. Conversely, this means the no policy alternative is less regretful when AVs are most impactful: opening up parking lots for redevelopment, encouraging people to live further from work, and reducing travel times. In this case, the power of TOD to outperform the baseline is limited when AVs already open up significant capacity everywhere. However, this same infill capacity causes regret in baseline policies when attempting to discourage movement beyond the region. It seems that there is still some additional regional demand that TOD can soak up.

As seen in the previous table on robustness, TOD does not always perform well in encouraging core growth because it opens up so much capacity in attractive inner suburbs. This tends to happen with values closer to the baseline parameter values: stronger preferences for work proximity and general access. In such cases, both the baseline and gas price scenarios perform better. If AVs loosen these preferences, fewer households choose the core in the baseline and gas price policy scenarios, generally weakening the relative power of these policies encouraging core growth relative to TOD. A similar pattern is noted for delayed TOD failing to encourage inner suburban growth. When infill capacity is increased from AVs, the delay in implementing TOD is less costly because those communities can already absorb the increased demand.

Distance to work, the value of accessibility, and travel times determined regions of high regret in four or more of the selected policies. The results indicate that if AVs encourage longer commutes, our existing policies for encouraging core and inner suburban development are blunted. Across the indicators, AVs often reduce the difference between the best-performing and the worst-performing policies. An exception is the uncertainty regarding the infill capacity that AVs will open up. As central areas of the DC region are quite attractive, any additional capacity can aid in holding down housing prices.

In terms of housing prices, the chosen equity indicator, gas prices are identified as a contingent policy. In general gas prices are a regretful policy whenever the draw to live in the center is already strongest. As gas prices encourage core living, the increased demand for limited core housing units can exacerbate housing costs. This is true when AV impacts are closer to baseline: lower distance to work, higher value of access, and lower travel times. Additionally, when AVs open up significant development capacity, higher gas prices steer households away from the inner suburbs where much of that capacity is available.

6.2.2 Measuring the relative effectiveness of scenario discovery

Table 6 provides the failure rate for the full LHS sample, the convex hull, and the exploratory scenarios. This is the proportion of all scenarios in which the policy produced a failed outcome. Color coding is used to indicate policies that perform particularly well or poorly (Green: <.1; light green: <= .25, light yellow >= .75, gold > .9). For instance, increased gas prices are a robust policy in all samples for promoting core household growth, failing only 2% of the time in the LHS sample, 11% of the time in the convex hull sample, and 0% of the time in the extreme points sample. On the other hand, the failure rates for increasing gas prices with respect to median unit price differ substantially between the three sampling approaches. With the LHS sample, the gas price increase fails 49% of the time; with the convex hull sample, the gas price increase fails 22% of the time; and with the extreme points sample, the gas price policy fails 75% of the time. This suggests that the sample does matter for measuring robustness and that the LHS sample might be preferable in some circumstances.

Table 6. Failure rates for all scenarios

Latin hyper-sample	Core Households	Inner households	Outer Households	Beyond Households	TEA Households	High Access Households	Median Unit Price
Base	.35	.24	.03	.11	.49	.81	.52
TOD	.50	.01	.03	.06	.56	.55	.06
Delay TOD	.63	.18	.31	.39	.15	.04	.32
Gas Price	.02	.95	.86	.71	.63	.58	.49
Delay Gas Price	.50	.62	.77	.73	.17	.02	.55

Convex Hull	Core Households	Inner households	Outer Households	Beyond Households	TEA Households	High Access Households	Median Unit Price
Base	.22	.11	.00	.00	.56	.78	.56
TOD	.44	.00	.00	.00	.44	.56	.22
Delay TOD	.67	.22	.33	.44	.11	.00	.56
Gas Price	.11	1.00	1.00	.78	.89	.67	.22
Delay Gas Price	.56	.67	.67	.78	.00	.00	.44

Extreme Points	Core Households	Inner households	Outer Households	Beyond Households	TEA Households	High Access Households	Median Unit Price
Base	.50	.50	.25	.50	.625	.75	.375
TOD	.75	.25	.25	.375	.625	.75	.00
Delay TOD	.50	.00	.25	.375	.25	.375	.375
Gas Price	.00	.50	.50	.25	.375	.125	.75
Delay Gas Price	.25	.625	.75	.50	.125	.00	.50

The robustness of each policy/indicator combination. Color coding is used to indicate policies that perform particularly well or poorly. Green: <.1; light green: <= .25, light yellow >= .75, gold > .9

To understand how the alternative approaches compared to scenario discovery, the binomial p-value compares the robustness scores for each scenario/policy pair. In this way, I sought to determine whether the robustness scores from the convex hull and the extreme points sample differed significantly from the robustness scores for the LHS sample. If the alternative samples produced similar results, on average, we should expect that the p-values should resemble those from 28 random subsamples—roughly evenly distributed between 0 and 1.

The CDF of the p-values for each is in Figure 5. Assuming that the LHS sampled scenarios provide an accurate picture of the uncertainty space, these p-values indicate the probability with which the alternative samplings would provide inaccurate estimates of robustness. The convex hull sampling performed relatively well, with all p-values higher than .2 and more than half greater than .8. While it is not a surprise that a subsample should resemble the overall sample; the performance of the convex hull is generally better than we should expect from even a true random sample of the LHS scenarios—all the p-values are greater than .2, whereas with a random sample we would expect roughly 5 robustness scores less than .2. This might result from a sampling approach that will not incidentally overdraw from one region of the uncertainty space. We can then conclude that the convex hull provides a relatively accurate, albeit imprecise measure of robustness. Of course, all scenarios included were also in the LHS sample. Exploratory scenarios, however, can provide very misleading impressions of robustness. Several scenario/policy pairs have p-values less than .05.

Of course, scenario discovery methods are also used to clarify the conditions under which a policy is likely to succeed. Both the convex hull and the extreme values samplings clearly cannot do this because they do not provide enough information about the center of the distributional range. Of course, even with the 100 LHS scenarios, not all PRIM boxes generated clearly identified parameters influencing the success of a policy. More than half the PRIM boxes captured over 80% of the failure futures, but often at the expense of including many successful futures. The majority of boxes were more than half successes. By the criteria listed above, one-quarter of the boxes produced clear policy suggestions. In practice, thresholds for density and coverage should be determined by the risk tolerance of decision-makers.

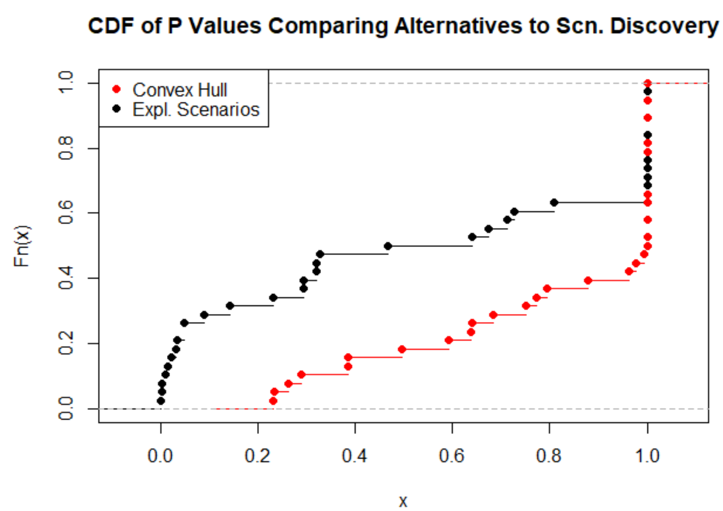

Figure 5. CDF of p values for sampling alternatives

6.2.3 Modeling resources

On average, a single SILO run took 134 minutes to simulate fifteen years on a twenty-core Windows server with dynamic memory up to 128G. Each processor was an Intel(R) Xeon(R) CPU E5-2667 v2 @ 3.30GHz. That means that the complete scenarios discovery run time was just over 1,133 hours of computer run time—more than 47 days. These run times are for the land-use model alone as it was not integrated with a transportation model at this time. Any time stopping, starting, pre-processing, and post-processing was trivial compared to the total run time. The model area is large by North American standards. SILO micro-simulates over 8,000,000 people and 5,000,000 households. Nonetheless, such a similar experiment would still require days of runtime in a smaller region. The 500 runs require more than ten times the computer hours than the 45 convex hull runs or 40 extreme points runs. This will remain true, even if modelers can reduce run times.

Nonetheless, if run times are sufficiently reduced, the magnitude difference between the run time for scenario discovery run times and the run time for other approaches will no longer be an inhibiting factor. Several approaches could have sped up the overall time used. First, I could have reduced the number of uncertain parameters and policies. The latter would have had more of an influence. To maintain the same density of scenarios in multidimensional space, the number of runs needs to increase exponentially with the number of uncertain parameters (see methodological appendix). With only three uncertain parameters, I would have explored the same density of the space with just 10 scenarios. Reducing the number of policies would also reduce the total run time, but only linearly.

The second approach would be to reduce the scope of the experiment temporally or geographically. If SILO was run for fewer years clear impacts might have been evident with far less model time. Similarly, the SILO version for Baltimore-Washington contains significant population centers beyond the region, as explained in the methods. By cutting out areas east of the Chesapeake Bay and western areas of the model region, the micro-simulated population would have been reduced by over 1,000,000 people. Simulations in smaller regions would have significantly smaller run times.

The final approach to reducing the run time is to use faster servers. Of course, model complexity has tended to increase with computing power, so the modelers may want to maintain less complex models for exploratory approaches like this one. This is also in line with established practice in exploratory modeling—most authors recommend faster-running, simpler models to increase exploration.

7 Conclusions

This research reinforces the need for urban modelers to increase the scenario count they use to explore the parameter space. The results of this analysis suggest that robust decision-making analysis using scenario discovery is a useful design for sorting through a high number of scenarios. My results indicate that full scenario discovery offers value over more limited explorations of uncertainty space in identifying contingent policies. Selecting limited scenarios at the edges of possibility can often overlook vast regions of robustness. Planners ought to be aware of extreme scenarios that break largely robust systems; however, they should also understand when the scenarios are rare outliers. Exploratory scenario exercises that incorporate modeling could easily give such false impressions. My results confirm the results of previous research, which indicates that sampling approaches significantly impact measures of robustness (McPhail et al., 2020). Additionally, utilizing large enough LHS samples also supports the deployment of PRIM to determine potential thresholds between policy options.

If the modelers cannot dedicate computational resources to simulating an LHS sample, they should consider simulating the convex hull. In this experiment, the run time for the scenario discovery runs was greater than an order of magnitude longer than the convex hull. Nonetheless, the convex hull sample performed comparably in identifying the robustness of policies. Modelers may also wish to restrict the number of uncertain parameters, as they would be able to explore a similar density of scenarios in the multi-dimensional space with far fewer model runs. Exploratory scenarios may initially be easier to experiment with in smaller regions with fewer agents to micro-simulate. Those smaller regions can provide a testing ground while computational power increases sufficiently to run land-use models like SILO in far less time.

Exploratory modelers might suggest a simpler modeling system to increase run times (Bankes, 1993). While there may be use cases for more aggregate modeling, the choice to use simpler models will depend on which indicators modelers choose to investigate. In particular, microsimulation models are better for estimating the distributional of outcomes between different populations, and thus more useful for studies of equity (Dawkins & Moeckel, 2016). Downsampling the microsimulation model may be an appropriate compromise when modelers can tolerate less precision and don't plan on analyzing the microscopic output (Llorca et al., 2020). Combining scenario discovery with microsimulation could provide a helpful tool in determining robust policies for regional equity.

The decision to use scenario discovery in urban planning depends on the goals of the planning process. If the primary goal of the planning is to inform key stakeholders about important uncertainties and consider how those uncertainties could play out, exploratory scenarios may still be the preferred approach. If the objective of planning is to determine robust and continent policies to include in long-range planning, scenario discovery has already exceeded the performance of many exploratory scenario exercises. For agencies that have the resources to conduct a rich stakeholder-driven process and execute several hundred simulations, there is great potential in combining the two approaches.

In considering the land-use impacts of AVs, this experiment confirms some concerns regarding the influence of AVs on household location choice decisions. Should autonomous vehicles devalue accessible locations, households will move further out than they otherwise would. More surprisingly, however, if autonomous vehicles free up urban space dedicated to auxiliary vehicle uses, such as parking, it won't always counteract core household dispersion. Rather, inner suburban communities possess a far vaster supply of easily redeveloped land, and, in the Baltimore-Washington Region, these communities are often among the most desirable places to live. On the other hand, when additional room is not provided in the inner suburbs, the core often benefits from households that prefer urban living with AV-enhanced access to the ring of suburban job centers.

In increasing the distance households are willing to locate from work, AVs may dull the effectiveness of smart growth policies. The highest regret scenarios, for both baseline and TOD, are associated with travel distances to work and accessibility preferences similar to current levels. The relative advantage of opening up new room for development closer to the core is diminished by the de facto opening of land on the fringe. On the other hand, if AVs do not allow for the redevelopment of parking, early TOD is essential in ensuring that land near transit stops is not underutilized.

This experiment also highlighted an understudied dimension of TOD policies that AVs are only bound to exacerbate. Not all TOD sites are created equally. Though TOD always performed well in preventing additional development on the margins of the study area, the development did not correspond to living in high accessibility areas. The Baltimore-Washington heavy rail and commuter rail network is suburban-oriented relative to older North American systems and context-insensitive TOD will likely open up significant development in locations that, despite their train stop, do not provide high regional transit accessibility. A resident in the non-rail Brightwood Park neighborhood in Washington is much more likely to drive less and live in a less energy-intensive home than a household living on the Reston Metro stop in suburban Virginia.

In reality, the competition for TOD living between the core and suburbs will be far less than the 50% increase assumed within my scenario. Nonetheless, MPOs and states would be wise to guide regional TOD strategies that prioritize sites that are currently, or will soon be, highly accessible via transit. Such an approach should not be used to preclude suburban or small-town TOD but to create an incentive structure to encourage transit-supporting context beyond the immediate site, such as more local bus routes of increased frequency. Specific sites might also be appropriate if they localized imbalanced regional travel patterns by providing a greater variety of opportunities in overwhelming residential areas.

According to the modeling and analysis presented above, the potential for AVs to revolutionize where people are willing to live will not necessarily exacerbate sprawl and associated environmental impacts. Both increasing willingness to travel and redevelopment of parking areas will serve to decrease housing prices at a regional scale. Prices, however, may increase even more quickly in high amenity areas that are suddenly easily accessible to even more jobs than before. Quality of place will become even more critical than before as households are increasingly free to live wherever they choose. The most important amenities will be those that people cannot easily travel to access, such as school districts for community safety.

This has important implications for planning policy. Though efforts to guide development, such as TOD, may be less effective overall, they will be all the more powerful in already desirable locations, such as the Baltimore-Washington inner suburbs. For some households, the most desirable place will be a house far away in the woods, but for many, the immediate appeal of specific neighborhoods, such as highly rated schools and low crime, will prevail if the prices are not too high. Though not included in this version of SILO, other quality-of-life factors may be crucial for location choice in some contexts. Though this experiment examined TOD policy specifically, AVs could provide the possibility for exciting, dense, walkable redevelopment anywhere with their ability to drop off passengers and depart to unseen locations.

This experiment also found that AVs could have potential equity implications concerning the cost of housing. Most clearly, anything that opens up new development capacity supports more generally affordable housing. This is true for both AVs allowing for the redevelopment of parking lots and conventional TOD policies. This research, however, also found that increasing gas prices could increase housing costs in an AV future. In encouraging core growth, they also encourage growth in the most capacity-constrained areas of the region. Gas prices are not a regretful policy when AVs have decreased the value of proximity and shortened travel times. But this is only because the gas price policy is no longer successful in centralizing growth.

This experiment has several limitations that also open up pathways for future research. First, SILO only simulates residential location choices. Future employment is taken exogenously. While previous experiments with SILO in the Baltimore-Washington region have assumed different distributions (Knaap et al., 2020), I chose not to do that in this experiment. Directly simulating commercial and retail decisions would provide a much more complete sense of potential AV impacts. Historically, businesses responded to widespread car ownership by choosing more decentralized locations, which further encouraged the dispersion of households. The potential for such dynamics should be tested with a fully integrated LUTI model.

A second limitation is the lack of full integration with a transportation model. In this experiment, I took travel times as an uncertain parameter, rather than simulating them in response to land-use changes. Though the overall pattern of households in the model was similar enough from run to run, it is unlikely to significantly change travel patterns, household agents were nonetheless unable to respond to travel time changes dynamically. Because I did not run a travel demand model, I was also unable to gauge the impact that the full integration would have on run times, but prior experience indicates that it would have surely inflated the already long run times.

This is also just a single experiment comparing scenario discovery to the convex hull and extreme points. While my results regarding different robustness scores conform with a previous experiment in a different domain (McPhail et al., 2020), I cannot be certain that the results are peculiar to the SILO model as instantiated in this region and the unique LHS sample. Additionally, I selected to examine the highest density box to capture all relevant uncertainties. Future experiments should test whether these results hold with different models, contexts, indicators, box selection, and sampling procedures. Providing practitioners with procedures to determine what sample to use in their case would be even more valuable.

Finally, this experiment remains a largely technical exercise in scenario discovery. I compare the results of a scenario discovery experiment to more limited approaches, but this work has not yet been translated to decision makers such as MPO board members selecting from various investment profiles. While the language of robust and contingent policies has already entered the discourse via exploratory scenarios, I cannot conclude whether these model results can be usefully applied. There are two potential challenges in that regard. First, the results of scenario discovery might be too technical for translation. Second, regional decision-makers might not possess the policy agility to apply adaptive policies, particularly when those policies, such as regional TOD, would require multi-party collaboration. Determining the value of exploratory scenarios to regional modeling and planning practice will thus require research that directly engages decision-makers.

Data availability

Project code is available on Github:
https://github.com/dengelberg/JTLU_ScenarioDiscovery.git.

Data from the Silo runs is stored on Google Drive:
https://drive.google.com/drive/folders/1sqoT_Grl9lcVE124mP5VbhcErX1qRzUp?usp=sharing.

Appendix

Appendix available as a supplemental file at https://doi.org/10.5198/jtlu.2024.2401.

References

Andrei, L., Luca, O., & Gaman, F. (2022). Insights from user preferences on automated vehicles: Influence of socio-demographic factors on value of time in Romania case. *Sustainability*, *14*(17), 10828. https://doi.org/10.3390/su141710828

Avin, U. (2007). Using scenarios to make plans. In L. D. Hopkins & M. A. Zapata (Eds.), *Engaging the future: Forecasts, scenarios, plans, and projects* (pp. 103–134). Cambridge, MA: Lincoln Institute of Land Policy.

Avin, U., & Dembner, J. (2000). Getting scenario-building right. *Planning*, *67*(11), 22–27.

Bankes, S. (1993). Exploratory modeling for policy analysis. *Operations Research*, *41*(3), 435–449.

Basu, R., & Ferreira, J. (2020). A LUTI microsimulation framework to evaluate long-term impacts of automated mobility on the choice of housing-mobility bundles. *Environment and Planning B: Urban Analytics and City Science*, *47*(8), 1397–1417. https://doi.org/10.1177/2399808320925278

Bootz, J.-P. (2010). Strategic foresight and organizational learning: A survey and critical analysis. *Technological Forecasting and Social Change*, *77*(9), 1588–1594. https://doi.org/10.1016/j.techfore.2010.06.015

Börjeson, L., Höjer, M., Dreborg, K.-H., Ekvall, T., & Finnveden, G. (2006). Scenario types and techniques: Towards a user's guide. *Futures*, *38*(7), 723–739. https://doi.org/10.1016/j.futures.2005.12.002

Brand, D. (2009). *Impacts of higher fuel costs*. Washington, DC: Federal Highway Administration. https://www.fhwa.dot.gov/policy/otps/innovation/issue1/impacts.cfm#o

Brown, A., Gonder, J., & Repac, B. (2014). An analysis of possible energy impacts of automated vehicles. In G. Meyer & S. Beiker (Eds.), *Road vehicle automation*. New York: Springer. https://doi.org/10.1007/978-3-319-05990-7_13

Bryant, B. P., & Lempert, R. J. (2010). Thinking inside the box: A participatory, computer-assisted approach to scenario discovery. *Technological Forecasting and Social Change*, *77*(1), 34–49. https://doi.org/10.1016/j.techfore.2009.08.002

Chakraborty, A., & McMillan, A. (2015). Scenario planning for urban planners: Toward a practitioner's guide. *Journal of the American Planning Association*, *81*(1), 18–29. https://doi.org/10.1080/01944363.2015.1038576

Childress, S., Nichols, B., Charlton, B., & Coe, S. (2015). *Using activity based model to explore possible impacts of autonomous vehicles*, 2493(1), 99–106.

Cordera, R., Nogués, S., González-González, E., & Moura, J. L. (2021). Modeling the impacts of autonomous vehicles on land use using a LUTI model. *Sustainability*, *13*(4), 1608. https://doi.org/10.3390/su13041608

Dawkins, C., & Moeckel, R. (2016). Transit-induced gentrification: Who will stay, and who will go? *Housing Policy Debate*, *26*(4–5), 801–818. https://doi.org/10.1080/10511482.2016.1138986

Friedman, J. H., & Fisher, N. I. (1999). Bump hunting in high-dimensional data. *Statistics and Computing*, *9*(2), 123–143.

Gelauff, G., Ossokina, I., & Teulings, C. (2017). *Spatial effects of automated driving: Dispersion, concentration or both?* Unpublished paper. Retrieved from https://doi.org/10.13140/rg.2.2.32766.48965

Gross, E. (2018). *Scenario discovery for a future of automated mobility on-demand in the urban environment* (master's thesis), Massachusetts Institute of Technology, Cambridge, MA.

Groves, D. G., & Lempert, R. J. (2007). A new analytic method for finding policy-relevant scenarios. *Global Environmental Change*, *17*(1), 73–85.

Hall, J. W., Lempert, R. J., Keller, K., Hackbarth, A., Mijere, C., & McInerney, D. J. (2012). Robust climate policies under uncertainty: A comparison of robust decision making and info-gap methods. *Risk Analysis*, *32*(10), 1657–1672. https://doi.org/10.1111/j.1539-6924.2012.01802.x

Jafino, B. A., & Kwakkel, J. H. (2021). A novel concurrent approach for multiclass scenario discovery using multivariate regression trees: Exploring spatial inequality patterns in the Vietnam Mekong Delta under uncertainty. *Environmental Modelling & Software*, *145*, 105177. https://doi.org/10.1016/j.envsoft.2021.105177

Jafino, B. A., Kwakkel, J. H., Klijn, F., Dung, N. V., Van Delden, H., Haasnoot, M., & Sutanudjaja, E. H. (2021). Accounting for multisectoral dynamics in supporting equitable adaptation planning: A case study on the rice agriculture in the Vietnam Mekong Delta. *Earth's Future*, *9*(5), e2020EF001939. https://doi.org/10.1029/2020EF001939

Kim, K.-H., Yook, D.-H., Ko, Y.-S., & Kim, D.-H. (2015). *An analysis of expected effects of autonomous vehicles on transport land ese in Korea*. New York: NYU, Marron Institute of Urban Management.

Knaap, G.-J., Engelberg, D., Avin, U., Erdogan, S., Ducca, F., Welch, T. F., …, & Shahumyan, H. (2020). Modeling sustainability scenarios in the Baltimore–Washington (DC) region: Implications for methodology and policy. *Journal of the American Planning Association*, *86*(2), 250–263. https://doi.org/10.1080/01944363.2019.1680311

Kolarova, V., Steck, F., & Bahamonde-Birke, F. J. (2019). Assessing the effect of autonomous driving on value of travel time savings: A comparison between current and future preferences. *Transportation Research Part A: Policy and Practice*, *129*, 155–169. https://doi.org/10.1016/j.tra.2019.08.011

Kwakkel, J. H. (2017). The exploratory modeling workbench: An open source toolkit for exploratory modeling, scenario discovery, and (multi-objective) robust decision making. *Environmental Modelling & Software*, *96*, 239–250. https://doi.org/10.1016/j.envsoft.2017.06.054

Kwakkel, J. H., Auping, W. L., & Pruyt, E. (2013). Dynamic scenario discovery under deep uncertainty: The future of copper. *Technological Forecasting and Social Change*, *80*(4), 789–800. https://doi.org/10.1016/j.techfore.2012.09.012

Kwakkel, J. H., Haasnoot, M., & Walker, W. E. (2016). Comparing robust decision-making and dynamic adaptive policy pathways for model-based decision support under deep uncertainty. *Environmental Modelling & Software*, *86*, 168–183. https://doi.org/10.1016/j.envsoft.2016.09.017

Kwakkel, J. H., Walker, W. E., & Marchau, V. A. W. J. (2010). Classifying and communicating uncertainties in model-based policy analysis. *International Journal of Technology, Policy and Management*, *10*(4), 299–315. https://doi.org/10.1504/IJTPM.2010.036918

Lempert, R. J., & Groves, D. G. (2010). Identifying and evaluating robust adaptive policy responses to climate change for water management agencies in the American west. *Technological Forecasting and Social Change*, *77*(6), 960–974.

Lempert, R. J., Groves, D. G., Popper, S. W., & Bankes, S. C. (2006). A general, analytic method for generating robust strategies and narrative scenarios. *Management Science*, *52*(4), 514–528. https://doi.org/10.1287/mnsc.1050.0472

Lempert, R., Syme, J., Mazur, G., Knopman, D., Ballard-Rosa, G., Lizon, K., & Edochie, I. (2020). Meeting climate, mobility, and equity goals in transportation planning under wide-ranging scenarios: A demonstration of robust decision making. *Journal of the*

American Planning Association, *86*(3), 311–323. https://doi.org/10.1080/01944363.2020.1727766
Litman, T. (2018). *Autonomous vehicle implementation predictions* (pp. 1–39). Victoria, BC: Victoria Transport Policy Institute.
Litman, T. (2022). *Understanding transport demands and elasticities how prices and other factors affect travel behavior*. Victoria, BC: Victoria Transport Policy Institute.
Llorca, C., Kuehnel, N., & Moeckel, R. (2020). Agent-based integrated land use/transport models: A study on scale factors and transport model simulation intervals. *Procedia Computer Science*, *170*, 733–738. https://doi.org/10.1016/j.procs.2020.03.163
Llorca, C., Moreno, A., Ammar, G., & Moeckel, R. (2022). Impact of autonomous vehicles on household relocation: An agent-based simulation. *Cities*, *126*, 103692. https://doi.org/10.1016/j.cities.2022.103692
Marsden, G., & McDonald, N. C. (2019). Institutional issues in planning for more uncertain futures. *Transportation*, *46*(4), 1075–1092. https://doi.org/10.1007/s11116-017-9805-z
Mckay, M. D., Beckman, R. J., & Conover, W. J. (2000). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, *42*(1), 55–61.
McPhail, C., Maier, H. R., Westra, S., Kwakkel, J. H., & Van Der Linden, L. (2020). Impact of scenario selection on robustness. *Water Resources Research*, *56*(9), e2019WR026515. https://doi.org/10.1029/2019WR026515
Meyer, J., Becker, H., Bösch, P. M., & Axhausen, K. W. (2017). Autonomous vehicles: The next jump in accessibilities? *Research in Transportation Economics*, *62*, 80–91. https://doi.org/10.1016/j.retrec.2017.03.005
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: Journal of the Transportation Research Board*, *2673*(9), 407–418. https://doi.org/10.1177/0361198119844463
Moeckel, R. (2016). Constraints in household relocation: Modeling land-use/transport interactions that respect time and monetary budgets. *Journal of Transport and Land Use*, *10*(1), 211–228.
Papa, E., & Ferreira, A. (2018). Sustainable accessibility and the implementation of automated vehicles: Identifying critical decisions. *Urban Science*, *2*(5), 1–14.
Schwartz, P. (1991). *The art of the long view*. New York: Doubleday.
Soteropoulos, A., Berger, M., & Ciari, F. (2019). Impacts of automated vehicles on travel behaviour and land use: An international review of modelling studies. *Transport Reviews*, *39*(1), 29–49. https://doi.org/10.1080/01441647.2018.1523253
Sperling, D. (Ed.). (2018). *Three revolutions: Steering autonomous, shared, and electric vehicles to a better future*. Washington, DC: Island Press.
Steck, F., Kolarova, V., Bahamonde-Birke, F., Trommer, S., & Lenz, B. (2018). How autonomous driving may affect the value of travel time savings for commuting. *Transportation Research Record: Journal of the Transportation Research Board*, *2672*(46), 11–20. https://doi.org/10.1177/0361198118757980
Stephens, T. S., Gonder, J., Chen, Y., Lin, Z., Liu, C., & Gohlke, D. (2016). *Estimated bounds and important factors for fuel use and consumer costs of connected and automated vehicles* (NREL/TP--5400-67216, 1334242; p. NREL/TP--5400-67216, 1334242). Washington, DC: USDOE Office of Energy Efficiency and Renewable Energy (EERE), Vehicle Technologies Office (EE-3V) https://doi.org/10.2172/1334242

Swartz, P. G., & Zegras, P. C. (2013). Strategically robust urban planning? A demonstration of concept. *Environment and Planning B: Planning and Design*, *40*(5), 829–845. https://doi.org/10.1068/b38135

Tadayon, M., & Shemer, L. (2013). *The Maryland statewide transportation model. Model documentation (version 1.0)*. Baltimore, MD: SHA State Highway Administration.

Taiebat, M., Stolper, S., & Xu, M. (2019). Forecasting the impact of connected and automated vehicles on energy use: A microeconomic study of induced travel and energy rebound. *Applied Energy*, *247*, 297–308. https://doi.org/10.1016/j.apenergy.2019.03.174

Thakur, P., Kinghorn, R., & Grace, R. (2016). *Urban form and function in the autonomous era*. Melbourne, Australia: Australasian Transport Research Forum.

van Delden, H., McDonald, G., Shi, Y., Hurkens, J., & van Vliet, J. (2011). *Integrating socio-economic and land-use models to support urban and regional planning*. Paper presented at the AGILE conference, Aug. 7–13, Salt Lake City, UT.

Wack, P. (1985). How medium-term analysis illuminated the power of scenarios for Shell management. *Harvard Business Review*, *November-December*, 139–150.

Walker, W., Haasnoot, M., & Kwakkel, J. (2013). Adapt or perish: A review of planning approaches for adaptation under deep uncertainty. *Sustainability*, *5*(3), 955–979. https://doi.org/10.3390/su5030955

Wegener, M., & Fuerst, F. (2004). *Land-use transport interaction: State of the art.* Dortmund, Germany: University of Dortmund, Institute for Spatial Planning. https://doi.org/10.2139/ssrn.1434678

Wellik, T., & Kockelman, K. (2020). Anticipating land-use impacts of self-driving vehicles in the Austin, Texas, region. *Journal of Transport and Land Use*, *13*(1), 185–205. https://doi.org/10.5198/jtlu.2020.1717

White, R., Engelen, G., & Uljee, I. (2015). *Modeling cities and regions as complex systems: From theory to planning application*. Cambridge, MA: MIT Press.

Xiang, W.-N., & Clarke, K. C. (2003). The use of scenarios in land-use planning. *Environment and Planning B: Planning and Design*, *30*(6), 885–909. https://doi.org/10.1068/b2945

Zegras, C., & Rayle, L. (2012). Testing the rhetoric: An approach to assess scenario planning's role as a catalyst for urban policy integration. *Futures*, *44*(4), 303–318. https://doi.org/10.1016/j.futures.2011.10.013

Zegras, C., Sussman, J., & Conklin, C. (2004). Scenario planning for strategic regional transportation planning. *Journal of Urban Planning and Development*, *130*(1), 2–13. https://doi.org/10.1061/(ASCE)0733-9488(2004)130:1(2)

Zhang, W. (2017). *The interaction between land use and transportation in the era of shared autonomous vehicles: A simulation model* (PhD dissertation), Georgia Institute of Technology, Atlanta, GA.

Zhang, W., Guhathakurta, S., Fang, J., & Zhang, G. (2015). Exploring the impact of shared autonomous vehicles on urban parking demand: An agent-based simulation approach. *Sustainable Cities and Society*, *19*, 34–45. https://doi.org/10.1016/j.scs.2015.07.006

Zhang, W., & Wang, K. (2020). Parking futures: Shared automated vehicles and parking demand reduction trajectories in Atlanta. *Land Use Policy*, *91*, 103963. https://doi.org/10.1016/j.landusepol.2019.04.024

Zhong, H., Li, W., Burris, M. W., Talebpour, A., & Sinha, K. C. (2020). Will autonomous vehicles change auto commuters' value of travel time? *Transportation Research Part D: Transport and Environment*, *83*, 102303. https://doi.org/10.1016/j.trd.2020.102303