tise
transition, innovation and sustainability environments

Co-funded by the Erasmus+ Programme of the European Union

On City Digital Twins and Agent-Based Models: An illustration

Master's thesis for obtaining the academic degree

Master of Science

in the study program

Erasmus Mundus Joint Master Degree in Transition, Innovation and Sustainability Environments

submitted by

Alan Rafael Dill

Department for Knowledge and Communication Management

at University for Continuing Education Krems

1st Advisor: *Dr. Pablo Lucas (University College Dublin)*

2nd Advisor: *Gabriela Viale Pereira, PhD (University for Continuing Education Krems)*

Krems an der Donau, Austria, July 2024

Statutory Declaration

I, *Alan Rafael Dill*, hereby declare,

1. that I have written my Master's thesis myself, have not used other sources than the ones stated and moreover have not used any illegal tools or unfair means,

2. that I have not used my Master's thesis or parts thereof as an exam paper in my domestic or any foreign country in any form to this date,

3. that, in case my Master's thesis concerns my employer or any other external cooperation partner, I have fully informed them about title, form, and content of the Master's thesis and have his/her permission to include the data and information in my written work.

Alan Rafael Dill

Acknowledgements

What you are about to read is the final piece of a puzzle I've been trying to assemble during this last two years. It is the result of a process of reading, reflection, and (I hope) learning that, although it was materialized only in the last few months, started a while ago. Here you will have contact to which we can call the palpable object, the final outcome, of this time abroad. Because the feelings, the dreams, and the disillusions that pertain to, and are part of, this experience, understandably, yet unfortunately, don't make the lines of a dissertation. This acknowledgement, above all, tries to convey what is in the thesis, but it cannot be read.

Now, that I have the chance to start reflecting on this experience, I have the impression that I lived many years inside these last two. Studying for two years in four universities, in four different countries, is really something. Sometimes, the *something* was pleasant, but sometimes, this *something* was hard to deal with. It was two years away from the food, the weather, the smell, the noises I was so used to. So many things we take for granted until we step outside, and leave behind the place where we grew up. The beauty of it is, then, to be able to learn how to value these things, but also that you, too, can learn how to live without them. And certainly, this is more true for some parts than for others.

It was also two years apart from the people I had around me since I was aware of my existence. Luckily, I was blessed enough to live in a time when we can connect easily, I cannot imagine how hard it could have been if I didn't have this accessible channel of communication with my family. I'll always be grateful for their understanding, and their support. I can only imagine how hard it has been for them, too. To my mom, I'm grateful for always being with me, even though thousands of kilometers away. My most inspired writing would never make justice to the new meaning of *saudade*, I learn during these two years living so far away from you. To my dad, I'm grateful for being a source of positivity and energy. As for my brother, I thank him for not forgetting me, and I only hope I was of help when things were not that great.

Here, in this new chapter of my life, I am certainly thankful for having the chance of meeting extraordinary people during my masters. I made connections with different, brilliant minds, and only a few things will make me happier than knowing that wherever they go, they will find the success and happiness they deserve. I'm more than grateful for being able to have shared this time with my proud South American people. To Valenttina Cardozo, I'm thankful and glad I met you. Your dancing, singing, joyous soul, is an energy that never fails to warm my heart. If, in the near future, Colombia joins *Mercosul*, I'll spread rumors that we were behind the scenes, working for the agreement. To Elissa Tokusato, I'm thankful for the friendship and support through these two years. I'm glad we could share the task of representing Brazil, I could never do it properly alone, at least not with my dance moves.

Finally, I would like to thank my supervisors: Dr. Pablo Lucas, and Dr. Gabriela Viale Pereira, for always being a source of encouragement, and patient with my struggles. If this thesis carries any merit, I share it with them. Naturally, any imprecision or misjudgment written here are my own. I want to thank professors in TISE consortium; and by doing that, I extend my heartfelt gratitude for the professors that I came across during my life. I remain convinced about the importance of public and quality education, I would not be able to get to a postgraduate level without the opportunity to study. To the people all around the world that fight to remove barriers to education, thank you.

Abstract

City Digital Twins (CDTs) are expected to play an important role for the realization of Smart Cities. However, existing CDTs are still lacks appropriate representation of social systems. Agent-Based Models (ABMs), on the other hand, have a long tradition in studying complex social systems. Our research explores to what extent ABMs can benefit CDT applications and how this integration can lead to more robust models. We conduct this research by exploring an existing ABM within the context of urban mobility, TransportVarese. This model is an ABM that aims to mimic transportation mode preference in the Italian city of Varese. By introducing a dynamic fare policy for public transport as a test case, we were able to generate a proof of concept for ABM and CDT integration. The results corroborate both the potential and limitations of integrating ABM and CDTs. While the possibilities of integrating physical aspects and real-time data into ABMs are promising, the lack of access to a CDT limits the scope of our conclusions. On the CDT side, the relevance of a theory encompassing a detailed description of human interactions with the physical environment is suggested as a possible future development. Additionally, the discussion covers the ethical implications and the need for a sociotechnical, transdisciplinary approach.

Keywords: City Digital Twin (CDT), Agent-Based Model (ABM), Integration, Sociotechnical Systems, Transdisciplinary

Table of Contents

List of Figures

List of Tables

1. Introduction

Government's role as provisioner of public services encompasses a challenging task, which is to design interventions that are capable of tackling specific public policy problems. According to Dunn (2017), policy problems are characterized as interdependent, subjective, artificial and unstable. By interdependent, it means that they are hardly isolated, but rather layered and interconnected with other problems, and social issues. Their subjective characteristic implies that they are not perceived the same way by different social actors, which is tightly connected to their artificial nature. Here, artificial means that problems are not naturally "out there", they do not exist independently of human judgment, but an understanding of how things are, and a desire of how things should be. Lastly, policy problems are considered to present instability, they are most likely to be transformed rather than solved. Apart from that, policy problems can be manifested with different levels of complexity. Dunn (2017) suggests we can classify them as well, moderately or ill structured. Well-structured problems are the ones that have only one decision maker involved, and there is a consensus on their part regarding value preferences, the alternatives to tackle the problem are of a small number and the outcomes are somewhat certain, if not deterministic. Moderately structured problems, in turn, have a few decision makers, and value preferences need to be negotiated, and yet the outcome of the limited alternatives are uncertain. Ill structured problems, in opposition, have many decision makers, value preferences are in dispute, the alternatives, and outcomes are unknown, or highly risky. Many, if not most, of the policy problems that are more apparent to society are of ill-structured nature. Urban mobility, for instance, is one of them. It involves a multitude of actors with considerable stakes in the table, all of them with a saying, and competing interests, and many possible approaches, although with high uncertain, or unknown outcomes.

The complex nature of ill-structured problems, such the ones we find in urban mobility, are the ones that can benefit from investigation through the lenses of Complexity Science. In this sense, to look at public policies as entwined in complex systems, means to acknowledge that in such systems there are a multitude of heterogeneous agents, in which the interaction between agents and the environment can present non-linear outcomes. Contrary to Cartesian models, where a phenomenon can be reduced to be the sum of its parts, in complex systems the phenomenon is more and different than the sum of its parts (Furtado et al., 2015). This opens the possibility to advance complex system methodologies, namely modelling and simulation, to study public policies. Specifically, social simulation — often used as a synonym of Agent-Based Modeling — can be utilized as a policy tool to understand, evaluate and generate alternatives to policy problems (Furtado, 2022). Among the available simulation techniques, the ones generated with Agent-Based Model (ABM) methodology, are common and well established in the Social Sciences. ABMs are computational models that take agents as entities of interest, that interact between themselves and the environment, by following a set of rules - which can evolve or not during a simulation run (Wilensky & Rand, 2015). Such interaction, at the micro level, generates patterns that emerge from the simulation. The possibility for the modeler to tweak aspects and variables of the model, enable policymakers with the rare opportunity to test policy (for example, what if questions) *in silico*. The applicability to day-to-day policymaking is, nonetheless, hurdled by several aspects, many of them of technical nature. For instance, Malleson et al. (2022) highlights that implementing ABMs with real time data is still a methodological barrier.

The opportunity to create models capable of capturing every relevant aspect of the targeted real-life phenomenon is expected to be realized with the implementation of Digital Twins (DTs). A DT is also a simulation technique, mostly used and implemented in the engineering context. According to Thelen et al. (2022), DT's definitions have grown, as many contextual applications of the technology have been developed. In their review of the DT concept, a broad and widely accepted definition can be found in Glaessgen and Stargel's (2012) paper, mobilized in the context of aerospace engineering. There, DT "*is an integrated multiphysics, multiscale, probabilistic simulation of an as-built vehicle or system that uses the best available physical models, sensor updates, fleet history, etc., to mirror the life of its corresponding flying twin*" (Glaessgen & Stargel, 2012, p. 7). In recent years, the DT technology started to gain traction in another research context, Smart Cities. There, it promises to aid policymakers, and city managers, by virtualizing the entire city area. City Digital Twins (CDTs) are, in turn, virtual representations of the city, that can aid decision-making by taking advantage of CDT's simulation capabilities to assess different interventions to policy problems (Nochta et al., 2021).

This Master thesis aims to explore the relationship and potential complementarities between the two simulation techniques presented previously, Agent-Based Modeling, and CDTs in the context of urban mobility. In the next chapters, we will present the state of the art in research with DTs and ABMs. Then we will discuss the prospects of integrating the two technologies. We then dive into Urban Mobility studies, specially related to the use of simulation as methodology. The next step will be to present the Agent-Based Modelling as a methodology, in order to illustrate in the next chapter a model for understanding commuters choice on transportation mode, which will serve not only for us to discuss the model results itself, but also its possibilities to be incorporated on a CDT. Finally, considerations and comments on future work are provided.

2. Research Context and Problem Definition

In 1997, Robert Axelrod, wrote a paper in which he advised what, in his vision, would advance social simulation as a research method in social sciences. He wrote that in order to understand the value of simulation in social sciences, one has to see it as a new way of conducting scientific research. Simulation, he wrote, is a third way of doing science, distinct from the traditional deduction and induction approaches (Axelrod, 1997). But to grasp how valuable social simulation can be for critical stages of the policy process, we offer a close look at the evaluation stage of a public policy. The process of policy evaluation is distinguished between ex post and ex ante analysis. An *ex ante* evaluation refers to the analysis and methods employed to measure the impact and outcomes *before* the policy alternative is implemented. The *ex post* refers to the analysis and methods used to assess the outcomes and impacts *after* the policy alternative was implemented. The "ideal standard" for (ex-post) policy evaluation would arguably be quite similar to the so-called "gold standard" in clinical trials. With this approach, analysts are interested in observing the difference in the variable(s) of interest between two groups of individuals, one of them receiving the treatment (in this case, the policy), while the other (the so-called control group) don't. Since the two groups are assigned randomly, researchers can attribute, with a measurable degree of confidence, the measured differences found between the two groups to be the likely outcome of receiving the treatment. This is why such method is called Randomized Control Trial (RCT). One considerable difference, though, between the ideal standard in policy evaluation and the gold standard in clinical trials is the level of control attainable in each context. Unlike in the Natural or Mathematical Sciences, the treatment the subject can receive often vary. For instance, let's imagine that we are going to offer a training program for a given population residing in a small state in Central Europe. It is very difficult to ensure the attendants of the training session receive the exact same information, delivered with the same clarity and effectiveness. Which is to say, it is hard to ensure every participant of the treatment group will receive the same treatment. For the sake of comparison, a biologist can apply to a plant the same 2 grams of a fertilizer it applied for every other plant that are part of the experiment. When measuring results of the training policy, many techniques to assess the outcome will be highly subjective, such as self assessments, attitudes, and motivations. All variables we, as social scientists, are unable to directly measure. At the same time, the biologist, to stay on the example, can rely on very accurate scales of growth of a plant, or acidity of the soil. Obviously, this doesn't mean we can't successfully evaluate policies, in fact there are several strategies to minimize these hurdles. And yet policy evaluation is costly if not only financially, but political as well. Public officials have a disincentive to acknowledge that its administration invested in a policy with no impact (or with a negative impact). Finally, designing evaluation can also come with serious ethical implication. Specially when designing *ex ante* evaluations can mean that one share of the population will not benefit, at least right away, from the policy, while others will.

Simulation poses the possibility to advance policy evaluations in the digital realm, avoiding the costs and the moral dilemmas that may arise. Much has been advanced in this direction with the application of ABMs to policy research, but in practice, it is still rather restricted to academic research (Malleson et al., 2022). As for DTs, despite its established application in engineering projects, only recently the concept of CDTs has become clearly defined in terms of its potential, and usability to improve (Cai et al., 2022). The technological promise is evident: CDTs can offer a comprehensive control and understanding of a cities' infrastructure, and dynamics. Yet DTs in the context of cities are still rather attached to the virtual representation of physical aspects of the cities, while human interactions tend to be rather relegated to overly simplified representations (Caldarelli et al., 2023).

Thus, on one hand we have a well established methodology for policy research, dedicated to understanding social phenomena at the individual level, which so far has been mainly circumscribed to academic research. And on the other hand, there is technology capable of delivering data at a rich level of detail and in real-time, which still is rather far from generating models that can capture the interaction between the physical and living entities. Birks et al. (2020) refer to this as Societal Twins. For Mariani et al. (2022), from an engineering perspective, there is a clear benefit of adding agents' behavior and interactions within the context of DTs, via the synergy between having simulated agents using the enriched environment of DTs. This can enable simulated agents to have more robust means to simulate their decision-making process. Another benefit is that the simulated physical counterpart can also be used to experiment enhanced functionalities from the interaction with the simulated agents (Mariani et al., 2022). The conjoint use of real time data, say from a DT, and the interaction of simulated agents inhabiting a virtual representation of a real city is worth being advanced. Our contribution hereby, within this scientific effort, lies then in the exploration and critical discussion of the link between Agent-Based Modelling and DTs for cities. To proceed in this direction, the following research question is suggested:

To what extent, City Digital Twins could benefit from Agent-Based Modeling? And in another way, how the interaction between these two technologies can help to build more robust models?

ABMs and DTs are, as any other technology, tools to solve specific issues, in specific contexts. Thus, to use them, analysts are required to be able to identify these contexts and specific issues in which its application brings actual benefits, or better put, that the social and economic costs are reasonable and in line with the best interest of the stakeholders. Given our scientific effort is inserted at the public policy discussion level, ethics, justice, and social values are of the utmost importance, however, it is beyond our scope to provide a discussion on this aspect. Yet, public policy is a broad topic, and considering what was said on specific problems and contexts, we understand that a most humble and scientifically reasonable approach to conduct this Master dissertation is to position it in a research field. According to Malleson et al. (2022), urban analytics, a field of research that use different types of data, combined with computational methods to study urban process, has been applying Agent-Based Modelling to study urban dynamics, such crime and mobility. While CDTs simulates cities' phenomena, these are expected to deliver relevant insights for urban analytics purposes. Yet, the lack of direct access to a CDT does hinder progress of this research project, as designing a whole new DT from scratch would be beyond the scope of this thesis. Thus, the departing point to address the aforementioned research question is a simple model used to understand urban transportation, specifically commuter choices, called TransportVarese (Maggi & Vallino, 2021) in order to illustrate and explore the (a) applicability of Agent-Based Modelling in a DT environment, and (b) to discuss the importance of detailed agent behavior in CDTs.

Deepening our knowledge on the intersection of Agent-Based Modeling and CDTs would allow the further deployment of Smart Cities. CDTs are growing in number and scope, but are still mainly financed by governments: according to a World Economic Forum (WEF) report on CDTs, governmental investments represent 66.7% of the total amount invested in this kind of technology. This in light that the market size for DTs is expected to reach 48.2 billion US dollars by 2026 (Cai et al., 2022). Therefore, advancing its technological use can help attract more investment in the sector from different sources and ensure that the investments made by public and private sectors are more cost-effective. Although, there is a self-evident benefit of ensuring that the technology does aid policymakers to deliver better policies, issues of functionality, safety, and governance remain to be addressed. Improving our knowledge on DTs and their use with ABMs can help better understand their possibilities and limitations, which in turn can better inform the debate on governance and privacy around DTs. Finally, when Axelrod (1997) wrote , he was also concerned with the maturity needed for simulations to reach their true potential as a research method. In terms of public policy studies, the integration of ABMs and CDTs does seem a promising example of such realization.

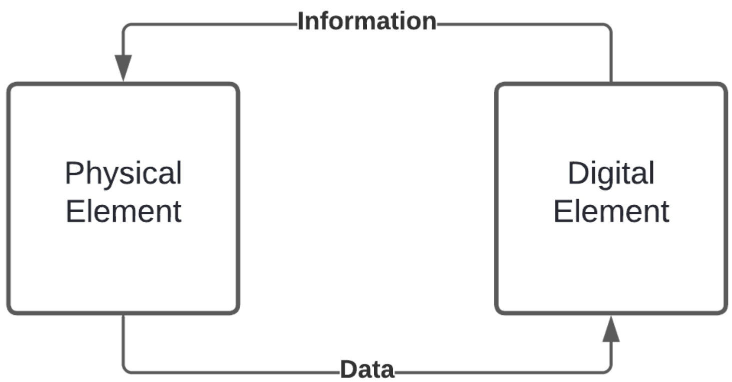
3. Theoretical background

3.1. Digital Twins

Michael Grieves (2023) tells that in the early 1970s, when he was working as a systems' programmer in a company, one of the salespeople there wondered if they couldn't help to solve a problem faced by the local telephone company, Michigan Bell. The issue consisted of landowners accidentally cutting through Bell's buried cable infrastructure, when digging their own property. The solution Michigan Bell found was to request landowners, before they would start digging, to call the company, who then send a crew to advise them about the cable location. A rather expensive solution that the salesperson who contacted Grieves thought could be best addressed by a computer. The idea Grieves came up with was to map the cable infrastructure digitally. Soon he realized that in order to carry out such a task would require computational capacity not available at the time. Yet, the idea of virtualizing physical objects persisted with him, and originated the DT model in the early 2000s, and yet the required computer power to fully run it was only met in the mid 2010s.

The rapid development of DTs technologies led also to different definitions proposed in the literature, as reviewed in Thelen et al. (2022), who offers a new definition adding more dimensions to the initial proposal by Michael Grieves. For the sake of our discussion, we opted to keep the most comprehensive and simple to illustrate. As shown in the Figure 1 below, the DT model proposed by Grieves (2014) encompasses (a) the physical element in the real space; (b) digital elements in the virtual space, and (c) data and information connections between the previous elements. Worth noticing is the bidirectional communication between digital and physical element, a necessity to the definition of a DT and differentiating characteristic from other modeling techniques (e.g., digital model, digital shadow), according to Thelen et al. (2022). Yet, Jones et al. (2020), in their review of DT publications, mention that the connection between virtual to physical elements is not always included in the DTs descriptions, despite it being present in the seminal definition made by Grieves. Notwithstanding, Jones et al. (2020) also understand the virtual to physical connection as fundamental in the DT definition, since it closes the loop and offers a benefit that no other previously existing model could offer. At the heart of it is the overall goal of trade information for resources in the real world (Grieves, 2023). And here it is important to keep in mind that the DTs model was created in the context of Product Lifecycle Management, and it was designed to match the creation, build, operate/sustain, and dispose stages in the engineering and manufacturing (Grieves, 2023).

Figure 1. The Digital Twin Model



Note: Adapted from Grieves (2014)

This is also represented in the classification of DT models, namely Digital Twin Prototype (DTP), Digital Twin Instance (DTI), and Digital Twin Aggregation (DTA). DTP is a Digital Twin born in the digital environment, there is no physical element (as in a physical prototype) to be explored of what the simulated object might be like in the physical space[1]. DTI is a Digital Twin which already is at the production phase, so the actual parts and its measurement exist only physically but digitally twined. DTA is then the aggregation of several instances in the Digital Twin environment, allowing for analysis of a population connected to the twined elements, forming the basis of the framework that can help with predicting failures, and evaluating performance.

As mentioned previously, the application of the DT concept, although still very much dominant, moved beyond the manufacturing domain. The literature review provided in Thelen et al. (2022) shows civil, mechanical, and energy as rising domain fields in which DTs are featuring. Do Amaral (2023), for instance, studied DT applications in all nodes of the energy supply chain (namely, generation, transmission/distribution, storage, and consumption). The authors found out that most of the studies are focused on the generation side of energy systems, mostly used to control these

[1] This is also a source of controversy: some understand Digital Twin as the mandatory existence of a physical element to be twined, Michael Grieves asserts that Digital Twins only need to be *intended* to be designed as a physical element. For him, it is key to understand that DT's main objective is to save resources in the physical environment, which is done by exploiting the informational resources in the digital environment.

systems. One example we found on the consumption side is best described by Arowoiya et al. (2024), who investigated DT applications in thermal comfort and energy consumptions in buildings, in which the technology is being used to manage, optimize and predict thermal comfort and energy usage. Another domain of research on DTs is centered around Healthcare Systems. In their systematic review in the field, Xames and Topcu (2024) found that the field is rapidly growing, yet very much attached to conceptualization, which also reflects the few studies on implementation. Nevertheless, DT applications were studied in the context of medical devices, patients health monitoring, disease prevention, and surgical assistance procedures. But also in the context of healthcare facilities, DTs were studied to improve staff management, and the specialized units, including preventive maintenance and emergency planning. Finally, linked to the concept of Smart Cities is the utilization of DTs to twin entire cities, which given our interest deserves a specific section to be discussed.

3.1.1. City Digital Twin

CDT is a concept often associated with the context of Smart Cities, precisely because DT capabilities adhere to the concept of smart cities, defined by Nam and Pardo (2011) as the intersection of human, institutional and technological factors. Here, CDTs act as a guide to smart city development, boosting sustainability and increasing life quality in the city (Shahat et al., 2021). Moreover, CDTs are expected to make cities smarter by providing the visualization of the key information on the city and a driver of innovation (Deren et al., 2021). Nonetheless, it is important to make some distinctions and to be more realistic with the concept and its opportunities and limits. For instance, the visualization feature that DTs offer should not be confused with other types of simulation techniques that are also very much attached to the visual aspect, and often part of the toolkit of city planners. Shahat et al. (2021) provide this distinction, discussing the characteristics of CDTs in comparison with Building Information Modeling (BIM) and 3D models. A 3D model is a digital visualization of a physical enterprise that can take the form of buildings, infrastructures, etc. BIM technology is a more robust 3D model, capable of handling information of all aspects of the simulated element. Data is, however, manually inserted, given there is no automated link between the physical and the digital. As we showed previously, the link between the two elements is a core definition of a DT, data is fed automatically to the model in real, or near real time, and there is a retrofitting link between the virtual to the real object. However, while the unprecedented amount of real time data gathered to replicate cities' physical twin via sensor and sophisticated IOT devices, Caldarelli et al. (2023) call attention to the fact that this does not guarantee a clear picture of a city. On the one hand, because not everything in a city is measurable and, on the other, because we still struggle to overcome problems with data, such as bias, accuracy, and related computational challenges. Yet, as the amount of data and its real-time availability is a key characteristic of CDTs, one has to take into account that more data to a model can, counterintuitively, reduce our explanatory capability of a given phenomenon. In fact, as Caldarelli et al. (2023) explain, it is not infrequent that simpler models yield better results. Batty (2018) also warns us that one has to keep in mind that a complete replica of the city is not possible, and not even desirable. A CDT, while a replica of the real entity, is an abstraction. If it was to be an exact copy, it would not be a representation of the system anymore, but the system itself. Leaving another issue of how to actually use a system so complicated as the one we try to act upon. Therefore, he argues, we cannot forget that DTs are still models that represent key aspects of reality while leaving behind many others. In the case of DTs of cities it must emulate not only physical objects such as buildings, trees, roads, but people and their complex interactions between themselves and the objects of the city. Also imbued in this discussion is the fact that CDTs implementations lack the representation of social systems (Batty, 2018; Caldarelli et al., 2023; Nochta et al., 2021; Shahat et al., 2021; Wan et al., 2023). In fact, Caldarelli et al. (2023) suggest that to model cities properly, we need to go beyond the physical assets, and actively include human interaction in their social, and economic aspects. One key aspect to it is to keep models informed by theory, as big data has not yet substituted the need for theory, according to Caldarelli et al. (2023), combining DTs with models informed by Complex Theory, can generate more trustworthy, calibrated and validated representations of systems, leading to deeper understanding of the phenomena of interest. Specially the models for urban policies, that must capture the evolutionary processes imbued in city developments, can benefit from a framework that takes into account bottom up mechanisms, such as the one supported by Complex Theory (Caldarelli et al., 2023).

Moving away from the technology driven discussion, Nochta et al. (2021) suggests a sociotechnical perspective on CDTs. Technology, they argue, is part of it, but if one wants to see the promised benefits of CDTs, one must envision a transdisciplinary and collaborative approach when designing a CDT. To design CDTs from a sociotechnical perspective means to ensure CDTs do not stay at the technical aspect, but that its design departs from policy problems structured from a practical approach, going beyond the stated overall policy objectives, taking into account the knowledge and the preferences from the local community, a process that requires a participatory design. In another token, to accommodate the contextual characteristics, such governance structures, must be considered when designing the CDT in order that the policies designed with the aid of the technology are actually possible to be effectively implemented. This also speaks to the often ignored costs, both organizational and individual, of learning and integrating CDTs in the policymaking process (Nochta et al., 2021).

In their review, Shahat et al. (2021) found Data Management, Visualization, Situational Awareness, Planning and Prediction, Integration and Collaboration to be the main themes in CDT studies. Data management is, in their findings, the kernel theme on CDTs, it refers to the challenge of integrating different types of data into city models. Next, Visualization stands out as a recurring theme and key feature of CDTs, and one of the main challenges is to visualize the dynamic of social systems. Situational Awareness is another theme identified in the authors review, and reflects on the great expectations of CDTs realization which is its ability to deliver a clear picture of city dynamics, which is closely related to the Planning and Prediction theme, given that not only analysts and managers can have an outlook on the city, but also that they can explore future scenarios and *what-if* questions. Integration and Collaboration is another theme covered, and it refers to the established challenge of integrating different models from the many aspects one might want to simulate. The authors mention that a direction found in the literature points to layering/coupling different models instead of having one "complete" model. In another way, this theme also relates to the possibility to engage different stakeholders in the policymaking. Virtualizing the important aspects of the city can help to democratize information access and bolster participatory practices.

So far, the application of DTs in urban contexts are still in their early steps. Reviewing studies on DTs, Masoumi et al. (2023) assessed the maturity level of CDTs, utilizing the Atkins maturity spectrum that measures the presence of five elements going from 0 to 5, in which 0 means the simple reality capture of the targeted physical object (the minimum functional DT) to 5, in which 3D modeling, and 2D mapping (element 1) are running along with persistent datasets and BIM technology (element 2), real-time data (element 3), bidirectional integration and interaction (element 4), and autonomous operations and maintenance (element 5). According to their classification, 90% of the CDTs reviewed are in the initial stage (52% in element 3, and 38% in element 2). In their overview of the challenges and trends in the field, the authors considered data types and sources, applied technologies and methods maturity and application. Regarding data types and sources, it was found that a majority of implementations use structured datasets, sourced mainly from official documents, followed by IoT sensors, but also including data from wearables and smartphones. Participatory and transdisciplinary methods of data gathering were rarely found in their review, which raise an important concern on legitimacy and governance for CDTs (see Helbing et al. (2023)). For the technologies applied, the authors found that Geographical Information System (GIS) was mainly used for spatial analysis and its integration with Building Information Modelling was rarely found. Since the sole use of GIS technology to simulate maps at the 3D level, the authors recommend that such integration of technologies must be advanced. The conclusion on maturity level of the CDT's that were analyzed is that more advancements are necessary for DTs to go beyond 3D modeling, and actually integrating humans and objects in the system. Finally, the authors classified the main fields of applications ranging from urban planning, transportation systems, to energy management, and natural disaster management, pointing to the necessity of more advancements in research on decision-making, policymaking, and evaluation.

3.2. Agent-Based Models

As we mentioned in the Introduction, ABMs are a representation of systems that take agents as units that interact between themselves and the environment, following a defined set of rules (Wilensky & Rand, 2015). Hence, as put by Klügl and Bazzan (2012), an ABM needs to address three elements, the agents, interactions, and the simulated environment. Agents have characteristics and are autonomous from other entities. The interactions are developed at the agent level and, given that the phenomenon of interest is derived from these interactions, it is central that they are explicitly modeled. The simulated environment refers to all the other elements to be modeled. It must not be confused with simulation infrastructure, where the simulation takes place, examples of simulation environments are NetLogo, Repast, or Swarm. In this sense, Klügl and Bazzan (2012) stress the distinction between the model and simulation characteristics of ABM, or ABMS, in their terminology. Simply put, while modelling is the creation of a model of a system, and simulation is the execution of the said model. An important feature that can be replicated, or generated with ABM, specially in the study of complex systems, is the emergence phenomenon, which is to say that the observed macro outcomes are derived from more than the sum of the interactions at the micro level (Wilensky & Rand, 2015).

This generative feature of ABMs is key when studying Social Science objects. Klein et al. (2018) argues that Agent-Based simulations in social sciences have distinguished advantages compared to other methodologies. In their view, ABM allows for understanding macro outcomes from the micro processes, this is often referred to as generative social science[2]. Besides, ABM has the capability to provide solutions to problems that are mathematically intractable due to the complex nature of the system it aims to represent. Contrary to traditional macroscopic approach that relies on mathematical description of the system, requiring users of the model to be well versed in the math language, ABM is more accessible and intuitive (Klügl & Bazzan, 2012). When comparing ABM with other types of simulation techniques, specially to other microscopic approaches to simulation, ABM distinguish itself by the high flexibility it offers when modeling agent behavior and its relationship with the environment, allowing modellers to control randomness and the model parameters that enable the discoveries of tipping points that not always are easily possible applying traditional methods (Klein et al., 2018; Klügl & Bazzan, 2012).

Certainly, as any approach, ABM also has its own shortcomings. Klügl and Bazzan (2012) comment that some of these methodological advantages can also incur drawbacks of ABMs. In this token, for instance, the higher degree of flexibility that it offers can also represent an additional challenge for inexperienced modellers. For instance, modelers have to choose what level of detail is going to be applied, and duly justify why some aspects are important to be detailed while others can be rather assumed to be of less significance. Given that there is no clear roadmap to build ABMs, in terms of design generating a model is often closer to an exploration process rather than the work of applying a protocol. In another sense, reproducibility of the models also comes as a weak spot, mostly related to the poor documentation of the models, something that the ABM community, though the application of the ODD (Overview, Design concepts, Details) protocol (Grimm et al., 2020), is trying to address. And, lastly, it is worth noticing that a good model is the one that appropriately mimics the real system, or, in other words, benchmarking how close the model is from the real system. This process is called validation and remains a challenge to both novices and experienced modellers. Collins et al. (2024) provide an overview of different techniques towards validation of ABM simulations. Above all, they stress the need to look into validation practices inside each field. And yet, validation also needs to be considered in terms of the purpose of the model. Edmonds et al. (2019) list 8 different types of modeling purposes; prediction, explanation, description, theoretical exploration, illustration, analogy, and social learning. In this token, one can

[2] Regarding the generative aspect of ABM, an interesting development is found in the paper *Inverse Generative Social Science: Backward to the Future*, by Epstein (2023) in which the author suggest departing from the macro behavior and, with the aid of artificial intelligence, arrive to the micro processes that originate the phenomena.

imagine that validation requirements for a model that aims to predict a system are higher, as one would expect from a model that wants to illustrate a system.

ABMs have been used to approach several topics in diverse disciplines. Makarov et al. (2022a, 2022b) developed an extensive review on the use of ABMs to study complex problems. The review covers the study of the spread of epidemics, pedestrian movement and evacuation, demography dynamics, transportation systems, geography and environment, land use, urban dynamics, reconstruction of history, Civil Conflict, Social Networks, and Economics. Particular to our interest is the study of Urban Mobility with the ABM methodology. Bastarianto et al. (2023) did a review of studies on urban transportation utilizing an ABM. Given that the research field typically involves simulating agents as travelers or vehicles, Urban Transportation Studies leverage the flexibility ABM provides, allowing for heterogeneous agent behavior, adaptability and learning. Their content analysis points to major themes in Urban Transportation, some of which include General Transport Modelling, Travel Behavior, Transport Policy. General Transportation Modelling was classified within the efforts of building theory and concepts through the implementation of micro and macro simulations to, for instance, analyze pedestrian and car movements. Travel Behavior describes the group of studies in which simulation driven by data is used to assess, as in Park et al. (2018), people's transport choices. Transport Policy theme reunites the corpus of research towards the creation and testing of policy scenarios

.

3.3. Integration of Agent-Based Models and Digital Twins

In this section we discuss the integration of the two technologies, ABMs and DTs, and its applications in CDTs. Yet, the literature we found is still tied with the concept of DTs, broadly speaking, and mostly related to the application with Multi-Agent Systems (MAS), which does not always translate to an ABM. Hence, it is no surprise that publications on the integration of these technologies are mainly focused on challenges and opportunities from broad applications in engineering, not in the application on CDTs per se.

One of the exceptions to this pattern is found in the work of Clemen et al. (2021). There, they explore the integration of a Multi Agent System and a DT of Hamburg traffic, in Germany. The project leveraged a sensor network in the city, coupled with an existing model of mobility deployed in a modelling platform called MARS. What is relevant for our discussion here is the methodological distinction they suggest to approach DTs and ABMs. Their proposition is to understand ABMs as

DTs, which departs from the conceptual definition of ABMs being comprised of Agents (A), as computational units that behave according to a set of rules and have agency in the environment, Entities (E), as computational units that can represent obstacles and barriers, and Layers (L), where Agents and Entities are placed. Therefore,

$$ABM = \{A, E, L\}$$

DTs, in turn, extend the definition of ABMs by incorporating Sensors (S), that can read the environment, and a Mapping (M) between the aggregation of Agent, Entity and Layer and Sensors. Hence,

$$DT = \{A, E, L, S, M\}$$

The proof of concept they implement and describe simulates the city of Hamburg, in which a population of agents is generated, each of these agents move from one point to another, a task that can be fulfilled by renting a bike in a bike rental station. In this sense, the ABMin MARS platform has agents (A) as a representation of people that commute around the city. These agents then decide to do it on foot, or utilizing a set of options, such as bicycles, or cars, the entities (E) of the system. An interaction between agents with other agents and entities is provided by the layers (L). The devices that provide an interface between bicycles in a bicycle rental station and the model are the sensors (S) that will be mapped (M) to the contextual entities, layers and agents. The implementation of such CDT is however very initial and describes a unidirectional flow of information between real and virtual world (i.e. modifications in the physical asset is messaged to the model, but modifications in the virtual asset doesn't affect the state of a physical asset). Among the challenges in the implementation, the authors highlight the complexity of dealing with synchronizing real time data in the simulation. Hence, the correction in the model provided by new up-to-date data was proven to work best at short term, since the risk of the simulated environment deviating from the real world increased as simulation time advances.

Challenges and opportunities of coupling DTs with Multi-Agent Systems (MAS), broad sense, are gaining traction in the literature. MAS is the umbrella field which ABMs are part of, and these two terms are often used interchangeably, although research with ABM methodology tends to focus on living beings, while MAS research this might not always be true (Malleson et al., 2022). Yet, the blurry distinction between MAS and DTs is also mentioned by Pretel et al. (2022) in their systematic literature review, in which they analyze the joint use of MAS and DTs. The similitude between DTs and MAS, according to the authors, can raise challenges to integrate the two technologies, although those challenges are not clearly pointed in the article. Their review shows that

DTs are being implemented along with MAS in two forms, mainly. On one hand, DTs are serving as databases of real time information about physical assets, a property of DTs called memorization, which in turn will be used by agents to make decisions. In another way, MAS are being implemented as the environment in which DTs will interact.

Barat et al. (2021) assert to have built a DT of the city of Pune in India, to test non-pharmaceutical interventions in the context of the COVID-19 pandemic. In this model, an ABMis developed to mimic the population characteristics and dynamic (i.e. age distribution, profession, movements and relationship between agents), places and points of contact of the city (offices, houses and apartments, vehicles) that represent the distribution of the different types of buildings in the city of Pune. Despite the authors claim to have a CDT of Pune, it is difficult to see the relationship with the concept of DT itself, there is no feedback mechanism between physical asset and the digital counterpart, and the representation of the city is rather abstract than a digital facsimile. To use the distinction posed by Clemen et al. (2021), the concept of Agent, Layer, and Entity is there, but there is no Sensor, and therefore no need for a Mapping between sensors and the other computational objects.

What seems to be clear is that the integration of Multi Agent Systems, such ABMs and DTs, is not only doable but beneficial. Mariani et al. (2022) explore the literature to offer a broad picture of the integration of MAS and DT's, illustrating how, from an engineering perspective, the implementation of DT's and MAS can be mutually beneficial, in a process they characterize as a *cross-fertilization* of the two technologies. Here, it is important to highlight that DT's exist, and are implemented, outside the MAS context. Apart from what we defined previously as DT, within MAS's context, the implementation of DTs leverages the concept of agents either by incorporating agents' feature of interaction with other computational entities, or by representing entities that interact with DT's.

Hence, what lies in the literature is that the integration is possible for one hand because there is a conceptual affinity between MAS, ABM and DTs, so closely related that it may be difficult to detach one from another when discussing an integration. In another way, the advancement of CDTs as a tool for policymaking will heavily rely on this capability of integrating human behavior in the system. This discussion is closely related to the necessity of models, and the role of theory when modeling a system of interest.

4. Methodology

In order to address the research question posed in Section 2, this research positions itself under the pragmatist worldview as described by Creswell & Creswell (2018). The pragmatist approach to research means a focus on the research problem and question rather than in specific methods. Hence, the statement that "*Pragmatism is not committed to any one system of philosophy and reality*"(Creswell & Creswell, 2018, p. 48), as one would find in constructivism, or post-positivism. Our interest is to explore the extent CDTs and ABMs can benefit from one another and generate better models. To our purpose, the pragmatism approach provides the flexibility to evolve the exploratory analysis we aim to do. Pragmatism is often embedded in mixed methods research designs. Although Creswell & Creswell (2018) classify case studies as methodology under the umbrella of qualitative studies, we are echoing the discussion on case studies made by Thomas (2011). In his definition, case studies are "*analysis of persons, events, decisions, periods, projects, policies, institutions, or other systems that are studied holistically by one or more methods.*" (Thomas, 2011, p. 513). Considering we do not have access to a CDT, we chose to approach the issue through the study of an ABM to address the research question. In the following sections, we are going to describe TransportVarese (Maggi & Vallino, 2017), an ABM designed to represent the transportation choices in the Italian town of Varese. In terms of the case study methodology we understand that the ideal scenario would be to develop an ABM within a CDT environment, or something very similar done in Clemen et al. (2021), where they managed to implement an ABM, connected to an infrastructure of sensors that allowed them to analyze the influx of shared bikes in a station. This is not possible at the moment. With TransportVarese, we will describe not only its functioning, but also we will implement an alternative test that can allow us to explore the research question at hand. Our understanding is that by studying the model and the test we propose, we will be "building a case", an illustration, which allows us to draw relevant conclusions regarding the proposed research question. The original model is provided in NetLogo, and open to modifications for scientific purposes. Our choice to use the original code instead of coding it from scratch is justified in the minor adjustments needed to the original script in order to implement the intended modification. And finally, we decided to study a model for urban mobility because for one, the transportation systems are a core characteristic of how cities develop, and a major factor considering urban planning (Wise et al., 2017), for the other we believe urban mobility is one of the fields where CDTs can most benefit from the integration of ABMs.

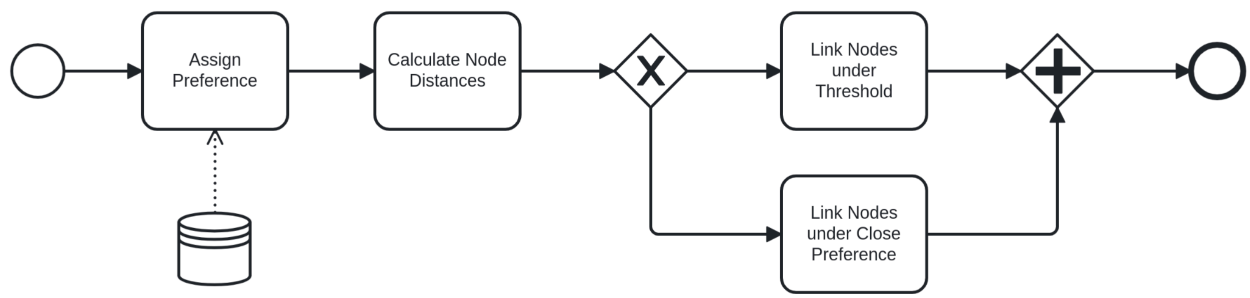
4.1. TransportVarese

One of the challenges for achieving smart cities is finding urban mobility solutions capable of alleviating the reliance on highly pollutant transportation modes. With the increasing growth of population, there will also be a higher demand for urban mobility. At the same time, cities are the main responsible for global CO_2 emissions (70%), in which urban transport plays an important role. Besides, cities are also a major responsible for particulate matter emission, which is directly linked to respiratory diseases (Bianchi Alves et al., 2023). To study the relationship between individual preferences and environmental impact, Maggi and Vallino (2021) present an ABMs that simulates passenger choices on urban transportation, and its impact on pollutant emissions. The authors build up from the MUSA model, developed by Natalini and Bravo (2014), aiming to perform an *ex-ante* evaluation of alternative policy and its effects on transportation choice in the United States of America context. Same as done in MUSA, TransportVarese test economic-based, and preference-based policies on transportation choices made by commuters. As individual choices have an impact on particulate matter emitted, the model is also able to demonstrate the impact of a policy or combination of both of them on emissions. Hence, their interest is to answer how personal choices on transportation mode affect emissions, and which policies are best suited to nudge individual decisions towards ecological options. However, differently than proposed by Natalini and Bravo (2014), TransportVarese only accounts for short trips, and generic enough to be applied to different urban contexts.

4.1.1. Setting Up TransportVarese

TransportVarese starts with a setup phase in which agents are assigned a preference for each of the transportation modes available. The transportation mode preference reflects the empirical data from 2011, concerning the city of Varese, provided by ISTAT (Istituto Nazionale di Statistica). The modellers assume that preference reveals subjective attitudes towards comfort and travel distance. The agent's transportation preference is also the key attribute to be used in the next task, which is to create social networks (see Figure 2). Agents are positioned in a virtual 3D graph that represents their preference for each transportation mode. Agents with close preferences are linked, respecting a threshold that is calculated to reproduce a power law distribution. For those agents that were not linked by the previous method, a new connection procedure is conducted solely based on preference similarity by on one of the transportation modes. Finally, the costs for each transportation mode are calculated.

Figure 2. **Social Network Creation**



Note: Adapted from Maggi and Vallino (2021)

4.1.2. Model Mechanisms

An agent chooses its transportation mode according to its *total satisfaction*, and *uncertainty* towards the transportation modes. Uncertainty takes into account the variation of *total satisfaction* on transportation modes over time, while *total satisfaction* is composed by a *personal* and *social satisfactions*. Hence, the *uncertainty*, U_t^i, is given by:

$$U_t^i = \sqrt{\left| N_{ik(t)} - N_{ik(t-1)} \right|},$$

where $N_{ik(t)}$ is the *total satisfaction* of agent i in regard to the transportation mode k in a time t. In turn, *total satisfaction*, N_{ik}, is computed as:

$$N_{ik} = \frac{\beta_i N_{ik}^s + \left(1-\beta_i\right) N_{ik}^p}{r_k},$$

where β_i is a randomly assigned value, ranging from 0 to 1, which serves as a weighting factor between *personal satisfaction*, (N_{ik}^p), and *social satisfaction*, (N_{ik}^s). While N_{ik}^p is given by the preference assigned from the empirical data. *Social satisfaction* is, in turn, given by:

$$N_{ik}^{s} = \frac{n_i^k}{n_i},$$

here, n_i^k is the number of agents using a given transportation mode in the agent's social network, while n_i, is the total amount of agents in this social network. Last but not least, in the total satisfaction equation, social and personal satisfaction are divided by r_k, a price index computing the relationship with the costs of a given transportation mode (k_{cost}) in terms of car costs (C_{cost}), as described below:

$$r_k = \frac{k_{cost}}{C_{cost}}$$

The natural consequence of this relationship is that the highest value the price index can get is for all the transportation means studied is 1, given that the index for private car, r_c, $\frac{C_{cost}}{C_{cost}} = 1$.

At the end of the computation, the agent performs a deliberation process that might keep or change their behavior towards its previous transportation choice. In the model, it is illustrated three combinations of factors in which the agent can change its transportation mode (*imitation*, *rational calculation*, and *social comparison*), and one in which it keeps the previous choice (*repetition*). The deliberation outcomes are a result of agent's *total satisfaction* and *uncertainty* levels. Hence, there are two thresholds that are observed here, U_{max}, as a maximum level of uncertainty, and N_{min}, as a minimum level of satisfaction. These two thresholds are defined by the modeller, and as in the case of TransportVarese, are used to calibrate the model in order to reproduce a very similar preference distribution as empirically measured.

Table 1. Deliberation Outcomes

N		U	**Deliberation Outcome**
$N_{ik} \geq N_{min}$	+	$U_{ik} > U_{max}$	Imitation
$N_{ik} > N_{min}$	+	$U_{ik} \leq U_{max}$	Rational Deliberation
$N_{ik} \geq N_{min}$	+	$U_{ik} \leq U_{max}$	Repetition
$N_{ik} < N_{min}$	+	$U_{ik} > U_{max}$	Social Comparison

Note: Maggi and Vallino (2021)

When the deliberation process falls into the Imitation outcome, it means that the agent is going to imitate the most common choice her neighbors made. Rational Deliberation means that the agent compute the satisfaction from each transportation mode and chooses the one that is most satisfactory. Repetition means that the agent will definitely not change its previous transportation mode, simply repeating its last choice. Social Comparison, in turn, checks the most popular transportation mode in its social network and compares it with its previous choice, deciding for the one that produces the highest satisfaction. Figure 3 provides an overview of the processes described so far.

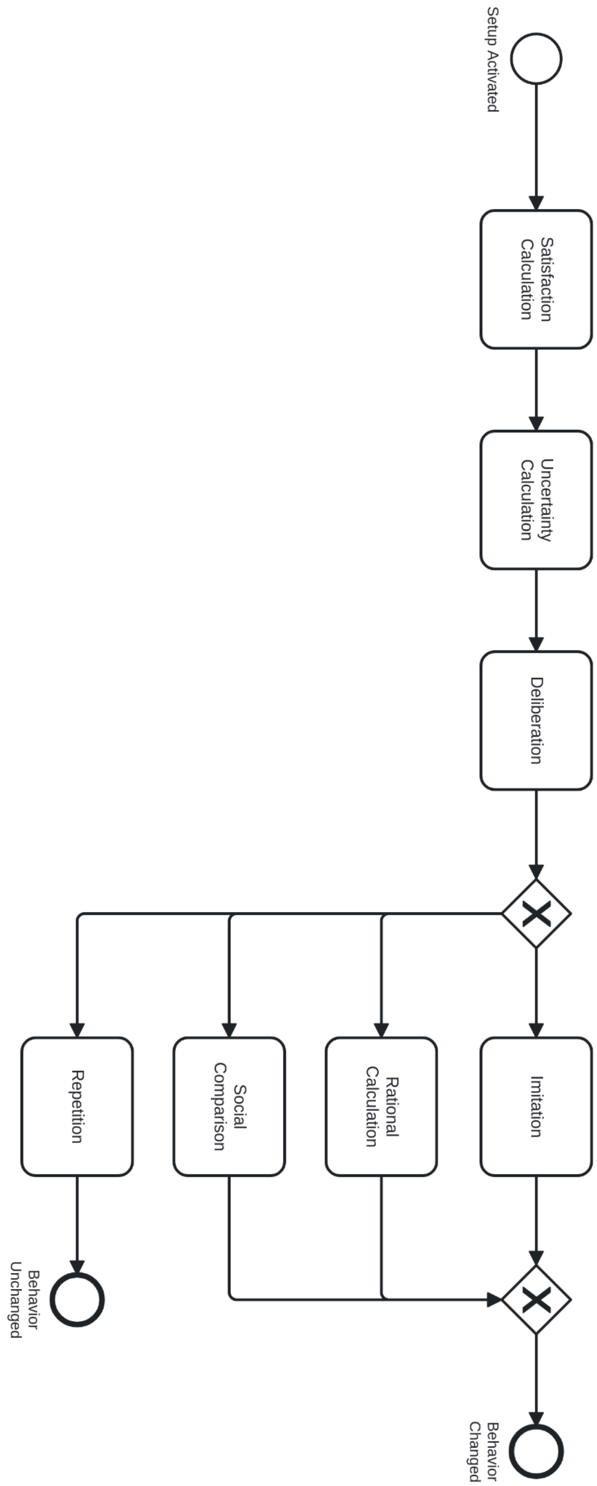
The model is said to have completed one unit of time (step) when all agents have decided on the transportation mode for the trip they have to make. The simulation goes on until it reaches what the authors call equilibrium, i.e., agents don't change their transportation choice. For TransportVarese, this equilibrium is known to be around 50 steps.

But TransportVarese also measures the impact agent decisions have on particulate matter emissions. The authors use empirical data to create an index (see Table 2) which calculates emission in relation to the most pollutant transport mode, the private car. They are described by the equations below:

$$C_{env} = 1 - \frac{C_{PM}}{C_{PM}} \qquad PT_{env} = 1 - \frac{PT_{PM}}{C_{PM}} \quad B_{env} = 1 - \frac{B_{PM}}{C_{PM}}$$

As we can see, the most pollutant transportation mode is given in a scale that varies between 1 and 0, being 0 the most pollutant and 1, the least. C_{env}, PT_{env}, and B_{env}, are respectively the environmental indexes of car, public transport and bicycle, while C_{PM}, PT_{PM}, and B_{PM} are the absolute values for particulate emissions.

Figure 3. Core Model Process



Note: Adapted from Maggi and Vallino

4.1.3. TransportVarese's Policies

There are two types of policies available for testing in TransportVarese, one of economic nature, affecting the costs of car usage, and another of preference nature, affecting the attitude towards transportation choices. According to Maggi and Vallino (2021), the price-based policy is assumed to incorporate and mimic policies that create economic disincentives to the use of highly pollutant means of transport, in this case the private car, examples are carbon taxes, road and park pricing, etc. This relationship is described in the following equation:

$$r_c^1 = r_c + \left(1 - C_{env}\right) * P_m$$

Here, r_c^1 is the updated value of r_c, the initial price index for cars, summed to the multiplication of P_m, the intensity of the pricing policy, that the modeller can vary from 0 to 1, and the difference between 1 and the environmental index of a car, C_{env}. Hence, if the policy is at its maximum intensity, $P_m = 1$, the price index for cars can double, given that $C_{env} = 0$, and $r_c = 1$.

The motivation based policies are policies that incorporate measures to reason with the population about green practices, and environment sustainability. It can include awareness campaigns, educative programs in schools, etc. In the model, the preference based policy also has an intensity level that can vary from 0 to 1, and affects directly the agent's initial preferences towards each transportation mode. In the expression below, x_i represents the initial preference for private car, y_i, the initial preference for public transportation, and z_i, the initial preference for bicycles, and P_q, the policy parameter that the modeller can vary from 0 to 1. As we can see, each variation in P_q, represent a decrease in agent's private car preference, and an increase of the same value in the preferences for public transport and bicycle.

$$x_i - P_q \qquad y_i + P_q \qquad z_i + P_q$$

4.2. Adding a new policy for TransportVarese

We propose to add a third policy to TransportVarese, and as our interest is to use the model as a case and discursive tool to answer the research question presented in Section 2, we are going to use most of the input data (see Table 2) and assumptions presented in the original model, only addressing the specificities of the added policy.

The model proposed by Maggi and Vallino (2017) shows that for the price based policy results are already satisfactory when the intensity applied revolves around 0.2 to 0.4, which translates to a shift in transportation choices, making, at 0.4 policy intensity, bicycles the most preferred transportation mode (57%), followed by public transport (25%), leaving private car the least preferred one (18%). Considering the initial preferences in Varese are 37% for bicycles, 19% for public transport, and 44% for private cars, the outcome is quite impressive. The simulation also reveals that, in theory, the preference for private car can be dropped to almost zero, if the policy intensity is set to the maximum. However, the authors remember, such a scenario is politically unfeasible, not to mention it does have an unjust impact across the population, leaving only those who can't afford to pay for the increasing costs of using private cars, the burden of greening the city. Preference-based policy positions itself away from the economic aspect, acting on preference through education and reasoning. This policy performs even better than the last one. At an intensity of 0.4, the preference based policy is able to shift the preference for bicycles to 73% of the agents, while declining the preference for cars to nearly zero. Interestingly, the preference for public transport remained steady at 19%. What is interesting here is not only the overwhelming shift towards bicycles but also the little to no change in behavior towards public transport. Maggi and Vallino (2017) argue that in the case of price policy this can be explained by the agent's initial preference being already skewed towards bicycles, hence, when prices hike for private car, the shift to bicycles prevails. For the motivation-based policy, the result is even more peculiar if we look at the (nonexistent) shift to public transportation. Agent's preference towards public transport remains unchanged, even with the intensity of the preference policy increasing up to 0.4. And, again, it is worth recall that with 0.4 intensity the simulation shows that is a level good enough to almost erase the preference for private car. The authors argue that this counterintuitive behavior can be explained by the persistent relevance of financial costs in agents decision-making, leading to agents that face increased preference for bicycle and public transport, still deciding for bikes.

The prevalence of an accentuated shift to bikes and the little adherence to public transport calls our attention for two main reasons. First, it seems far-fetched that in different contexts, such policies can realistically overwhelmingly push commuters to bicycles. In a context where, for example, bike lanes are unsafe or violence and robbery is frequent, the incentive to take bikes may be arguably smaller. Besides, a sudden influx of bikes riders also creates other types of friction in the city, and not every city has the infrastructure ready to adjust to the new scenario. In the case which the shift to bicycles does not happen, one can expect a change of commuter behavior towards public transport, pressuring the public transport system, creating yet another type of problem, that may lead to shortages and crowded buses, metros, and trams. Second, public transport is still a solution to car centric cities, where walking or cycling is not feasible at all. Hence, if TransportVarese is correctly simulating commuting behavior one might argue that price policy and preference policy may not be enough, and can actually reinforce inequalities in urban mobility, generating other dysfunctions in the transportation system. To address the lack of incentive towards public transport as shown by the simulated data for the case of Varese, we propose to add a third policy test. Travel Demand Management is a strategy used to reorient patterns of use a given transportation system has, such strategies allow transportation providers to better allocate resources, by inducing commuters to use the transport system in different times of the day, releasing the demand pressure in peak times, and by extension improving comfort for users (Halvorsen et al., 2016). One of these strategies is the deployment of dynamic pricing (or fare), a technique that takes into account supply and demand concepts in price formation of a given service, and many studies have concentrated efforts in defining appropriate methods to calculate the dynamic fares (Saharan et al., 2020). Our effort is humbler, given the importance the economic variable has on agent's decision-making, we wonder if such policy is also capable of inducing commuters to shift their preference towards public transport when prices vary. Therefore, for the additional test we are proposing, we want agents to see a variable price index for public transport during the simulation run. Thus, we expect to be able to test if price variation is able to nudge commuters choices towards public transport. Our hypothesis is that this addition to the model may help us to discuss the research question, after all, in order to implement such policy it is necessary that the entity taking decisions on price fares, is able to see a clear picture of buses, tram, and metro demand and occupancy, in real time. In this sense, a DT that represents the transportation system flows is of the utmost importance to calibrate the pricing policy of public transportation. Although idea of managing the transportation system through flexible pricing is not new, but derived from demand management studies in transportation systems (Halvorsen et al., 2016; Peer et al., 2016; Saharan et al., 2020), here, the idea of dynamic pricing serves us in two purposes, from one hand, it allows testing the sensibility of the model to price changes for public transportation and its implication on commuter preferences, and by doing so providing an illustration of how a CDT, in this case focused on urban transportation, and ABMs can work alongside each other.

As one can imagine, the core of the policy is to present different fares according to travel times. It means that commuters that travel in specific times, are offered different fares. Commuters traveling in peak hours are surcharged, and commuters traveling in off-peak hours are rewarded with fare reductions. We suspect that at the same time it creates an incentive for people to change their commuting habits, it can help to attract new adopters of public transport. However, to fully test this policy in the TransportVarese is not really possible. First, given that the original model is conceptualized to represent a day, and not hours in a day, we cannot set the policy at the granular level, e.g. defining the hours in which the surcharge and discounts take place. Second, following this limitation of the model, we cannot be incrementally reducing or increasing fares depending on the amount of commuters in a given time. The original model also would not account for supply of public transport. All of these relevant aspects are left out of the modeling, but the core concept of offering agents different fare prices can be modelled.

The implementation is carried out in two moments. First, at the setup process, there is a step to check whether the modeler is setting the dynamic fare policy on. In case the policy is in place, all agents register a price index for public transport. To the best of our knowledge, there is no clear figure on what proportion of commuters travel in peak hours, therefore we assume that 50% of the commuters travel during peak and off-peak hours. Thus, we modeled that half of the population commuting see an increase in public transport fares, while the other half, a decrease. Both the increase and decrease is in the order of 30% of the standard fare, this proportion follows the findings of a study for the case in Sydney, in Australia (Douglas et al., 2011) where a similar policy was evaluated. Hence, the model works for that half of the commuting population is able to see an increased price, and the other half a decreased price. When the time comes to calculate each agent's satisfaction with each transportation mode, it uses the reduced/surcharged fare. In the table below, we summarize all the input data and variables utilized to run the model. In the next section, we will present the results of the simulation. Each set of scenarios, or variable modification, was run 5,000 times and their results averaged and displayed at one decimal point, without rounding.

Table 2. **Variable Values and Description**

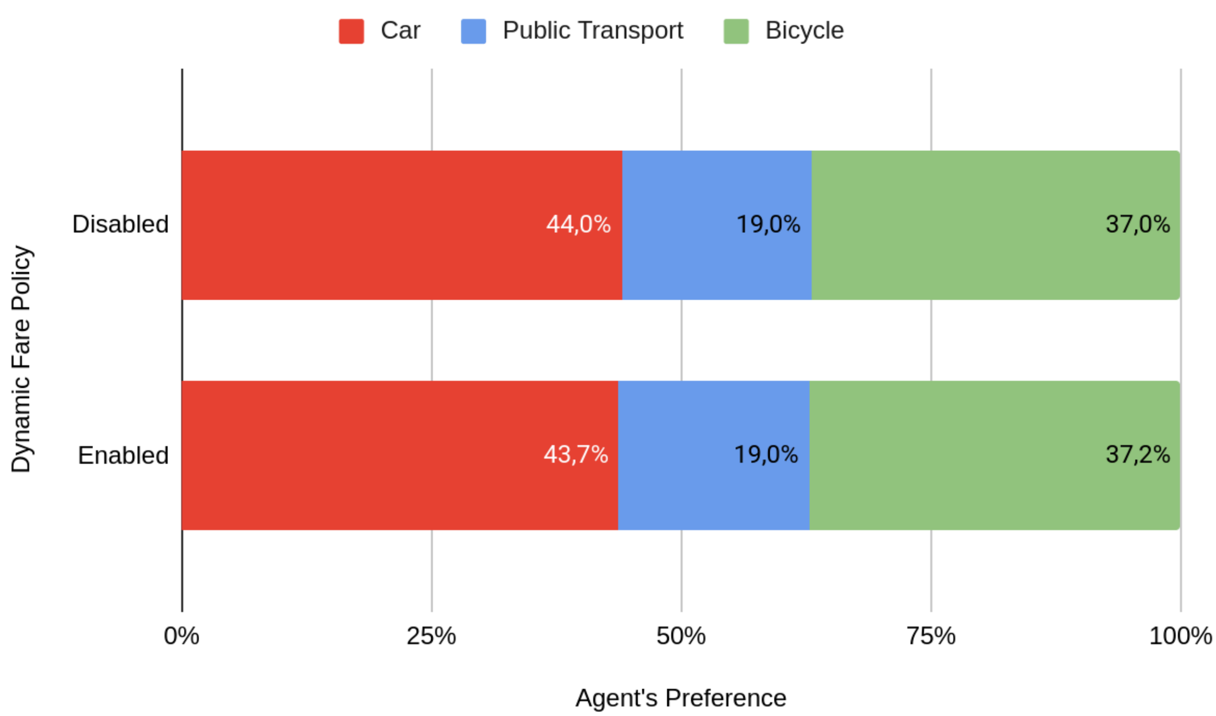
Variable	Value	Source	Description
Price Index — Car	1		A price index that varies from 0, to 1, having car costs as value reference.
Price Index — Public Transport	0,26		A price index that varies from 0, to 1, considering the average ticket price and having car costs as value reference.
Price Index — Bicycle	0,13		A price index that varies from 0, to 1. Bike costs are estimated to be 13% of car costs.
Environmental Index — Car	0		An Environmental Index, having car's emission as value reference. The Environmental Index for cars is 0, the most pollutant, and considers emissions weighted by car type.
Environmental Index — Public Transport	0,44	Maggi and Vallino (2021)	An Environmental Index, having car's emission as value reference, see equation x. The Environmental Index for public transport is calculated taking into account vehicle type and average utilized capacity.
Environmental Index — Bicycle	1		An Environmental Index, having car's emission as value reference, see equation x. In terms of emissions, considering only the activity of riding a bike is zero, and therefore the Environmental Index is at the maximum.
Uncertainty (maximum)	0,4		The uncertainty threshold represents the maximum value an agent can tolerate, and The total Satisfaction threshold represents the minimum value an agent accept, before considering changing her transportation mode. Values selected to calibrate TransportVarese to mimic the mobility preference of Varese city
Total Satisfaction (minimum)	0,7		
Link Distance	0,5		Variables used to create the social network and mimic a power law link distribution
Random Position	5		
Fare Rate	0,3	Douglas et al. (2011)	Rate of discount and surcharge applied to public transport fares, when dynamic fare policy is enabled

5. Results

5.1. The Dynamic Fare Policy

The first test we conduct with the new policy, and as any of the other scenarios we present here, set the dynamic fare policy to 0.3, meaning a surcharge and discount of 30% in the price of public transportation being shown to some agents, as described in the previous section. The primary intent with this test is to visualize whether the dynamic fare alone is already enough to nudge agents towards public transportation, or not. The results are displayed in the bar chart below.

Figure 4. TransportVarese on Dynamic Policy



Note: Baseline results (Disabled) were extracted from Maggi and Vallino (2021).

As illustrated in Figure 4, there is no change in preference towards Public Transport, and only a slight tradeoff between Cars and Bicycles. In fact, as Maggi and Vallino (2021) do not disclose how many runs they did to arrive in their final results, nor if how they rounded them, one might attribute the change we see to be small enough to be considered a rounding difference. At first, it can strike as surprising that almost no change is seen in agent behavior towards public transport, but if we recall their argumentation to explain why the price policy has little effect on public transport, they mention that agents changing their transportation mode will, most likely, change to bikes simply because it is the second most preferred transportation mode at the start of simulation, and besides, bikes are the cheapest transportation mode compared to any other simulated option (Maggi & Vallino, 2021). Also, let's not forget that the price of public transportation not only decreases, but also increases in the same proportion. We were hoping that a small change in price could be enough to start a cascading effect within the social networks to create a new pattern towards public transportation. This led us to wondering if there is any threshold capable of changing an agent's behavior towards public transport. In the graph below, we plot increasing fare rates and the resulting preferences after 5000 runs.

Figure 5. **TransportVarese preferences with different Dynamic Policy rates**

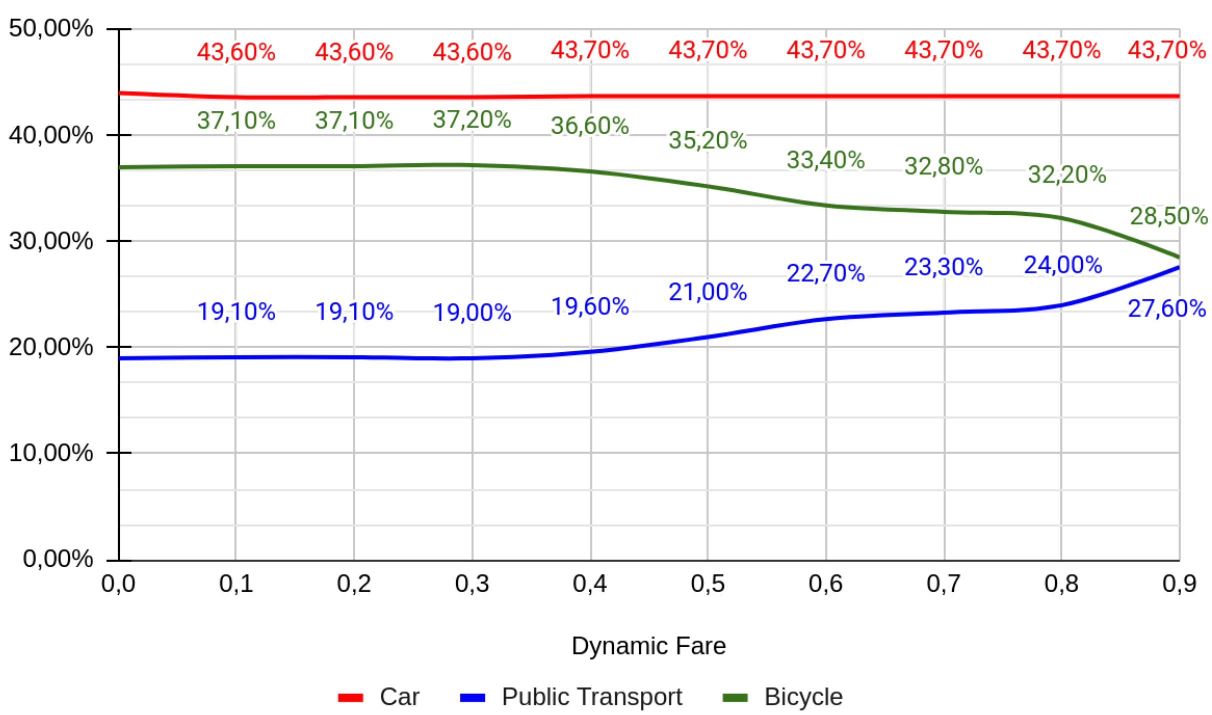


Figure 5, shows results do not vary sensibly when the dynamic fare is adjusted from 0.1, to 0.4. From this threshold onwards, the preference for public transportation start to slowly increase, representing an increase of 3.1% at a discount/surcharge rate of 0.6. From 0.6 to 0.8 the preference curve for public gets flattened, but faces an accentuated increase at 0.9. At this last rate, is capable of almost equalizing the number of public transportation users, with bike riders. Surprisingly, the preference for car remains unaltered, making public transport and bicycle preference curves to mirror each other.

If instead of the average, we plot all 5,000 runs for each policy rate, we would be able to see how preference is distributed. That's what we did in the violin plots below. To better visualize them, we separated the plots by transportation mode. Violin plots are useful visualization tools to assess distribution and central tendency on data, they are composed by a mirrored density curve (wider parts, are the ones where we found more data points), and a box plot at the center, which helps us identify the median value. These plots are convenient here to compare how these features change over different parameters (in our case, the dynamic fare rate).

Figure 6. Violin Plot for Car Preference with different Dynamic Policy rates

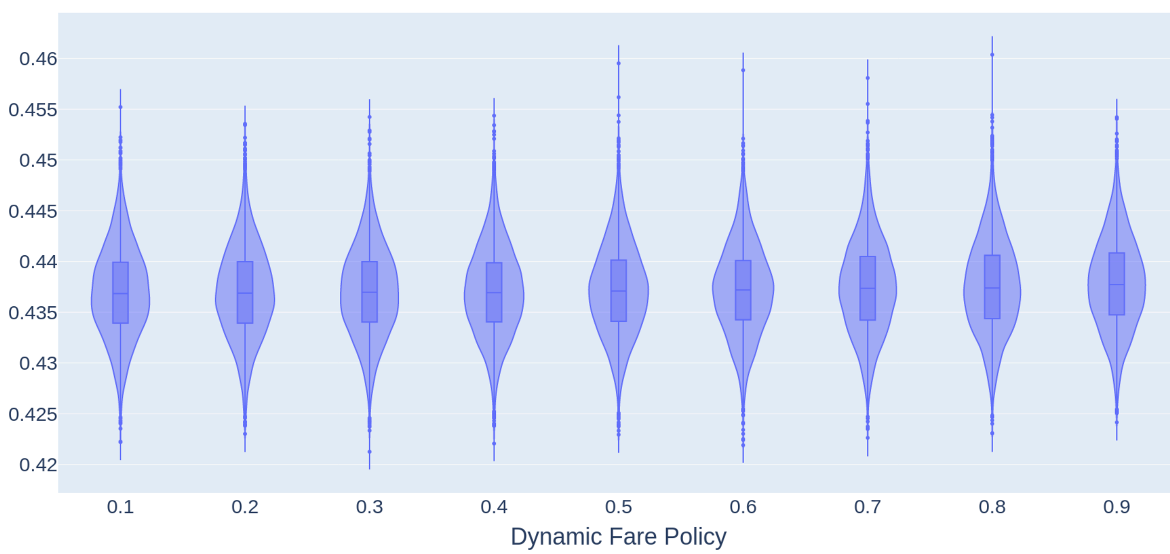


Figure 6 shows a violin plot for dynamic fares ranging from 0.1, to 0.9. As we saw in figure 5, the average preference remains stable at 43.7%. And here we see that throughout all simulated fare rates, simulation results are uniformly distributed around the median. When we plot the same graph, but for bicycles (Figure 7), we note that as the fare rates increase, not only the preference for bicycle start to plummet, but the simulation results become more scattered, and the range of values, wider.

Figure 7. Violin Plot for Bicycle Preference with different Dynamic Policy rates

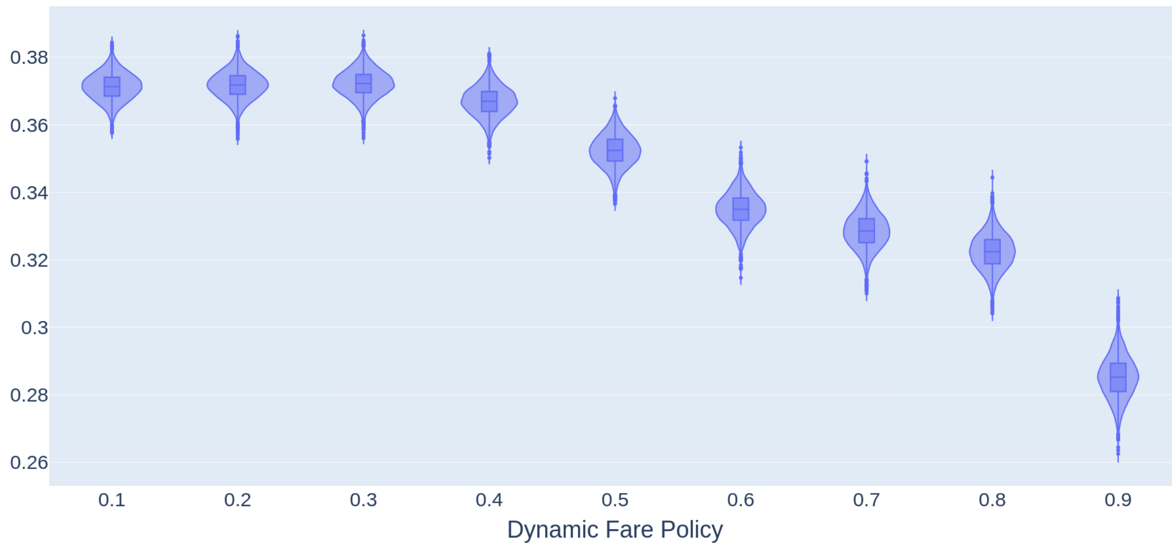
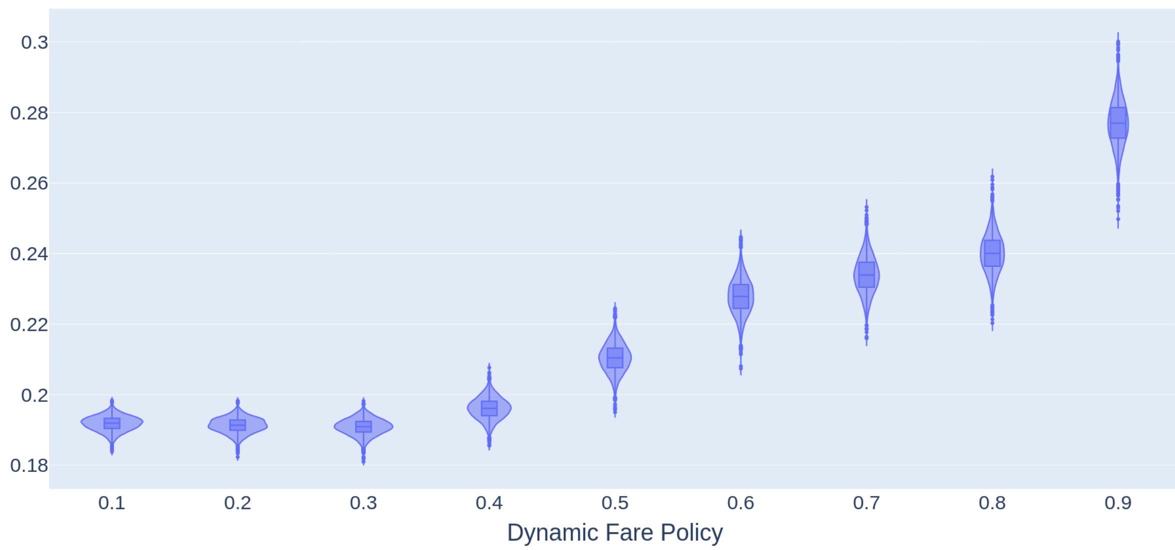


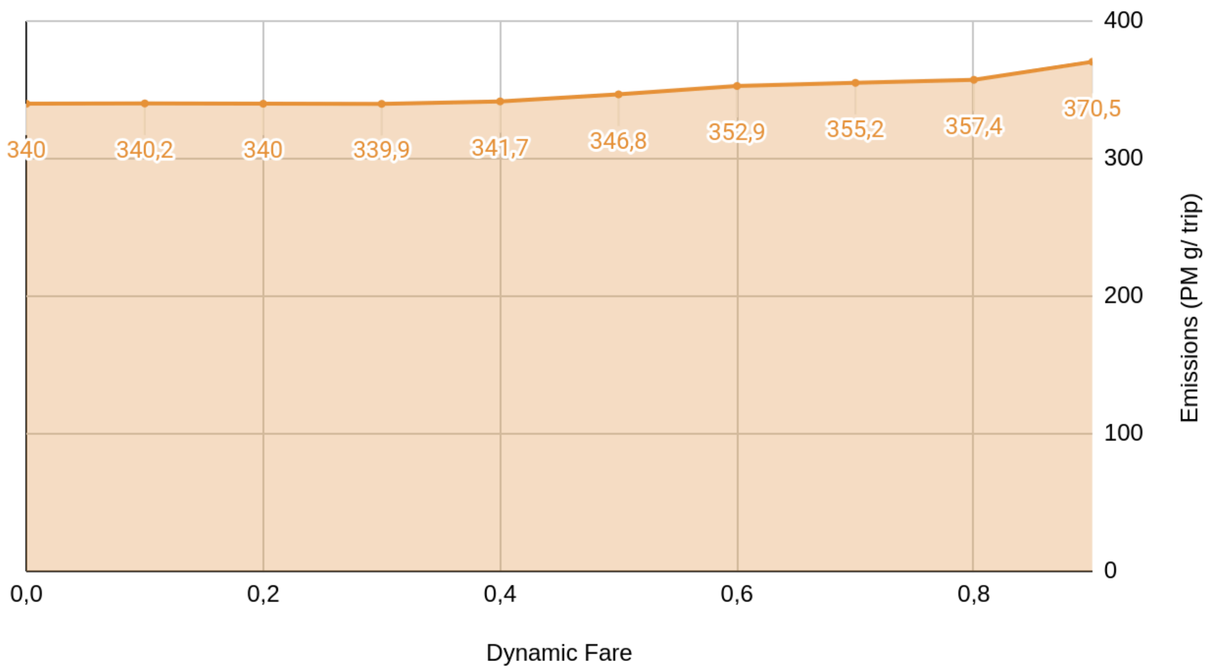
Figure 8. **Violin Plot for Public Transport Preference with different Dynamic Policy rates**



What we pointed out for Figure 7, we can see even clearer in Figure 8. Note that from 0.4 rate, it is not only the starting point for the increase on public transport adoption, but also the tipping point for distribution shapes. From 0.5 to 0.9, the median preference for public transportation grows, but the width of the violin (more uniform), indicates that simulated preferences are less concentrated around the median, also corroborated by the increasing range of values. The key takeaway here is that as we increase the fare rate, both preference distributions, for public transport and bicycle, present more variability.

The mirror effect we see in the graph, clearly shows the dynamic fare policy alone, assuming the initial Varese preference, is insufficient to shift car users to greener options. The phenomenon has a clear implication for the environment, given that a growing number of agents decide by bus instead of bikes, but nearly no users change from car to bus or bike. The graph below shows what the dynamic fare policy alone seems to cause to the estimated particulate matter emission. An increase in dynamic fare policy intensity can lead to an increase of almost 9% of emissions.

Figure 9. TransportVarese emissions with different Dynamic Policy rates



In the next two sections, we single test each original TransportVarese policy along with our newly implemented dynamic fare policy. Finally, we test the five scenarios tested in the original paper, and compare the results with the dynamic fare policy enacted. All the following tests depart from the policy intensities suggested in Maggi and Vallino (2021).

5.2. The Dynamic Fare Policy and the Price Policy

To test the Dynamic Fare Policy, is to combine two different economic policies at the same time. However, they do not shadow each other, given for one hand price policy affects solely prices for car use, and, for the other, the dynamic fare only affects bus prices and yet the highest price is compensated by the lowest price, proportionally. In the previous section, we saw that the dynamic fare policy failed by itself to nudge car commuters to public transportation. Here we expect that with the increasing car costs, commuters that drive cars will have a bigger incentive to adopt public transport.

Table 3. Price Policy and Dynamic Fare Policy

Dynamic Fare Policy	Price Policy	Preference Policy	Emissions (PM g/ trip)	Car	Public Transport	Bicycle
Disabled	0,1	***	325	**42,0%**	**19,5%**	38,5%
Enabled	0,1	***	328,7	41,6%	19,4%	**38,8%**
Disabled	0,2	***	310	**40,0%**	**20,0%**	40,0%
Enabled	0,2	***	318,4	39,8%	19,8%	**40,3%**
Disabled	0,8	***	140	9,0%	**28,0%**	62,0%
Enabled	0,8	***	154,6	**9,40%**	27,2%	**63,2%**

Note: The highest proportion between the test (Enabled) and baseline (Disabled) are in bold. Baseline results were extracted from Maggi and Vallino (2021).

However, it was not the case either. When we compare the values between dynamic fare policy enabled and disabled, we note no accentuated difference. Commuters leaving car as transportation mode keep going with bikes as a second choice and, if we look closely, although very slightly, the proportion of public transport adopters is higher when the policy is not in place.

5.3. The Dynamic Fare Policy and the Preference Policy

We saw in section 4.1.3 that the preference policy interferes with the personal preference agents have towards all the simulated transportation means. It adds up to the agent's preference for public transport and bicycle, while subtracting from the agent's preference for car at the magnitude of the policy intensity. Considering the results we have so far, it is unlikely that a proportional change in preference towards public transport and bicycles will be enough to switch car users to public transport.

Table 4. **Preference Policy and Dynamic Fare Policy**

Dynamic Fare Policy	Price Policy	Preference Policy	Emissions (PM g/ trip)	Car	Public Transport	Bicycle
Disabled	***	0,1	320	41,0%	19,0%	38,0%
Enabled	***	0,1	328,4	**41,9%**	19,0%	**39,0%**
Disabled	***	0,2	310	37,0%	**19,0%**	41,0%
Enabled	***	0,2	300,8	**37,9%**	18,1%	**43,8%**
Disabled	***	0,4	110	8,0%	**18,0%**	73,0%
Enabled	***	0,4	95,1	**8,30%**	12,2%	**79,4%**

Note: The highest proportion between the test (Enabled) and baseline (Disabled) are in bold. Baseline results were extracted from Maggi and Vallino (2021).

And as shown in the table below, our intuition was right. However, we notice that what dynamic policy was able to do was to sway commuters to bicycles in a more accentuated fashion than observed in the baseline. This result is interesting because in the original paper from Maggi and Vallino (2021), the sole intensity variation in preference policy barely affect public transport choice, and here, using the same parameters and with the dynamic fare policy activated we observe a change of a considerable magnitude as we increment intensity in association with. The explanation to the phenomenon in the paper was that agents that change from car to another means of transportation mostly decide for bike, given it is the cheapest option. In our case, we suspect that a share of agents that were commuting by bus and perceived a price increase in the transportation fare, had an extra incentive (the increase in preference for bikes) to change to bicycles as well.

5.4. Mixing Policies

Finally, we present the five scenarios created in Maggi and Vallino (2021), adding the dynamic fare policy to them. We are interested in any difference the model can yield, following the combination of different measures suggested, plus the dynamic fare policy. As the authors made evident in the original model, price based and preference based policies work best together. In their case, a price policy, which increases cost for car users, needs to be coupled with a preference policy to incentivize and create awareness on green mobility, because by itself it needs a considerable intensity in order to have an effect on transportation choice. Furthermore, increasing costs to car users certainly has limits, it is quite unpopular, thus, politically unfeasible. Conversely, the cost benefit appeal such policies have, can also be moderated by non-economic factors (such social status, community practices, etc.), all aspects that can be treated within the scope of a preference based policy. And certainly preference policies face their own limitations, a study on crime prevention showed that public awareness works best when paired with other policy instruments (McGuire et al., 2021). The use of different policy instruments can help to alleviate the downsides of a single solution, while highlighting the potentials each initiative can bring to tackle the policy problem. Also being a type of economic policy, the dynamic fare policy is a pricing strategy that can also benefit from concurrent policy implementation. In the table below, we summarize a series of tests combining the price policy, and preference policy, along with the Dynamic Fare Policy.

Table 5. Policy combinations

Scenario	Dynamic Fare Policy	Price Policy	Preference Policy	Emissions (PM g/ trip)	Car	Public Transport	Bicycle
I	Disabled	***	***	340	**44,0%**	19,0%	37,0%
	Enabled	***	***	339,9	43,7%	19,0%	**37,20%**
II	Disabled	0,1	0,1	310	37,0%	19,0%	41,0%
	Enabled	0,1	0,1	311,5	**38,9%**	**19,5%**	**41,5%**
III	Disabled	0,2	0,2	120	6,0%	**24,0%**	69,0%
	Enabled	0,2	0,2	116,6	**6,7%**	21,2%	**71,9%**
IV	Disabled	0,8	0,2	70	**1,0%**	**23,0%**	74,0%
	Enabled	0,8	0,2	79,8	0,50%	21,9%	**77,5%**
V	Disabled	0,2	0,4	70	**2,0%**	**19,0%**	78,0%
	Enabled	0,2	0,4	53,7	1,4%	12,7%	**85,8%**

Note: The highest proportion between the test (Enabled) and baseline (Disabled) are in bold. Baseline results were extracted from Maggi and Vallino (2021).

What we see in the table above is just a confirmation of what was shown in the previous tests. Variations from the baseline are marginal, as with policy intensity at 0.1, for both the price and preference policy. As soon as we increase both to 0.2 we start to see an increase in preference for bicycles. The difference from the baseline, as we observed in the previous tests, is however more accentuated for bicycles. The policy seems to be more effective to convince public transport users to go for bikes, instead of convincing car users to leave their vehicles and adopt mass transport.

In the next section, we try to understand these results in light of what we have been describing so far.

6. Discussion

We started this thesis emphasizing the applicability of Complexity Science to study public policies. Essentially, the advantage of the methodological toolkit that the Complexity framework can offer for policymakers. A scientific paradigm best suited to study the systems in which public policies are trying to act on. Such systems can rarely be predicted, as a physical system can normally be, but, with the aid of Complexity approach, can certainly be better understood. It is in this context that simulations are suggested as a useful tool. DTs and ABMs are both simulation techniques, but with different trajectories when it comes to policy studies. As we noted in section 3, while ABMs have an extensive trajectory in policy research, DT is a technology that has been mostly deployed in the engineering context, where the application is embedded in more predictable, and stable systems. Meanwhile, its application to emulate city environments holds a promise that is yet to be fulfilled. In the engineering context, DTs mostly deal with variables that can be directly measured and controlled, and data collection is digital and mostly automated. Contrary to its application in physical systems, CDTs has to process several aspects of the city that goes beyond the physicality of the city, such social and economic process, in which data is not always digital or automated (Batty, 2024). Hence, when we want to discuss the possibility of these two technologies complementing each other, we are assuming that there is something that distinguishes them both. However, as we presented in section 3.3, the concept of DT and ABMs can get mixed in the literature. One side of the issue is that there are similarities between ABMs and CDTs. As we pointed out in section 3.2, ABM is flexible enough to support different types of model entities, with various degrees of autonomy; and although it may not heavily rely on the physical representation as DTs would, its application integrated with GIS is far from uncommon. The key characteristic here is that DTs are expected to have a two-way real-time communication between the real and the digital (see Figure 1). Changes in the real twin impact the digital, and changes in the DT impact the real. This is not the case for ABMs that have regardless of purpose aims to represent the real system with some degree of fidelity. When we presented the results from Masoumi et al. (2023) in section 3.1.1, that measured the maturity level of CDTs following the Atkins maturity scale showing that most of the implementations of CDTs falls under “element 0”, representing a model that only performs reality capture, we can assume that a CDT, at its very initial stage can, in theory, be an ABM. We say this in theory because reality capture can mean different things. If we think only in terms of physical aspects, a satisfactory representation of roads and buildings might be enough to fall into this category. And probably that was the meaning inferred to reality capture, especially if we are aware that the addition of social and economic processes to DTs is a pressing challenge (Batty, 2024; Caldarelli et al., 2023). Hence, CDTs can benefit from ABM to the extent which serves to appropriately represent these processes, and give meaningful insights to policymakers about urban complex systems. Our work with TransportVarese does contribute in this direction, highlighting what CDTs and ABMs can do separately but also where these may integrate.

What is missing for integrating CDT and ABM

Without access to a CDT, or, a CDT that already implements an ABM incorporating social processes, working from an ABM seemed to be the appropriate first step to examine our research question. The test implemented in TransportVarese was a sufficient proof-of-concept to provide insights into the discussion regarding CDTs and ABMs, as it allowed us to infer the possibilities if we were to have a CDT coupled with TransportVarese. One evident advantage of having a CDT, is that we would have been able to move from a more illustrative model, as defined by Edmonds (2019), to one more representative of the real system, by adding more variables and real data to it. In terms of policy evaluation, experimenting with ABM models with more and fine-grained variables, means replacing or refining the assumptions modelers have proposed. Such a possibility potentially improves the theory sustaining the models, while also giving policymakers the ability to test different policy instruments, having a complex systems approach. In the case of TransportVarese many factors that can influence commuters to either take a bike or a bus are assumed to be of less relevance, or simplified. Comfort, for instance, is not explicitly modeled, but understood to be part of the preference component, which is to say that opinions on comfort are revealed in the act of choosing one transportation over the others. There is no need for a thorough analysis to see that this is an oversimplification: e.g., comfort may not even be part of agent's decision-making. It might be that for a certain context the comfort offered by public transport is above average and still other factors are weighing in to sway commuters from taking a bus or train, such as prices, safety, or cultural traits. CDTs can offer is an unprecedented variety and volume of data. But, equally important, as ABM has so far been mostly restricted to academic research (Malleson et al., 2022), CDTs can offer ABMs a chance to fulfill its promise as policy tools. In fact, this is already a reality in some domains. For instance, a private company, OpenSapce[3] has been implementing models of crowd dynamics to DTs of transportation hubs, such as in St-Pancras station, in London, UK (Rail

[3] https://open-space.io/

Safety and Standards Board, 2020). It is important to mention that developing a DT can happen either if we already have the real system we want to model, or if we just want to design a system before implementing the real twin. The classification of DTs suggests that a DT Prototype is the digital model before the simulated element exists physically (Grieves, 2023). The rationale is to save time and resources by developing the initial model as much as possible in the digital environment, before producing the physical artifact. For public policies, this is of great value, if we can design interventions *in silico*, as many times need to adjust and test scenarios, we can actually deploy policies only when stakeholders are comfortable with the simulated outcome. ABMs can help with this. While an aerospace engineer can test a physics model in a turbine prototype, a policy analyst can deploy as many models available of agent behavior interacting with, to stay in the urban mobility context, buses, trams, bike lanes, etc. Therefore, while the prototype of a DT in the engineering context is the soon-to-be physical object, in the public policy context, the prototype is the emergence pattern generated by the interaction between agents and the simulated environment, which forms a complex systems' environment. Besides, CDTs can speed up the process of finding solutions for ABMs and data assimilation, given the fact that there is no established mechanism to incorporate real time data into the models (Malleson et al., 2020, 2022). Thus, one can expect that one major opportunity CDTs might face by adding ABMs to represent social and economic processes is that by doing so it will also inherit the challenge to incorporate real time data to it. Nonetheless, the incorporation of real, or near real-time data to ABM has been shown not only possible, but a must to improve models. Preiss et al. (2022) adapted a model to predict hospitalization during COVID-19, which was able to increase its accuracy with the addition of near real time data.

What we can say about the designed tests

Transportation is about overcoming challenges of available space and its physical constraints (Rodrigue, 2024). In TransportVarese, the only physical aspect mentioned is distance. Contrary to the original model, MUSA, TransportVarese only considers short trips in the model. By short, it means travel distances no longer than 10 kilometers, and shorter than 15 minutes. And yet the only implication is that all the simulated trips are of short distance. The absence of physical aspects in the model have some non-trivial implications, which we already touched in the previous results section of this thesis. Some examples include: absence of modeled factors of non-economic nature to explain bike adoption, and low acceptance of public transport. Also, the impossibility to check how different transportation preference is spread through the different parts of the city. Complementary to this discussion, Winters et al. (2011) conducted a survey in the context of metropolitan Vancouver, Canada, to investigate factors that influence bike travels. In their analysis, they found strong motivators and deterrence related to physical aspects of the city. Factors such road safety, terrain flatness, pavement quality, illumination, and bike lanes location are significant elements to be observed by city planners when designing policies and cycle routes. In another, more recent study, O'Reilly et al. (2024) analyzed a survey conducted in 31 countries and associated the responses with the respondent's geospatial data. Here, it is worth mentioning that the authors were also interested in seasonal differences on bike adoption, another physical aspect that comes to play in transportation systems (Rodrigue, 2024). Although O'Reilly et al. (2024) found out that demographics (gender and income) and sociocultural aspects (pro environmental attitude) to be the strongest predictors on deciding to bike, physical aspects such living in elevated locations, presented less tendency to ride a bike, while elevated temperatures in summer also a played as a deterrent in user's decision to bike. Besides, they were able to show the correlation between infrastructure and bike adoption, yet they warn policymakers that these patterns are not universal across countries and had to be accompanied by educational (or, to stay with the terminology used in TransportVarese, preference-based) policies that create a culture around biking as a transportation mean. Both studies provide empirical evidence that corroborate the importance of physical aspects -of both built infrastructure and the natural environment- on the user's decision-making to travel by bike. Considering the strong results towards bike adoption in TransportVarese, it is of highest importance to be able to model these features. Furthermore, even if the goal of the model is solely to measure transportation mode preference, one might expect some dynamic preference changes over time. Especially when we start adding more physical aspects to the model. For instance, if TransportVarese is right about the overwhelming bicycle adoption after price and motivation policies are enabled, one can expect that either the infrastructure would be designed to accommodate the sharp increase in bike transit, or, on the contrary, this shift may be moderated by the hurdles imposed by the environment that were not appropriately accounted for. Many if not all of these aspects and variables are hard-to-capture only using ABMs. Although possible, there is a significant technical effort to incorporate such information accurately. It makes better sense to connect the ABM with a CDT, benefiting a model such as TransportVarese by providing a platform that incorporates the accurate physicality of the city. And here we are echoing the vision exposed in Mariani et al. (2022), in which agents of Multi Agent Systems interact with DTs to update their information and adjust their decision-making. This extra power of integrating an ABM with a CDT can provide a whole new range of testing possibilities to policymakers, allowing them to experiment different sets of policy instruments before actually choosing which one to be implemented in the real world.

In the test we proposed, the suitability of a CDT for TransportVarese can become even more evident, Dynamic Pricing depends on the information on supply and demand, thus, heavily reliant on data. The examples we found, however, are using data of past flows of commuting to generate fixed discounts and surcharges. A CDT, in this case, would, for example, account for the number of buses circulating, the occupancy rate, the number of passengers waiting in bus stops, in real time, and could potentially offer different fare rates in real-time, adjusting itself even based on traffic flow, road blockages (accidents, demonstrations, safety hazards, etc.). This diversified data we could extract from a CDT could also help us explain why in TransportVarese, bicycles keep being preferred to public transport, and why public transport cannot attract car users. In the context of our tests, we can only speculate that the dynamic fare policy might not be working accurately in this model, because it fails to incorporate any other aspect that can influence future decisions. Also, we have to remember that the original aim of Transport Demand Management policies is to reduce the stress put on public and private transportation systems . Particularly during peak hours, with the intention to make the system run more efficiently with the resources it already has at its disposal and with minimal intervention.

Although we found no study that directly tried to establish the relationship between dynamic fare and increases in public transport adoption, as we mentioned in the methodology, we expected that price variation would be the catalyst of pattern change through a social network of commuters. However, we were only partially right. When applied alone, the dynamic price policy showed little to no effect in transportation choice. This was the case when policy intensity ranged from 0.1 to 0.4 (see Fig. 5). Thus, applying the dynamic fare policy in the model, at a standard rate of 0.3 discount/surcharge, as we found in the literature, showed no effect in mode preference. But, when the 0.4 threshold is surpassed, the pattern does change. The preference for public transportation increases, but at expense of bike adoption, and not car users. Besides, policy intensity means the proportion of discount and surcharge on fares, in other words, an intensity of 0.8 represent a surcharge and discount of 80%. Interestingly, applying the dynamic fare policy in the model, at a standard rate of 0.3 discount/surcharge, together with price and preference policies (see Table 5), resulted in the opposite of what we observed with the dynamic fare policy applied alone, causing a reduced public transport adoption. We see that this phenomenon is most noticeable in scenarios with a predominance of preference policy rather than with price policy (see Tables 3 and 4). We can conclude that price variation may be not enough to attract new public transport adopters, however we remind the reader that any other dynamic aspects that can influence decision-making (e.g., bus accessibility, road quality, fleet age, etc.) is constant in the model, tied to the initial perception commuters have for each transportation mode.

The test proposed of a model that falls under the abstract typology is suggested by Squazzoni (2012) and pointed out by Tolk et al. (2022) that, in order to advance to simulation based policy support, we need more than good examples: robust and complete theories of agent behavior are required. Which also resonates with Caldarelli et al. (2023) comment that big data did not replace the need for theory. Specially for CDTs, they argue, there is a gap in "*representing the complete set of relevant interactions between physical assets, processes, and systems*" (Caldarelli et al., 2023, p. 14). In the literature, it seems clear that the utility of a theory is mainly that of an explanatory ability of the agent's behavior. Thus, we suspect that a Middle Range Theory (MRT) could be developed and implemented in the computational environment of a complex system such as the case study of this thesis, benefiting the development of more robust and applicable ABM models integrated with CDTs. MRT is a theoretical development made by the sociologist Robert Merton, which aims to develop an explanatory approach to an interrelated system of logical assertions, broad enough to generate explanations for a particular social phenomenon. The empiricism related to the MRT would help with creating hypotheses by analyzing concrete observations of a social context, so that a MRT would stand for an "*observation to be theoretically informed and theory to be empirically grounded*" in the social sciences (Mills et al., 2010, p. 559). Merton proposed a paradigm for analysis which one should follow to develop an MRT, guiding the researcher to build an MRT that can be used to fit and test assumptions of a theory and so can add further explanatory power to that theory. The thesis thus concludes then just before the starting point towards the development of a testable middle range theory approach to the integration of ABMs and CDTs. This could be one of the future works stemming from this research project.

To what extent the research question has been addressed with the ABM

This study is limited to the extent of the test we were able to implement in TransportVarese, thus the contribution it offers to answering the research question is strongly attached to it. The discussion we provided here is tangential to the two-way relationship between CDTs and ABMs. However, given the limited time and lack of access to a CDT, we were able to illustrate how ABMs can benefit from CDTs more thoroughly than we were able to describe the extent in which ABMs can benefit CDTs. We tried to distance ourselves from a pure technical description on how the integration of the two technologies can happen, although, in our Theoretical Background section, we did provide a glimpse of the technical hurdles it faces, specially the issue with real-time data being incorporated in ABMs. As posed by Mariani et al. (2022), to advance the integration of the different layers of DTs and Multi-Agent Systems there is a need to work on the existing platforms, in both fields, in order to identify the functionalities and models that can be extended. Our work did not aim to accomplish such a task, but certainly contributes to this direction by offering an illustration using an ABM to show how a CDT could be integrated to a simulation model involving agent behaviors in an urban mobility context. And yet the discussion, and applicability, of such integration were constrained to the urban transportation domain. Not every simulation of social process needs to incorporate a CDT, neither every CDT would benefit from incorporating an ABM model. The reason is that models can have different purposes, so it is possible that results found in another modeling context may offer different insights that may — or not — justify integrating an ABM with a DT. In other words, some scenarios may present no clear advantage of implementing a model where agents interactions are mediated by physical objects. There can also be scenarios where there is no clear advantage of simulating physical objects and its interaction with agents. Thus, the integration of these two technologies and research methods depend on each case. Our illustrative policy test serves as a proof of concept to aid the discussion and applicability of how ABMs could be integrated in a DT environment. For this specific purpose a model of Urban Mobility is an appropriate scenario, given the relationship of space, and its physical features.

What it could mean to have ABM and CDTs integrated for policymaking

Regarding policymaking, there are some potential societal benefits and challenges that remain to be solved. For instance, there are specific hurdles such as: ethical and governance aspects exclusively linked to having a CDT connected to an ABM. While it is beyond the scope of this thesis, we find it appropriate to briefly discuss this topic. According to Anzola et al. (2022), ethical concerns of computational simulations have been identified, but not fully developed in terms of its implications. With the continuous advancements of such models and its implications in how we act upon its outputs, there is a reasonable imperative to not only identify these ethical issues, but to actively minimize and control risks, both from the research as for the policy perspective. Ethical issues cross through the entire life cycle of designing a simulation and some core aspects of simulation, e.g., the interdisciplinary (i.e., different epistemic, and attitudes towards what is ethical), and the technology dependence (i.e., opacity of models, lack of access to high performance technology) can also raise specific ethical questions (Anzola et al., 2022). Although the authors were discussing ethics looking at Agent-Based Social Simulations, CDTs, as a simulation technology, can also be permeated by most of the same ethical issues. Writing on DTs potentials and limitations for smart cities, Helbing and Argota Sánchez-Vaquerizo (2022) highlight significant issues related to data privacy and mass surveillance due to extensive data collection, not to mention the perils of societies being subject to technocrat system of control that inhibits participation, instead of fostering it. Or worse, authoritarian regimes where the government leaves no choice to its citizens but to be intrusively monitored. Adding ABMs to CDTs only brings another layer of privacy concerns: who (public and/or private sector(s)?) and how it should be managed. Given that some of these models would no longer serve as a scientific exploration, as we plan to use these models to aid the management of cities and its environment, one must address questions related to: a) how representative these models are?; b) how accurate is their data?; c) how we should use the output of these models to make decisions?; and d) how can we explain the model outcomes?

Most prominently, as demonstrated with the tests in TransportVarese, the potential of having a DT capable of representing the real city context can be invaluable for testing and selecting policymaking alternatives. With TransportVarese, which is a rather simple model, we were still able to test three different types of policies *in silico*. As our objective was to use the model as a support to a discussion on the possibilities, and limitations, of integrating that to a DT that mimics the urban mobility of Varese. This opportunity was then about looking critically at the model, highlighting its limitations, and the many simplifications and assumptions behind its conceptualization, before it connects to a DT. For instance, we questioned how the increasing preference for bicycle use could generate other types of dysfunctions in the city. The idea that a price policy and preference policy can act as a catalyst of green transportation is an option, among many others — which could only be tested within the ABM. This illustrates the role we have as policy analysts when designing a model for a policy problem. Batty (2024) uses the expression"human in the loop", to denote the role of the scientist, deciding on what variables are important, what perspective it incorporates, and what are left out in CDT. When using CDTs and ABMs as tools for policy modeling, who designs the system has an idea of how the problem is structured (i.e., what the problem really is, and what should be tackled), implying a set of assumptions and predetermined values from the system designer. In the case of TransportVarese, the policies implemented in the original model aiming to green the city are directed against car use, implying a value judgment towards it, and, therefore, a correction of trajectory through the implementation of a policy that penalize its use. A different approach to the same problem would be to invest in alternatives to car use, expanding public transport services, modernizing the fleet, etc. Then the question a CDT integrated with an ABM could help to answer would be of route optimization, and frequency of transportation services. Framing the problem this way, one of the underlying assumption is that the problem of urban transportation is linked to a poor supply of mass transportation in the city. Yet another option to sway people from highly pollutant transportation modes would be to subsidize electric cars, and provide the appropriate infrastructure of Electrical Vehicle (EV) chargers (EVC). Inviting considerations that could be tested using a CDT integrated with an ABM that could answer question like: how far should charging stations be from each other? What are the environmental aspects influencing the adoption of a particular EV and EVC? Note that here the problem no longer poses a matter of access to mass transportation, but the type of private vehicle that citizens should be incentivized to purchase, and the type of infrastructure that should be put in place. Many other research questions could be designed and tested prior to real-world implementation. All of these policy alternatives have intrinsic value preferences, and an idealistic view of how the system should be, what is just, and desirable. Although these different forms of structuring the problem may all point to the overall objective of reducing emissions in the city, they may have different outcomes, and are probably appealing to different segments of society. ABMs cannot address, by itself, the "human in the loop" issue pointed out in CDTs, but can certainly be of help to test and communicate these assumptions in a way that creates a shared understanding of what the problem actually is and what are the implications of tackling it in different ways.

As posed by Nochta et al. (2021), in order to build legitimacy and applicability of such solutions, practitioners need to also consider CDTs from a sociotechnical perspective, which implies to analyze the social dynamics between technology and social actors. This perspective, they argue, points to the need for interdisciplinary, collaborative approaches. But we can go beyond and assert that if models really aim to generate relevant outputs, with legitimacy and support from all stakeholders, it needs to aim for transdisciplinary. Scholz and Steiner define interdisciplinary as the *merging of concepts and knowledge from different disciplines* (2015, p. 529). Transdisciplinary, on the other hand, is a process in which stakeholders from science and practice learn from each other when dealing with a specific real world problem. The result of this integrated generation of knowledge are socially robust orientations. These socially robust orientations refer to the shared understanding (of all stakeholders involved, from science, and society) of patterns of causality of a given issue (Scholz & Steiner, 2015). What is important to mention here is that all the actors involved have to be part of the problem definition, the causation chain and the definition of the relevant factors of the system it aims to move from the status quo to a shared desired notion of how things can be. Therefore, ABMs can function to CDTs, apart from contributing to include social processes to the city model, as a mechanism to communicate and illustrate the variables and processes that are part of the perceived problem. In this sense, ABMs can help to build better models to the extent they are able to capture variables informed by theory and practice. Batty holds a view that in our understanding captures well the direction we can push to use simulation to design better public policies. He says that in the outlook of, the challenge of simulating a city should encourage us to keep pushing for DTs. However, echoing Marc Kac's, he remembers, one can not forget to take models not so much as its usefulness to predict, but as its function to "*polarize thinking and pose sharp questions*" (Kac, 1969 *apud* Batty, 2024, p. 193). It is, perhaps, the case that the way we can expect our models, and hi-end technology to work in our favor, is to start building them from the outside, as the tools we use to explore the complexity of our environment.

7. Final Remarks

In a moment in time where technology is increasingly expected to solve the existential threats we face, it is important to take a step back and inquire about the limitations and actual potentials of such tools. During the course of this dissertation we presented two types of simulation technologies, DTs, and ABMs, explained their main characteristics and how they are currently being used to either study policy, or provide useful information for decision makers. This master thesis addressed a fundamental topic on the development of CDTs. A technology which is expected to revolutionize the way we manage cities, expanding the possibilities to deliver better public policies, and featuring as an essential component to achieve smart and sustainable cities. The prospect of integrating these two technologies, comes from an identified pitfall in the CDTs to appropriately incorporate the city's social and economic processes. ABMs have an extensive methodological tradition in studying social and economic phenomena. Hence, the potential suitability for ABMs to fill the gap in CDT applications, resulted in a research question that guided our thesis in an attempt to convey the extent in which ABMs could benefit implementations of CDTs. As any technological application must be understood in its context, we framed the discussion in a specific topic of Urban Mobility. Overall, the goal was to discuss the possibility to join synergies between a nascent technology, CDT, with a most established research methodology, departing from a simple ABM, as an illustrative case in order to answer our guiding question.

For one side, our literature review showed us the novelty and potential CDTs (CDTs) represent to city planners, policy analysts and decision makers. However, we also saw that in order for this realization to be possible, there is still a considerable amount of technical and research work to be done. Although the concept of DT has been around for more than 20 years, there is a long way to consolidate the research field, the characterizing notion of DTs is still under dispute. This struggle to arrive in a precise definition was also reflected when we tried to make a clear distinction of ABM in relation to CDTs. Most of the challenge resided in conceptual similarities, and the fact that the bidirectional communication between the real and the virtual system, a key aspect to differentiate DT from other types of digital models, was not present in most of the existing CDTs we had access to. In another way, we were interested in how integrating ABMs and CDTs could help to advance more robust models. This study contributed to initially addressing what are the possible results and the possible limitations of such integration via a theoretical proof of concept. The exploration of this concept happened with an adaptation we were able to implement in an existing model, TransportVarese. The original model was developed to test policies that could help cities to move from highly polluting transportation modes, to greener options. In this sense, TransportVarese provided two different policies, one of economic nature that worked to increase costs for private car use. And another policy to work directly with people's preference for transportation modes. The outstanding performance of policies in the original model to incentivize greener transportation, specially in the case of bike adoption, got us wondering what measure could help to boost public transport adoption. Although Transportation Demand Management is not really designed to increase adoption, but to better use the system resources, we were curious if there was any new pattern that could arise when we start varying prices for public transport. This additional policy, a dynamic fare policy worked for our goals in two levels, for one it helped us to question the original model, the omitted variables and broad assumptions and the implications of such policies to the transportation system. In another hand it served to justify the existence and suitability of DT, given the fact that Travel Demand Management, and precisely, fares calculation, rely in accurate data on supply and demand, in different times of the day. That said, price variations could be enough to start a new pattern of commuting behavior, yet the emergence of this new pattern could also justify a refinement to the model to a point where we could see the necessity of an auxiliary source of data (e.g. a Digital Twin). It was possible to suggest a different system for dynamic pricing based on real-time information, leveraging the system that DTs concept would offer. Realizing the data flow from real to virtual, furthered from the digital to virtual, as illustrated in Figure 1. Thus, a change of preference for public transport in the real twin, would create a new demand for public transport in the DT, which in turn would feed back to the real twin with a new dynamic fare, which in turn would adjust the preferences for public transport once more. As we described in the results section, this change in preferences did not happen significantly. The scenarios considered led to the stagnation in public transportation preference, not to mention they were working backwards the objective of the original model intent of engineering policies that lead to more sustainable transportation modes. But whenever we succeeded in increasing preference for public transportation, we did so by swaying bike adopters, and not car users.

TransportVarese is far from being a definitive model to simulate transportation mode choice, but it proved useful to demonstrate in which ways CDTs can benefit from Agent-Based Modeling. Such results corroborate the illustrative power ABMs can offer, helping planners to advance policy design, and the urgency to integrate refined real data to test and make explicit the assumptions and causal links in the system. Echoing the calls to understand CDTs as sociotechnical systems, we touched ethical issues embedded in this discussion and reinforced the need for transdisciplinary approaches in CDTs design. In this sense, we see ABM not only contributing to CDTs as a methodological instrument to represent social and economic process, but as a tool to communicate and negotiate models with all stakeholders. This derives from an understanding that models are not to be taken as absolute truths or source of objective answers, but refined mechanisms to understand and generate common knowledge.

7.1. Future work

While some advancements were made regarding what one technology can add to the other, additional technical work and research is needed. One way to advance future developments is to keep refining the model to meet the possible data inputs a CDT can offer. For TransportVarese, this could mean to include variables such as the number of buses, passengers demand requirements. The model could be redesigned to simulate hours of the day, instead of entire days, which in turn could help us to test impacts of TDM appropriately, specially its impact on public transportation adoption. Besides, it would be interesting to be able to also measure how much emissions are saved with the TDM policy implemented, which in theory can happen not only if there is an increase in public transportation adoption, but simply because there is vehicle circulation can also be functioning on and on demand basis. Another route could be to develop a middle range approach to this integration. Such development, of a whole theory of agent behavior was also suggested here, and would help to catalyze the interactions between agents and physical space in deeper level of detail. To the best of our knowledge, MRT integration with a CDT is a task has yet to be taken.

8. References

Anzola, D., Barbrook-Johnson, P., & Gilbert, N. (2022). The Ethics of Agent-Based Social Simulation. *Journal of Artificial Societies and Social Simulation*, *25*(4), 1. https://doi.org/10.18564/jasss.4907

Arowoiya, V. A., Moehler, R. C., & Fang, Y. (2024). Digital twin technology for thermal comfort and energy efficiency in buildings: A state-of-the-art and future directions. *Energy and Built Environment*, *5*(5), 641–656. https://doi.org/10.1016/j.enbenv.2023.05.004

Axelrod, R. (1997). Advancing the art of simulation in the social sciences. *Complexity*, *3*(2), 16–22. https://doi.org/10.1002/(SICI)1099-0526(199711/12)3:2<16::AID-CPLX4>3.0.CO;2-K

Barat, S., Parchure, R., Darak, S., Kulkarni, V., Paranjape, A., Gajrani, M., Yadav, A., & Kulkarni, V. (2021). An Agent-Based Digital Twin for Exploring Localized Non-pharmaceutical Interventions to Control COVID-19 Pandemic. *Transactions of the Indian National Academy of Engineering: An International Journal of Engineering and Technology*, *6*(2), 323–353. https://doi.org/10.1007/s41403-020-00197-5

Bastarianto, F. F., Hancock, T. O., Choudhury, C. F., & Manley, E. (2023). Agent-based models in urban transportation: Review, challenges, and opportunities. *European Transport Research Review*, *15*(1), Article 1. https://doi.org/10.1186/s12544-023-00590-5

Batty, M. (2018). Digital twins. *Environment and Planning B: Urban Analytics and City Science*, *45*(5), 817–820. https://doi.org/10.1177/2399808318796416

Batty, M. (2024). Digital twins in city planning. *Nature Computational Science*, *4*(3), 192–199. https://doi.org/10.1038/s43588-024-00606-7

Bianchi Alves, B., Bou Mjahed, L., & Moody, J. (2023). *Decarbonizing Urban Transport for Development* (Mobility and Transport Connectivity Series). World Bank.

http://hdl.handle.net/10986/40373

Birks, D., Heppenstall, A., & Malleson, N. (2020). Towards the Development of Societal Twins. *24th European Conference on Artificial Intelligence*.

Cai, C., Ying, C., Peng, H., Nan, M., Jingtao, Z., & Yuxiong, Z. (2022). *Digital Twin Cities: Framework and Global Practices* (Insight Report). World Economic Forum. https://www.weforum.org/reports/digital-twin-cities-framework-and-global-practices/

Caldarelli, G., Arcaute, E., Barthelemy, M., Batty, M., Gershenson, C., Helbing, D., Mancuso, S., Moreno, Y., Ramasco, J. J., Rozenblat, C., Sánchez, A., & Fernández-Villacañas, J. L. (2023). The role of complexity for digital twins of cities. *Nature Computational Science*, *3*(5), Article 5. https://doi.org/10.1038/s43588-023-00431-4

Clemen, T., Ahmady-Moghaddam, N., Lenfers, U. A., Ocker, F., Osterholz, D., Ströbele, J., & Glake, D. (2021). Multi-Agent Systems and Digital Twins for Smarter Cities. *Proceedings of the 2021 ACM SIGSIM Conference on Principles of Advanced Discrete Simulation*, 45–55. https://doi.org/10.1145/3437959.3459254

Collins, A., Koehler, M., & Lynch, C. (2024). Methods That Support the Validation of Agent-Based Models: An Overview and Discussion. *Journal of Artificial Societies and Social Simulation*, *27*(1), 11.

Creswell, J. W., & Creswell, J. D. (2018). The Selection of a Research Approach. In *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (5th Edition). SAGE Publications.

Deren, L., Wenbo, Y., & Zhenfeng, S. (2021). Smart city based on digital twins. *Computational Urban Science*, *1*(1), 4. https://doi.org/10.1007/s43762-021-00005-y

do Amaral, J. V. S., dos Santos, C. H., Montevechi, J. A. B., & de Queiroz, A. R. (2023). Energy Digital Twin applications: A review. *Renewable and Sustainable Energy Reviews*, *188*, 113891. https://doi.org/10.1016/j.rser.2023.113891

Douglas, N. J., Henn, L., & Sloan, K. (2011). *Modelling the ability of fare to spread AM peak passenger loads using rooftops*.

Dunn, W. N. (2017). Structuring Policy Problems. In *Public Policy Analysis: An Integrated Approach* (6th ed., pp. 68–117). Routledge. https://doi.org/10.4324/9781315181226

Edmonds, B., Le Page, C., Bithell, M., Chattoe-Brown, E., Grimm, V., Meyer, R., Montañola-Sales, C., Ormerod, P., Root, H., & Squazzoni, F. (2019). Different Modelling Purposes. *Journal of Artificial Societies and Social Simulation*, *22*(3), 6.

Furtado, B. A. (2022). Simulation Modeling as a Policy Tool. In M. Howlett, *The Routledge Handbook of Policy Tools*. Taylor & Francis. https://doi.org/10.4324/9781003163954-27

Furtado, B. A., Sakowski, P. A. M., & Tóvolli, M. H. (2015). A Complexity Approach For Public Policies. In B. A. Furtado, P. A. M. Sakowski, & M. H. Tóvolli (Eds.), *Modeling Complex Systems for Public Policies* (pp. 17–36). IPEA. https://repositorio.ipea.gov.br/handle/11058/4051

Glaessgen, E. H., & Stargel, D. S. (2012, April 16). *The Digital Twin Paradigm for Future NASA and U.S. Air Force Vehicles*. 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference - Special Session on the Digital Twin, Honolulu, HI. https://ntrs.nasa.gov/citations/20120008178

Grieves, M. W. (2014). Digital Twin: Manufacturing Excellence through Virtual Factory Replication. *White Paper*, 7.

Grieves, M. W. (2023). Digital Twins: Past, Present, and Future. In N. Crespi, A. T. Drobot, & R. Minerva (Eds.), *The Digital Twin* (pp. 97–121). Springer International Publishing. https://doi.org/10.1007/978-3-031-21343-4_4

Grimm, V., Railsback, S. F., Vincenot, C. E., Berger, U., Gallagher, C., DeAngelis, D. L., Edmonds, B., Ge, J., Giske, J., Groeneveld, J., Johnston, A. S. A., Milles, A., Nabe-Nielsen, J., Polhill, J.

G., Radchuk, V., Rohwäder, M.-S., Stillman, R. A., Thiele, J. C., & Ayllón, D. (2020). The ODD Protocol for Describing Agent-Based and Other Simulation Models: A Second Update to Improve Clarity, Replication, and Structural Realism. *Journal of Artificial Societies and Social Simulation*, *23*(2), 7.

Halvorsen, A., Koutsopoulos, H. N., Lau, S., Au, T., & Zhao, J. (2016). Reducing Subway Crowding: Analysis of an Off-Peak Discount Experiment in Hong Kong. *Transportation Research Record: Journal of the Transportation Research Board*, *2544*(1), 38–46. https://doi.org/10.3141/2544-05

Helbing, D., & Argota Sánchez-Vaquerizo, J. (2022). *Digital Twins: Potentials, Ethical Issues, and Limitations* (SSRN Scholarly Paper 4167963). https://doi.org/10.2139/ssrn.4167963

Helbing, D., Mahajan, S., Fricker, R. H., Musso, A., Hausladen, C. I., Carissimo, C., Carpentras, D., Stockinger, E., Argota Sanchez-Vaquerizo, J., Yang, J. C., Ballandies, M. C., Korecki, M., Dubey, R. K., & Pournaras, E. (2023). Democracy by Design: Perspectives for Digitally Assisted, Participatory Upgrades of Society. *Journal of Computational Science*, *71*, 102061. https://doi.org/10.1016/j.jocs.2023.102061

Jones, D., Snider, C., Nassehi, A., Yon, J., & Hicks, B. (2020). Characterising the Digital Twin: A systematic literature review. *CIRP Journal of Manufacturing Science and Technology*, *29*, 36–52. https://doi.org/10.1016/j.cirpj.2020.02.002

Klein, D., Marx, J., & Fischbach, K. (2018). Agent-Based Modeling in Social Science, History, and Philosophy. An Introduction. *Historical Social Research / Historische Sozialforschung*, *43*(1 (163)), 7–27.

Klügl, F., & Bazzan, A. L. C. (2012). Agent-Based Modeling and Simulation. *AI Magazine*, *33*(3), Article 3. https://doi.org/10.1609/aimag.v33i3.2425

Maggi, E., & Vallino, E. (2017). *TransportVarese* (1.1.0) [NetLogo]. https://www.comses.net/codebases/5419/releases/1.1.0/

Maggi, E., & Vallino, E. (2021). Price-based and motivation-based policies for sustainable urban commuting: An agent-based model. *Research in Transportation Business & Management*, *39*, 100588. https://doi.org/10.1016/j.rtbm.2020.100588

Makarov, V., Bakhtizin, A., & Epstein, J. M. (2022a). Agent-based modeling for a complex world. Part 1. *Economics and the Mathematical Methods*, *58*(1), Article 1. https://doi.org/10.31857/S042473880018970-6

Makarov, V., Bakhtizin, A., & Epstein, J. M. (2022b). Agent-based modeling for a complex world. Part 2. *Economics and the Mathematical Methods*, *58*(2), Article 2. https://doi.org/10.31857/S042473880020009-8

Malleson, N., Birkin, M., Birks, D., Ge, J., Heppenstall, A., Manley, E., McCulloch, J., Ternes, P., Albrecht, S. V., & Woolridge, M. (2022). Agent-based modelling for Urban Analytics: State of the art and challenges. *AI Communications*, *35*(4), 393–406. https://doi.org/10.3233/AIC-220114

Malleson, N., Minors, K., Kieu, L.-M., Ward, J. A., West, A., & Heppenstall, A. (2020). Simulating Crowds in Real Time with Agent-Based Modelling and a Particle Filter. *Journal of Artificial Societies and Social Simulation*, *23*(3), 3.

Mariani, S., Picone, M., & Ricci, A. (2022). About Digital Twins, Agents, and Multiagent Systems: A Cross-Fertilisation Journey. In F. S. Melo & F. Fang (Eds.), *Autonomous Agents and Multiagent Systems. Best and Visionary Papers* (pp. 114–129). Springer International Publishing. https://doi.org/10.1007/978-3-031-20179-0_8

Masoumi, H., Shirowzhan, S., Eskandarpour, P., & Pettit, C. J. (2023). City Digital Twins: Their maturity level and differentiation from 3D city models. *Big Earth Data*, *7*(1), 1–36. https://doi.org/10.1080/20964471.2022.2160156

McGuire, J., Evans, E., & Kane, E. (2021). What Works in Public Awareness Campaigns? A Scoping

Review. In J. McGuire, E. Evans, & E. Kane (Eds.), *Evidence-Based Policing and Community Crime Prevention* (pp. 417–433). Springer International Publishing. https://doi.org/10.1007/978-3-030-76363-3_11

Mills, A. J., Durepos, G., & Wiebe, E. (2010). Middle-Range Theory. In *Encyclopedia of Case Study Research* (pp. 560–561). SAGE Publications, Inc. https://doi.org/10.4135/9781412957397

Nam, T., & Pardo, T. A. (2011). Conceptualizing smart city with dimensions of technology, people, and institutions. *Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times*, 282–291. https://doi.org/10.1145/2037556.2037602

Natalini, D., & Bravo, G. (2014). Encouraging Sustainable Transport Choices in American Households: Results from an Empirically Grounded Agent-Based Model. *Sustainability*, *6*(1), Article 1. https://doi.org/10.3390/su6010050

Nochta, T., Wan, L., Schooling, J. M., & Parlikad, A. K. (2021). A Socio-Technical Perspective on Urban Analytics: The Case of City-Scale Digital Twins. *Journal of Urban Technology*, *28*(1–2), 263–287. https://doi.org/10.1080/10630732.2020.1798177

O'Reilly, R. K., Kollmann, A., Cohen, J. J., & Reichl, J. (2024). Macro-factors driving bicycle adoption as a primary transport mode across Europe. *Travel Behaviour and Society*, *34*, 100669. https://doi.org/10.1016/j.tbs.2023.100669

Park, B. H., Abdul Aziz, H. M., Morton, A., & Stewart, R. (2018). High performance Data Driven Agent-based Modeling Framework for Simulation of Commute Mode Choices in Metropolitan Area. *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, 3779–3784. https://doi.org/10.1109/ITSC.2018.8569232

Peer, S., Knockaert, J., & Verhoef, E. T. (2016). Train commuters' scheduling preferences: Evidence from a large-scale peak avoidance experiment. *Transportation Research Part B: Methodological*, *83*, 314–333. https://doi.org/10.1016/j.trb.2015.11.017

Preiss, A., Hadley, E., Jones, K., Stoner, M. C. D., Kery, C., Baumgartner, P., Bobashev, G., Tenenbaum, J., Carter, C., Clement, K., & Rhea, S. (2022). Incorporation of near-real-time hospital occupancy data to improve hospitalization forecast accuracy during the COVID-19 pandemic. *Infectious Disease Modelling*, *7*(1), 277–285. https://doi.org/10.1016/j.idm.2022.01.003

Pretel, E., Navarro, E., López-Jaquero, V., Moya, A., & González, P. (2022). Multi-Agent Systems in Support of Digital Twins: A Survey. In J. M. Ferrández Vicente, J. R. Álvarez-Sánchez, F. de la Paz López, & H. Adeli (Eds.), *Bio-inspired Systems and Applications: From Robotics to Ambient Intelligence* (pp. 524–533). Springer International Publishing. https://doi.org/10.1007/978-3-031-06527-9_52

Rail Safety and Standards Board. (2020, March 24). *A real-time digital twin of St-Pancras station, and the journey to create an emotion-optimised railway*. https://www.rssb.co.uk/about-rssb/insights-and-news/blogs/real-time-digital-twin-of-st-pancras-station-and-journey-to-create-an-emotion-optimised-railway

Rodrigue, J.-P. (2024). Transportation and Geography. In *The Geography of Transport Systems* (Sixth Edition). Routledge. https://transportgeography.org/contents/chapter1/

Saharan, S., Bawa, S., & Kumar, N. (2020). Dynamic pricing techniques for Intelligent Transportation System in smart cities: A systematic review. *Computer Communications*, *150*, 603–625. https://doi.org/10.1016/j.comcom.2019.12.003

Scholz, R. W., & Steiner, G. (2015). The real type and ideal type of transdisciplinary processes: Part I—theoretical foundations. *Sustainability Science*, *10*(4), 527–544. https://doi.org/10.1007/s11625-015-0326-4

Shahat, E., Hyun, C. T., & Yeom, C. (2021). City Digital Twin Potentials: A Review and Research

Agenda. *Sustainability*, *13*(6), Article 6. https://doi.org/10.3390/su13063386

Squazzoni, F. (2012). What is agent-based computational sociology all about? In *Agent-based computational sociology* (1st Edition, pp. 1–32). Wiley & Sons. https://doi.org/10.1002/9781119954200

Thelen, A., Zhang, X., Fink, O., Lu, Y., Ghosh, S., Youn, B. D., Todd, M. D., Mahadevan, S., Hu, C., & Hu, Z. (2022). *A Comprehensive Review of Digital Twin -- Part 1: Modeling and Twinning Enabling Technologies* (arXiv:2208.14197). arXiv. http://arxiv.org/abs/2208.14197

Thomas, G. (2011). A Typology for the Case Study in Social Science Following a Review of Definition, Discourse, and Structure. *Qualitative Inquiry*, *17*(6), 511–521. https://doi.org/10.1177/1077800411409884

Tolk, A., Clemen, T., Gilbert, N., & Macal, C. M. (2022). How Can We Provide Better Simulation-Based Policy Support? *2022 Annual Modeling and Simulation Conference (ANNSIM)*, 188–198. https://doi.org/10.23919/ANNSIM55834.2022.9859512

Wan, L., Nochta, T., Tang, J., & Schooling, J. (2023). Introduction. In *City Digital Twins: Fundamentals, conceptualisation and challenges* (1st Edition). ICE Publishing. https://doi.org/10.1680/dtsc.66007.001

Wilensky, U., & Rand, W. (2015). What Is Agent-Based Modeling? In *An Introduction to Agent-Based Modeling: Modeling Natural, Social, and Engineered Complex Systems with NetLogo*. MIT Press. http://ebookcentral.proquest.com/lib/ucd/detail.action?docID=3339969

Winters, M., Davidson, G., Kao, D., & Teschke, K. (2011). Motivators and deterrents of bicycling: Comparing influences on decisions to ride. *Transportation*, *38*(1), 153–168. https://doi.org/10.1007/s11116-010-9284-y

Wise, S., Crooks, A., & Batty, M. (2017). Transportation in Agent-Based Urban Modelling. In M.-R. Namazi-Rad, L. Padgham, P. Perez, K. Nagel, & A. Bazzan (Eds.), *Agent Based Modelling of*

Urban Systems (pp. 129–148). Springer International Publishing. https://doi.org/10.1007/978-3-319-51957-9_8

Xames, Md. D., & Topcu, T. G. (2024). A Systematic Literature Review of Digital Twin Research for Healthcare Systems: Research Trends, Gaps, and Realization Challenges. *IEEE Access*, *12*, 4099–4126. https://doi.org/10.1109/ACCESS.2023.3349379