Delft University of Technology

Impact of Migration and Urbanization on Cities

an Agent-Based Model on the effects of Migration on the city of The Hague

Master thesis submitted to Delft University of Technology in partial fulfilment of the requirements for the degree of Master of Science in Engineering and Policy Analysis at the Faculty of Technology, Policy and Management by

Author:
J.H. (Jochem) Vlug (4165993)

Committee:

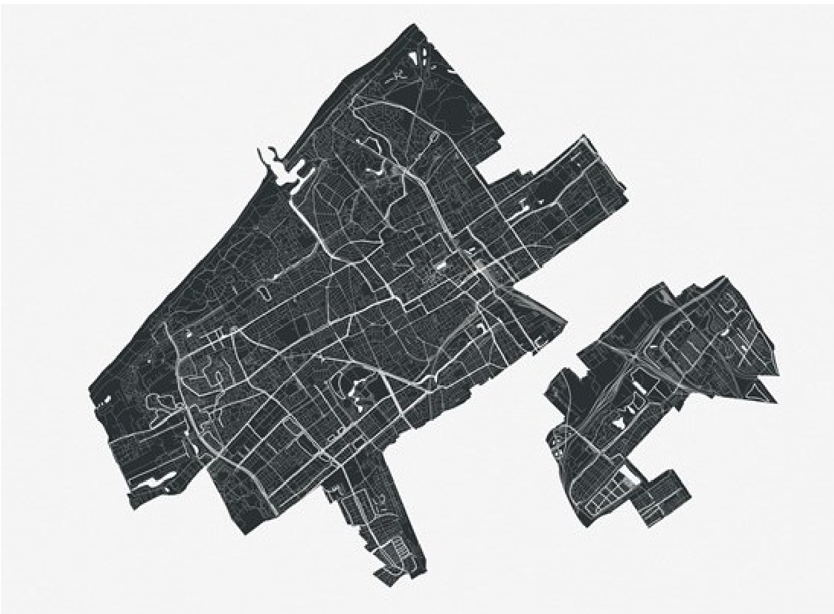
Chairperson:	Dr.ir. I. (Igor) Nikolic
First Supervisor:	Dr.ir. T. (Trivik) Verma
External Supervisor:	ir. M. (Mikhail) Sirenko
External Supervisor:	Prof. dr. B.A. (Bartel) Van de Walle

September 14, 2020

TU Delft Delft University of Technology

Abstract

Migrants are moving in great numbers towards urban areas as a result of urbanization and drivers such as political-, economical- and climate crises. The influx of new people brings new challenges for cities and municipalities to provide for a suitable environment for new and existing citizens whilst preventing adverse effects to occur. Current literature describes the dynamics of the city using effects such as segregation, gentrification and urban decay. However, the changing character of urban areas is not linked with the decision-making behavior of citizens in current literature. Furthermore, the dynamics of urban areas are rarely observed from a meso-scale perspective. This research aims to describe the relation between migration patterns and changes to the fabric of the city. A conceptual model on the dynamics involved in observing decision-making of citizens and observing changes to the city fabric is made. Using real-world data of the city of The Hague in The Netherlands, a Data-Driven Agent-Based model is made in Netlogo. Data analysis on the outcomes of the model using Python provide a better understanding of the impact of migration on the changes to the city fabric. The model output describes the impact on the city caused by the influx of migrants. More migrants entering the city can drastically increase homelessness or the "pushing out" of citizen groups because of housing shortages. This is most prevalent for the poorest citizens, since social rent housing is first to become scarce when more people enter the city. An increase in income inequality or a decrease in average migrant income results in even more pronounced housing shortages. The model simulates 6 policy interventions for the municipality of The Hague. The model shows that all policies that increase housing options significantly increases well-being of citizens. However, the policies that try to improve safety and health do not show an increase in well-being of citizens. By means of historical validation the model is tested for validity. The model is able to describe much of the decision-making behavior of migrants and citizens. However, for certain ethnicities, the model is currently unable to describe the decision-making process in such a way that it can represent the observed real-world moving patterns. A framework has been presented that can be used to better explain the impact of migration on cities. Furthermore, the conceptual model can be applied to different aspects of city dynamics, for instance the transport or energy sector. The ABM shows that current housing options and expected influx of migrants become problematic within the coming 10 years. An indication of the usefulness of proposed policy interventions is presented. More research and surveys are needed to better understand these mechanics and their relation to changes to the fabric of the city. Additional research into the changes to the neighborhood amenities and businesses as a result of changes in citizen composition can aid policy making and understanding the developments of the ever-changing city landscape.



Acknowledgements

Dear reader,

You are about to embark on a journey with me, Jochem, of reading the final phase of my time as a student. After a lot of long work days (and nights), many discussions with fellow student colleagues, supervisors, friends and family, a lot of trial and error, new insights, things learned, mistakes made and errors discovered, the "finish line" is on the horizon. I hope this document will be as fun to read as it was to make (and believe me, it is not only you that will struggle to read the whole thing). But before I will let you go on your journey with me, there are some people I would like to thank:

First and foremost, I would like to thank the amazing committee that helped me through the process of writing my thesis. Starting from refugees in Syria, this project has taken many turns but thanks to the amazing support of my supervisors I was able to keep my sanity and find a course to follow.

As the (former) chair of my committee **Bartel** used his experience and expertise to ask the right questions at the right to help me find my way in defining the goal of my research.
Furthermore, I would like to thank **Trivik** for all his help, guidance, and reading through everything I always send him. His articles on writing scientific reports were a major help and his vision was always there to guide me when I was lost.
I am very thankful for the support of **Igor** when I was in need of advice on modelling, coping with complexity, the role of psychology and philosophy in my research, when helping me crash the HPC server and last but not least was able to always give some proper life advice.
Mikhail, your help has saved my research from going down the drain. A perfect mix of critical questions and poking at my methods helped me stay focused and aware of mistakes. In times of need your amazing coding skills helped me tame the Python.
I would also like to thank some colleagues for supporting me in my process, sometimes by having fun, drinks or playing a game late at night, but you were also there to discuss technical issues in modeling, long conversations on reporting and research methodologies and the outline of the Thesis project in general. In no particular order, I would like to mention **Anne**, **Brennen**, **Dirk**, **Fuuk**, **Jaromir**, **Jason**, **Joep**, **Kevin**, **Marieke**, **Pietro**, **Rob**, and **Vyshnavi**.

This project would not have been possible if it wasn't for the help and support of my family. My parents, **A** & **W**, thank your for always have a nice warm meal ready in the weekends, the endless discussions, the holidays and trips, the financial and mental support and for being awesome people.
My sister, **Lotte**, for keeping up with my nonsense (when I tend to go on a rant on something about climate change), for always reading the articles I sent you, helping me with my thesis, job interviews, for the fun dinners, tennis sessions, board games and all other shenanigans we enjoy.
Gwen, my partner (in crime), thank you for always being there for me, staying patient and supportive throughout the process. Thanks for keeping me focused when I slack off, and rested when I have done enough. Thanks for sticking around even during these tense times, I can't wait for the next five years with you.
Gerda, **Ruben** & **Sharona**, thank you for being the loveliest family-in-law that keep me grounded, loved and fed. Thanks **Sarah** for being the cutest and fluffiest of dogs. Woof.

Last but not least, I would like to thank all my friends (which are too many to all give a shout out). I love each and everyone of you, and can't wait for the pandemic to be fully over so we can meet like old times again.

Contents

1 Introduction **5**
- 1.1 Urbanization . . . 5
- 1.2 Relevance of research . . . 5
- 1.3 Uncertainty in changes to the city . . . 6
- 1.4 Urban Science . . . 6
- 1.5 Decision-making models for migration . . . 7
- 1.6 Agent-Based Models on City Sciences . . . 8
  - 1.6.1 Migration . . . 8
  - 1.6.2 Decision-making of migrants within cities . . . 8
  - 1.6.3 Other models describing city dynamics . . . 9
- 1.7 Current state of literature and research . . . 9
- 1.8 Knowledge Gap . . . 9
- 1.9 Research Question . . . 10
- 1.10 Research Approach . . . 10

2 Data **12**
- 2.1 Data Sources . . . 12
- 2.2 Data Preparation . . . 13
- 2.3 Changes to the city . . . 15

3 Methods **16**
- 3.1 Model Overview . . . 16
  - 3.1.1 Scope of the System . . . 16
  - 3.1.2 Conceptual Model . . . 17
  - 3.1.3 Model procedure . . . 21
- 3.2 Model Implementation . . . 22
  - 3.2.1 Agent properties . . . 23
  - 3.2.2 Agent Decision Logic . . . 23
  - 3.2.3 The Agent-Based Model . . . 24
- 3.3 Policy Implementation . . . 26
  - 3.3.1 Current alternatives . . . 26
  - 3.3.2 Building More Houses . . . 26
  - 3.3.3 Transform Existing Housing Into Smaller Properties . . . 27
  - 3.3.4 Re-purpose Industry/Business Zoning . . . 27
  - 3.3.5 Increasing Health Amenities . . . 27
  - 3.3.6 Improving Safety . . . 27
  - 3.3.7 Change Social Housing Prices . . . 27
- 3.4 Model Verification . . . 28

4 Results **29**
- 4.1 Result Data Overview . . . 29
- 4.2 Relation Between Migration Patterns and Changes to the City . . . 31
  - 4.2.1 The Parameter Space . . . 31
  - 4.2.2 Outcome Differences from Income Inequality . . . 31
  - 4.2.3 Outcome Differences from Influx of Migrants . . . 32
  - 4.2.4 Outcome Differences from Income of Migrants . . . 34
  - 4.2.5 Outcome Differences from Randomized Agent Properties . . . 36

4.2.6 Outcome Differences from Inflation and Price Change . . . 38
4.2.7 Spatial Differences . . . 39
4.2.8 Temporal Differences . . . 40
4.2.9 Exogenous Factors Changing the City Fabric . . . 41
4.3 Validity of the Model . . . 42
4.3.1 Historical Validation . . . 42
4.3.2 Validity Findings . . . 45
4.4 Opportunities and Adverse Effects for The Hague . . . 45
4.4.1 Policy Levers . . . 46
4.4.2 Transforming Houses to Rental Options . . . 47
4.4.3 Improving Healthcare Facilities . . . 48
4.4.4 Increase Social Housing Availability . . . 49
4.4.5 Build More Houses . . . 50
4.4.6 Allow Mixed-Use Zones . . . 51
4.4.7 Improve Safety, Reduce Crime . . . 53
4.4.8 Spatial Observations . . . 54
4.4.9 Temporal Observations . . . 56
4.4.10 Changing the Fabric with Policies . . . 57

5 Conclusion **60**

6 Discussion **62**
6.1 Further research . . . 67
6.2 Limitations . . . 68

References **70**

A Appendix I: Time Planning **76**

B Research Approach **78**
B.1 Desk Research / Theory Formalization . . . 78
B.2 Conceptual Model . . . 78
B.3 Data gathering and preparation . . . 78
B.4 Agent-Based Model . . . 78
B.5 Data Analysis . . . 79
B.6 Policy Analysis . . . 79
B.7 Time Planning . . . 79
B.8 Research Flow . . . 79
B.9 Reporting . . . 79

C Background Literature **81**
C.1 Transport & Mobility . . . 81
C.2 Urban Planning & Housing . . . 81

D Neighborhoods of The Hague **82**

E UML **86**

F Social Group Composition **87**
F.1 Social Groups . . . 87
F.2 Factors for Categorization . . . 88
F.3 Calculation of Categorization . . . 88

G Assumptions **89**

H Model Narrative **91**
H.1 The actors . . . 91
H.2 Model Procedure . . . 91
H.2.1 Input . . . 92
H.2.2 Initialization . . . 92
H.2.3 Simulation . . . 93
H.2.4 Output . . . 94
H.3 Decision-making Logic . . . 94
H.4 Exogenous Factors . . . 95

I Model Verification **96**
I.0.1 Recording and tracking agent behavior . . . 96
I.0.2 Single-agent testing . . . 97
I.0.3 Interaction testing in a minimal model . . . 98
I.0.4 Multi-agent testing . . . 98

J Model Simulation Appendix **102**
J.1 Model Simulation Settings . . . 102
J.1.1 Unexpected Outcomes . . . 103
J.1.2 New Simulation Runs . . . 103
J.2 Simulation Run Parameters . . . 103
J.3 Simulation Descriptive Statistics . . . 103
J.3.1 Full Data Descriptive Statistics . . . 103

K Key Performance Indicators **105**
K.1 Homelessness . . . 105
K.2 Measuring on Neighborhood Level . . . 106
K.3 Available Housing Options . . . 106
K.4 Average Income & Housing Value . . . 106
K.5 Average Utility . . . 106
K.6 Citizen Count . . . 106
K.7 Social Group Prevalence . . . 106
K.8 Ethnicity Prevalence . . . 107

L Sensitivity Analysis **108**
L.1 Exogenous Sensitivity . . . 108
L.1.1 Variation in Standard Deviation of Income Distribution . . . 109
L.1.2 Multiplier of Influx of Migrants . . . 112
L.1.3 Average Spendable Income of Migrants . . . 113
L.1.4 Randomized Attributes of Citizens . . . 115
L.1.5 Inflation . . . 117
L.1.6 Aggregation Size of Household Agents . . . 119
L.1.7 Sensitivity of Parameters over Time . . . 122
L.2 Endogenous Sensitivity . . . 123

Chapter 1

Introduction

1.1 Urbanization

With the increase of globalization around the world, many people are migrating towards urban areas (Harvey, 2008). It is estimated that around 85% of all people on Earth will live in urban areas in 2050 (Ritchie & Roser, 2020; United-Nations, 2019). Cities in The Netherlands observe more (young) people moving towards the cities and less people leaving the city for a number of years now, making the cities more crowded and as a result, housing prices skyrocket (van Amsterdam et al., 2015).

The rise in international migration towards European cities is the result of political instability, civil unrest, climate change[1] and social, economical, and other drivers (Carling & Collins, 2018; Doocy, Lyles, Delbiso, Robinson, & Team, 2015; Balcilar & Nugent, 2019; Yazgan, Utku, & Sirkeci, 2015; McLeman & Hunter, 2010; Dun & Gemenne, 2008; Cohen, 2006; Winchie & Carment, 1989). An example of the recent past is the Arab Spring, which resulted in a large influx of migrants from the Middle East towards Turkey (Yazgan et al., 2015; Duvell, 2019).

Because of this influx of new citizens to urban areas, cities are expanding in both size and dimensions. With these expansions, cities begin to face new challenges now and in the near future (Cohen, 2006). In short, as Bettencourt, Lobo, Helbing, Kühnert, and West formulates it, "The inexorable trend toward urbanization worldwide presents an urgent challenge for developing a predictive, quantitative theory of urban organization and sustainable development" (Bettencourt et al., 2007). For future planning, policy and decision-making, cities need to cope with the uncertainty of complex topics such as transport, climate, resiliency, sustainability, spatial planning, segregation, polarization and inequality. However, many cities and municipalities are currently not prepared to handle these complex challenges and many more are not yet aware of the problems they will face in the near future (Bao-xing, 2003; Wenge, Zhang, Dave, Chao, & Hao, 2014).

1.2 Relevance of research

The importance of coping with future changes in cities caused by an increased influx of people moving to urban areas is acknowledged by the Sustainable Development Goals (SDGs) from the United Nations. Goal number 11 is to "Make cities and human settlements inclusive, safe, resilient and sustainable" (Economic & Council, 2019; Nations, 2020a). Other SDGs that are closely related to this topic are SDG number 1, 8, 9 and 10 which look at poverty, work & economics, industry & infrastructure and inequality (Nations, 2020b). These pillars are considered the focus points of attention for the coming 10 years and are noticeably present when looking at the dynamics of urban areas. In more recent debates such as the G20, mass migrations towards urban areas has been high on the agenda as well. In their report, Charles, Guna, and Galal state: "G20 countries need to strengthen institutional capacity at the city level, empowering cities to determine the legal status and protect the rights of all migrants" (Charles et al., 2017).

However, making policy decisions requires information and understanding of the dynamics that are involved in the changes of the urban fabric. To deal with the influx of new citizens and at the same time provide care

[1] An example interactive display of climate migration, made by the New York Times, can be found here.

to a sustainable, resilient and safe housing environment for all citizens in urban areas, it is of importance that policy makers are able to make informed decisions on policy by using as much information on the state of the city and its inhabitants (Jin, Xu, & Yang, 2009). In the near future, policy makers will have to make decisions based on an expected growth of cities that is bigger than the cities have coped with in the past. A better understanding of the interactions that make the fabric of cities is needed to prepare for this future influx of new citizens whilst preventing negative side-effects such as segregation, inequality and falling quality of life. With a lack of insight in the changes of the urban fabric, policy makers are challenged by the lack of resiliency their city faces in the light of the expected influx of migrants (Kirbyshire, Wilkinson, Le Masson, & Batra, 2017).

The research that follows in this report has been conducted as part of the Master's Thesis for *MSc Engineering and Policy Analysis* of the faculty Technology, Policy and Management of the University of Technology in Delft, The Netherlands. Before the literature is discussed in more detail, the relevance of the work with regards to the Master's program is highlighted. First, by looking at the different Sustainable Development Goals, it becomes clear that migration is a topic which can be categorized as a *Wicked Problem* (Nations, 2020a, 2020b). This means that the nature of migrations: its causes and effects are deeply rooted in systems which are too complex to change or even completely define (Head, 2008). Furthermore, the implications on society of the influx of migrants has been discussed in this introduction, making it a relevant social and societal problem. By further analyzing this problem, new insights can be found with regards of better understanding this *wicked problem*. By doing so, this research aims to provide insight in the complex mechanisms of the influx of migrants on cities and therefore aims to provide policy recommendations on the subject matter. This approach is explained in more detail with the help of defining the research question and sub questions, which are highlighted in Section 1.9. However, first a literature study has been conducted to present the current state-of-the-art knowledge on migration simulation.

1.3 Uncertainty in changes to the city

It is clear that the future will bring an increase in migration and urbanisation. As shown, the societal relevance and importance of coping with the influx of new citizens in urban areas and preventing negative side-effects from occurring as a result of said influx is paramount. However, the relevance of studying the impact of migration to study changes to the city and its form, citizens and shape, is a second topic. This research focuses on better understanding the impact of migration flows on the changes it produces to the city.

An example of such changes can be observed in the city of The Hague, The Netherlands, in which the amount of new migrants entering the city since 2008 has been much higher than expected, causing more housing shortages (Ontwikkeling, 2012). Furthermore, with the influx of many new people with different backgrounds, cultures and other differences, changes to the urban fabric are expected and many cities acknowledge they are not ready to cope with these changes nor have the insight what changes to expect when new migrants will arrive in the city (MacDonald & Sampson, 2012; McFarlane & Rutherford, 2008). Because little is known to the decision-making of migrants with regards to finding suitable housing locations, the municipality is unable to prevent inequality and changes to the neighborhood fabric from occurring.

When looking at the history of *De Schilderswijk*, a neighborhood in The Hague, the problems that can arise from the influx of migrants become more apparent. This neighborhood has had a long history of being a place where new migrants reside when they arrive in the city. This neighborhood has seen many reforms and policies have tried to combat inequality, social isolation and polarization. However, literature shows it has not always been successful, in part caused by a lack of understanding of the interactions between citizens in these parts, causing the neighborhood to become "a revolving door for migrants" (Kritsioudi, 2015; R. Kloosterman & Lambregts, 2001; Hoekstra, 2018).

1.4 Urban Science

To better understand how cities change and evolve, the field of Urban Science (which looks at the dynamics that can be observed, measured and calculated in cities) is looked at in more detail. Cities and urban areas have been a field of study for decades (Forrester, 1970; Schelling, 1978; Winchie & Carment, 1989), but a new era of city science and a new paradigm has started to emerge. As Batty describes it, "In a world now dominated by communications and in a world where most people will be living in cities by the end of this century, it is high time we changed our focus from locations to interactions, from thinking of cities simply as idealized morphologies to thinking of them as patterns of communication, interaction, trade, and exchange; in short, to thinking of them as networks." (Batty, 2013). This means, that instead of looking at the physical form of cities such as buildings and roads to understand the dynamics of cities, one should focus on the behavior of its agents such as those who live and work there. By studying the behavior, one can interpret the change from one state to another and might even be able to assess the estimated future conditions of cities given a certain scenario.

The dynamics of cities can explain changes to the composition of cities caused by the interactions of the citizens living in the city. This means that there is a way to define the structure of a city, not only by its physical form but also the consistency or residents (Meerow, Newell, & Stults, 2016). The urban fabric of a city is defined as the shape or form of the physical city, which only looks at the buildings and infrastructure. However, the fabric of the city can also include the social and political aspect, taking into account the shaping of the city caused by the humans living in it (McFarlane & Rutherford, 2008). There can not be a single definition of a city throughout history, since the city is always changing (Netto, 2017). For the sake of clarity, in this research, whenever the term fabric is mentioned, this refers to the latter definition which also takes into account the role of citizens that live in the city.

An example of interaction causing changes to the fabric of cities that has been studied frequently is *gentrification*. This phenomena describes the influx of people from a higher social class into old (low income) neighborhoods. This leads to an increase in neighborhood safety, the presence of more amenities and ultimately to higher value of land and housing prices. The latter then causes the original, lower class, residents of the neighborhood to be pushed out (Atkinson, 2002; Nara & Torrens, 2005; Betancur, 2014; Zuk et al., 2015).

1.5 Decision-making models for migration

By better understanding the dynamics of cities, the underlying changes to the city are more easy to observe. However, another key element of understanding the dynamics that result in changes to the city, is the decision-making process of citizens in the city. Moreover, understanding the way migrants decide to move to certain places can help understand the link between the decision and the resulting changes to the city.

Simulating dynamics within cities has been a topic of scientific research for decades (Alfeld, 1995). Batty proposed that a clear strategy is needed to tackle the complexity of city sciences based on network science (Batty, 2013; Newman, 2018). This complexity can be found in the *complexity theorem*, which defines the system as a Complex Adaptive System (CAS), which means the interactions of individuals can be regarded as a complex system of effects caused by these interactions and the interactions on the environment around individuals (Lansing, 2003; Liu et al., 2007; Sayama, 2015; Menezes, Evsukoff, & González, 2013; Chan, 2001; van Tongeren, 2014).

In their book, Lane, Pumain, van der Leeuw, and West show the link between complexity science and city sciences (Lane et al., 2009). Many complex adaptive systems are modeled using a modeling technique called Agent-Based Modeling (ABM) which allows for interactions between agents and their environment and other agents, which makes it possible to observe emergent behavior (Dam, Nikolic, & Lukszo, 2013). Emergence is the phenomena where new behavior can be the result of interaction between multiple parts or actors in a system. For example, when people all decide to go to work at the same time, using the same roads, traffic jams start to appear as a result of bad driving behavior (Wilensky & Payette, 1998).

Many simulation approaches have been taken to capture the complexity of dynamics of cities. To get a grasp of the scientific research efforts made so far, an overview of current models has been made by reviewing literature.

On the macro scale, literature finds that shorter distances, economic prosperity, and cultural similarity (e.g. shared language) attract forced migrants to a certain destination, as does the probability of being granted asylum (Frith, Simon, Davies, Braithwaite, & Johnson, 2019). On the micro-scale, research shows that decision-making relates to the situation of the migrant, its environment and its social network (De Jong & Gardner, 2013; Haug, 2008; Perez, Dragicevic, & Gaudreau, 2019).

Furthermore, researchers find that population growth follows the same pattern, regardless of factors that might influence behavior of citizens (Bettencourt et al., 2007). This scaling of population sizes of cities has a non-linear effect on wealth, innovation and crime (Bettencourt, Lobo, Strumsky, & West, 2010; Bettencourt, Lobo, & West, 2009). Through observation of progress through history, and comparing different regions in the world, Bretagnolle, Pumain, and Vacchiani-Marcuzzo have identified several patterns of city system dynamics, some of which also relate to population and dimension growth of the city (Bretagnolle et al., 2009). However, the aforementioned research focuses on the city-wide effects of growth, and tend to look at the city as a macro-scale system. The behavior that is observed is interesting, but misses meso-scale granularity.

Other models have looked at the meso-scale regarding population growth. Multiple examples of Cellular Automata (CA) models on the meso-scale are present in literature (Jantz, Goetz, Donato, & Claggett, 2010; Torrens, 2003; Nara & Torrens, 2005; Dabbaghian, Jackson, Spicer, & Wuschke, 2010). This field of study is useful to formulate a series of possible futures of population growth of cities, on the meso-scale. However, the models are incapable of modeling interactions between citizens and thus are not able to define the consequence of actors' behavior within a city on the growth of the city.

1.6 Agent-Based Models on City Sciences

Agent-based models focus modeling and simulating the behavior of actors (so-called agents) which can interact with each-other, themselves, and the environment that they exist in. Doing so, it is possible to better observe the resulting behavior based on a set of rules that define the way an agent can act (Bonabeau, 2002). From these observed behavior, a better understand of real-world phenomena can be formed or confirmed. Using ABMs to better understand human behavior has been done for many years, even back in 1978 (Schelling, 1978).

By studying literature, a list of relevant models and research have been found that can be used as a framework for defining the behavior of citizens in cities. Furthermore, this literature lays the foundation for better understanding the dynamics between citizens and the city, and between migrants and their move to within the city.

On the micro-scale, agent-based models on migration are more prevalent. An overview of literature on ABMs in urban research was made by Chen, in this paper, a broad scope of models is reviewed. Many models focus on transport and housing in urban areas (Chen, 2012). Another paper by Klabunde and Willekens reviews Agent-Based models for migration behavior (Klabunde & Willekens, 2016). Because of the vast quantity of Agent-Based Models in literature and their broad appliances within the city science domain, this chapter reviews the models that are most relevant for the design of a conceptual city model from a meso-scale design view. The work and scientific contribution to the field of each paper is defined as well as the possibilities for further research.

1.6.1 Migration

Many of the papers in migration research focus on the macro-scale of migration flows. Furthermore, a lot of the studies only focus on the current status-quo, and describe the current situation or try to predict the future. These approaches do not look at the meso-scale, nor do they focus on the changing of the fabric of a city. Examples of interesting research in this field are: the defining of drivers and key factors in migration patterns on an urban level (Carling & Collins, 2018), A computational model of predicting key spatial patterns of conflict-induced forced displacement (Edwards, 2008), The use of data to construct agent-based simulations of refugee movements (Suleimenova, Bell, & Groen, 2017),Forecasting migration patterns from Syria towards certain settlements and camps, using drivers to predict behavior (Huynh & Basu, 2019), Looking the effect of housing prices on in-city migration patterns from (mostly) citizens with an immigrant status (Kashnitsky & Gunko, 2016).

1.6.2 Decision-making of migrants within cities

During the literature review, two studies were especially relevant for the research at hand since they combine both aspects of complexity that are considered for this research. These papers have combined the decision-making behavior of citizens on the changes of the city landscape. In other words, these models look at the influence of citizen's decision to move on the composition of citizens within (parts of) the city. First, Tomasiello, Giannotti, and Feitosa describes the decision-making of citizens and the resulting accessibility to transport. This model measures the average travel time of citizens to their work location and correlates this accessibility to income class. What is interesting, is the decision-making behavior of agents in the model, which shows potential to be applied to the decision-making behavior of migrants entering a new city (Tomasiello et al., 2020). The second paper, by Perez et al., does exactly that. By looking at the factors included in the decision-making process of migrants, an estimating can be made of migration behaviors of migrants. This research showed the influence of certain "preferences" of migrants on the expected movement behavior (Perez et al., 2019).

1.6.3 Other models describing city dynamics

More (Agent-Based) models have been made, and research has been conducted to better understand the dynamics of and within cities. However interesting these models are, they are out of scope of the research at hand. Therefore, mentions to the most relevant works are shown in Chapter C. These papers are out of scope of the research, but could be taken into consideration when looking at different aspects of the dynamics of cities in follow-up research.

1.7 Current state of literature and research

Currently, literature on the decision-making of migrants for deciding a place to live is available, and much has been written about changes to the fabric of cities where effects such as segregation, gentrification, urban decay and inequality are topics much covered. However, the changing character of neighborhoods has not been linked with the decision-making of citizens in their pursuit to find a suitable place to live. A better understanding of urban dynamics might be found by observing both the decision-making and resulting changes in citizen composition of neighborhoods, as well as the resulting changes to the fabric of neighborhoods.

This research aims to better understand the dynamics of the city by focusing on the decision-making of migrants and the resulting changes to the city fabric. A flowchart provides a better grasp of the complexity of the problem, and further defines an overview of the parts of the system that are in the scope of the proposed research. In Figure 1.1, this flowchart schematically presents the problem. As the diagram shows, there is a feedback loop between the composition of citizens in an area and the availability of amenities. Furthermore, it is expected that the state of the fabric of a neighborhood not only influences housing price, but also (indirectly) influences the demand for certain amenities.

1.8 Knowledge Gap

By studying the current literature in the modeling field, it is unclear what influence the influx of new citizens (and especially migrants) can have on the fabric on the city. Currently, literature is divided if changes to citizen composition in neighborhoods change the fabric of the neighborhood or if the relation is the other way around (where the fabric defines which citizens are attracted to a certain neighborhood) and to what extend the causalities between these effects is present. Furthermore, there is a lack of models looking at decision-making of migrants moving into the city that include changes to the fabric of the city to the simulation. It is therefore unknown if the decision-making of migrants changes based on changes to the fabric of the city, and if this effect can explain phenomena such as gentrification and urban decay. In other words, more research is needed to better understand the link between the decision-making behavior of migrants entering urban areas and the resulting changes to the fabric of the city.

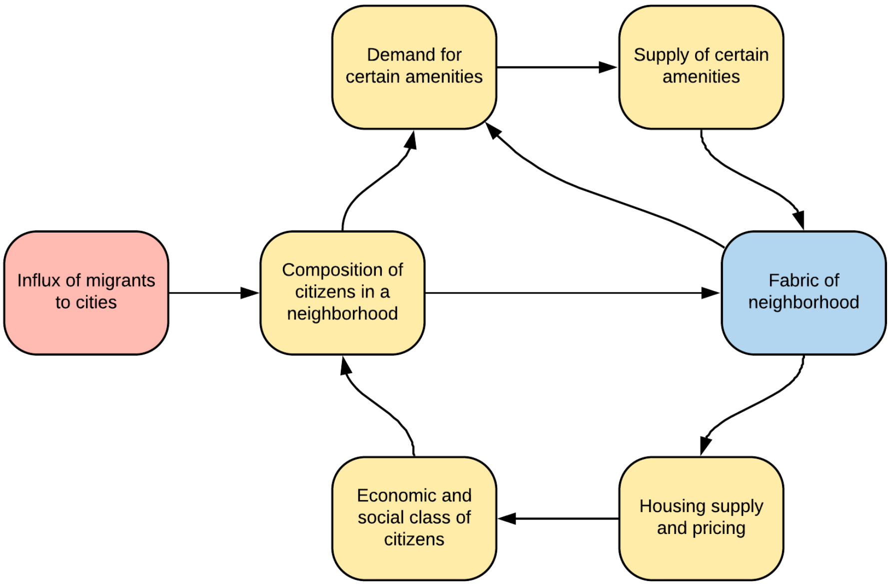


Figure 1.1: Overview of the problem statement.

1.9 Research Question

The current gap in knowledge offers possibilities for further research. And thus, this gap leads to a research proposal. The research focuses on the missing information regarding the influence of changes to the fabric of neighborhoods when looking at the decision-making behavior of migrants moving into a new city. In other words, the research aims to better understand the dynamics between the influx of migrants and the changes to the fabric of the city. To do so, the goal is to develop a conceptual model describing the dynamics of citizens moving in a city and the resulting changes to the urban fabric of that city. Subsequently, using this conceptual model as a framework, the goal is to make an Agent-Based model using statistical (demographic) data as input which can provide insights to the changing of the urban fabric on the meso-scale. To further guide and structure the goals of this research, a research question is proposed.

"What are the relations between migration patterns from the influx of migrants and changes to the fabric of neighborhoods in the city of The Hague?"

1.10 Research Approach

To answer this research question and further define the different parts of missing knowledge, multiple sub questions are formulated. However, first, the type of research that will be conducted is defined. By looking at the knowledge gap, the research will be of an exploratory nature since a new theory that combines current knowledge will be made. According to Edmonds et al., there are 7 possible purposes for making a model. The proposed research in this paper is in line with the *description* modeling purpose, because the goal of the model is to describe the changes of the urban fabric given an influx of migrants (Edmonds et al., 2019). A suggested approach for this modeling purpose is the *Keep It Descriptive Stupid* (KIDS) technique, where the modeling focuses on being descriptive of the observed system rather than being elegant or simplistic (Edmonds & Moss, 2005).

By making a hypothesis, a better understanding of the goal of the research and the intended knowledge that is sought after can be gathered. The hypothesis of the research is that given the decision-making of citizens is based on their preferences in amenities and services but also similarity in citizens living in the same area, people tend to form more homogeneous neighborhood compositions over time. Resulting from this trend, the fabric of neighborhoods will start changing. A similar theory has been posed by Schelling, hypothesizing the uniform change of social class over time (Schelling, 1978). However, in addition to this theory, the research hypothesis assumes this relation to withhold not only for social groups but also ethnicity. Therefore, the hypothesis in short assumes that the observed changes to the city fabric are a result of citizens looking for similarity in their preferred neighbors, whether it is similarity in income, social group, education, background, language or ethnicity. To test this hypothesis and put the research goal to practice, the main research question is accompanied by four sub questions.

The sub questions for this research are:

1. What components and interactions form the basis for a model that describes the impact of migration on a city?

2. How can the dynamics of components and interactions in cities be formalized in an Agent-Based model?

3. How do the dynamics and interactions observed in the model correlate to historical real-world observations?

4. What are the opportunities and adverse effects for the city of The Hague given a mass influx of migrants in the near-future?

After defining the goal of the research by stating the research question and subsequent sub questions, first the data used for simulating the behavior of migrants in the city is highlighted. This is done in the Data section (Chapter 2). Afterwards, the Methods section (Chapter 3) explains the methodology of making a conceptual model of the system and Agent-Based model to simulate the system, this chapter also highlights the verification of said model. In the Results section (Chapter 4), the results from the Agent-Based model are shown and analysed. To conclude, the Discussion section (Chapter 6) highlights the most important findings of the research along with suggestions to further research and shortcomings of the current approach. For more information on the details of the research, the Appendix (Chapter A-F) provides a detailed report of steps taken in the modeling and data analysis of the project.

Chapter 2

Data

As the current gap in knowledge and subsequent research questions show, the goal of this research is to get a better understanding of the dynamics that change the city fabric and in particular, the role of the influx of migrants to the city within the observed changes.

In order to reach this goal, data is needed to observe the system and describe what is happening. Real-world phenomena captured in statistics can be used to simulate the behavior that is observed in the city and therefore be used to better explain the relation between the influx of migrants and the resulting changes to the city fabric. This chapter outlines the sources of data, and subsequent processing of the data, to be used in the research of defining the relationship between the influx of migrants and changes to the city fabric.

To get a better understanding of the changes of the city fabric, data is required not only on the properties of the migrants, but also more information on the city is needed. Using data to describe the current state of the city fabric is needed to observe changes. Furthermore, data also needs to describe the role of current citizens in the system, as to get a better understanding of the role of migrants and the role of current citizens respectively in the changing of the city fabric.

2.1 Data Sources

Governmental bodies in the Netherlands gather a lot of data. This data is published publicly on websites from municipalities and on the website of the Central Bureau for Statistics (CBS) and is free to use (Statline, 2020a). The case study for observing changes in the city fabric is applied to the city of The Hague. Fortunately, the municipality uses its resource to gather data on the current (and past) status of the city, and shares this data publicly on their Open Data platform (of The Hague, 2020).

The data on citizens, the city, and migrant influx can be used to get a better understanding of how the observed system works. Figure 2.1 schematically overviews the relations observed in data using an XLRM schematic (R. J. Lempert, 2003). Here, we can see that modeling the (uncertain) influx of migrants and perceptions of citizens and observing the output of the model can result in a better understanding of changes to the fabric of the city.

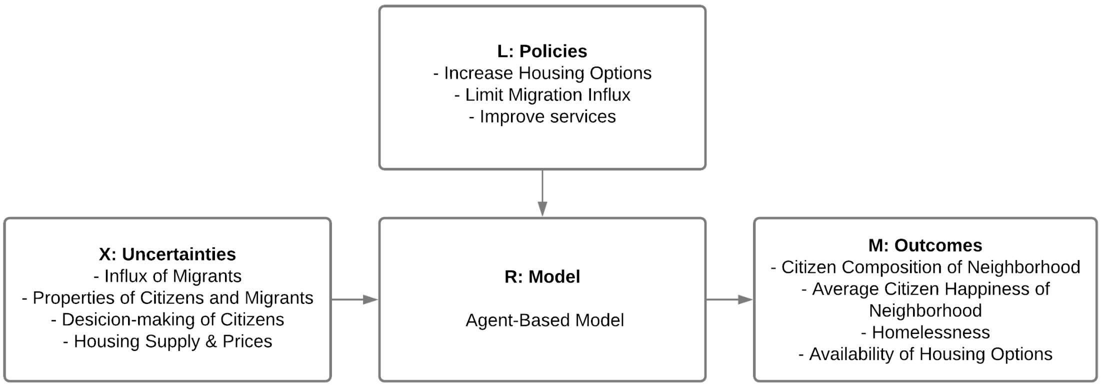


Figure 2.1: XLRM diagram of the observed system. The inputs (uncertainties) require data to model the system to generate outcomes.

To build a suitable model, data is needed that describes the influx of migrants, that describes the current state of the city fabric, that explains the perceptions and decision-making of citizens and migrants and that describes the possible policies the municipality can use to influence the outcomes.

In the model, six datasets are identified to be used. An overview of these datasets and which variables will be used for each respective part of the model is shown in Figure 2.2. First, two datasets are extracted from The Hague municapality Open Data project (of The Hague, 2020). The two datasets are "Ruimtelijke Kengetallen", a source of all types of data about neighborhoods in The Hague and secondly, "OV Haltes", which is a list of all public transport stations in The Hague and their locations within the city.

Three datasets originate from CBS, namely the "Wijken & Buurten" dataset (containing a very broad range of data on neighborhoods in the Netherlands), "Wijk en Buurtkaart" (which has similair data but includes geographical information on the dimensions of the neighborhood) and lastly "Opleidingsniveau naar buurt 2017" (a research to estimate the average education degree in each neighborhood) (Statline, 2020b).

Finally, using a case study on the opinions of residents in neighborhoods of The Hague, the importance (and significance) of different factors for the happiness of residents has been determined. This data is used to define the significance of different factors in the utility function (Center, 2020).

2.2 Data Preparation

The data from the data sources is raw and needs to be processed to be used for modeling. The steps involved are described here. Fortunately, the data provided by the Open Data projects from the municipality and CBS is very "clean", which means not a lot of steps need to be taken to work with the data since it is already in a usable format.

The CBS data is collected for the whole country of the Netherlands, but for this research only the data concerning The Hague is used. First, the datasets are imported from the CBS servers and then, a subset of data is made only containing information of the municipality of "'s-Gravenhage" (which is the official name of The Hague used by CBS). The datasets from the municipality use Dutch variable names and use a different style of separating columns, so a few minor data translations are done in Python using Pandas (McKinney, 2011).

The next step is to sort and categorize data by year. There is a range of years in which data is gathered, and over time, some of the data gathering techniques have changed. For this reason, the earliest data that will be used is from 2015. The data from 2014 and earlier was collected in a different format and is also not relevant for the research.

After categorizing the data on their year of origin, the data needs to be categorized by its location. Fortunately, most data is collected on a neighborhood level, which means the statistics are based on averages of each

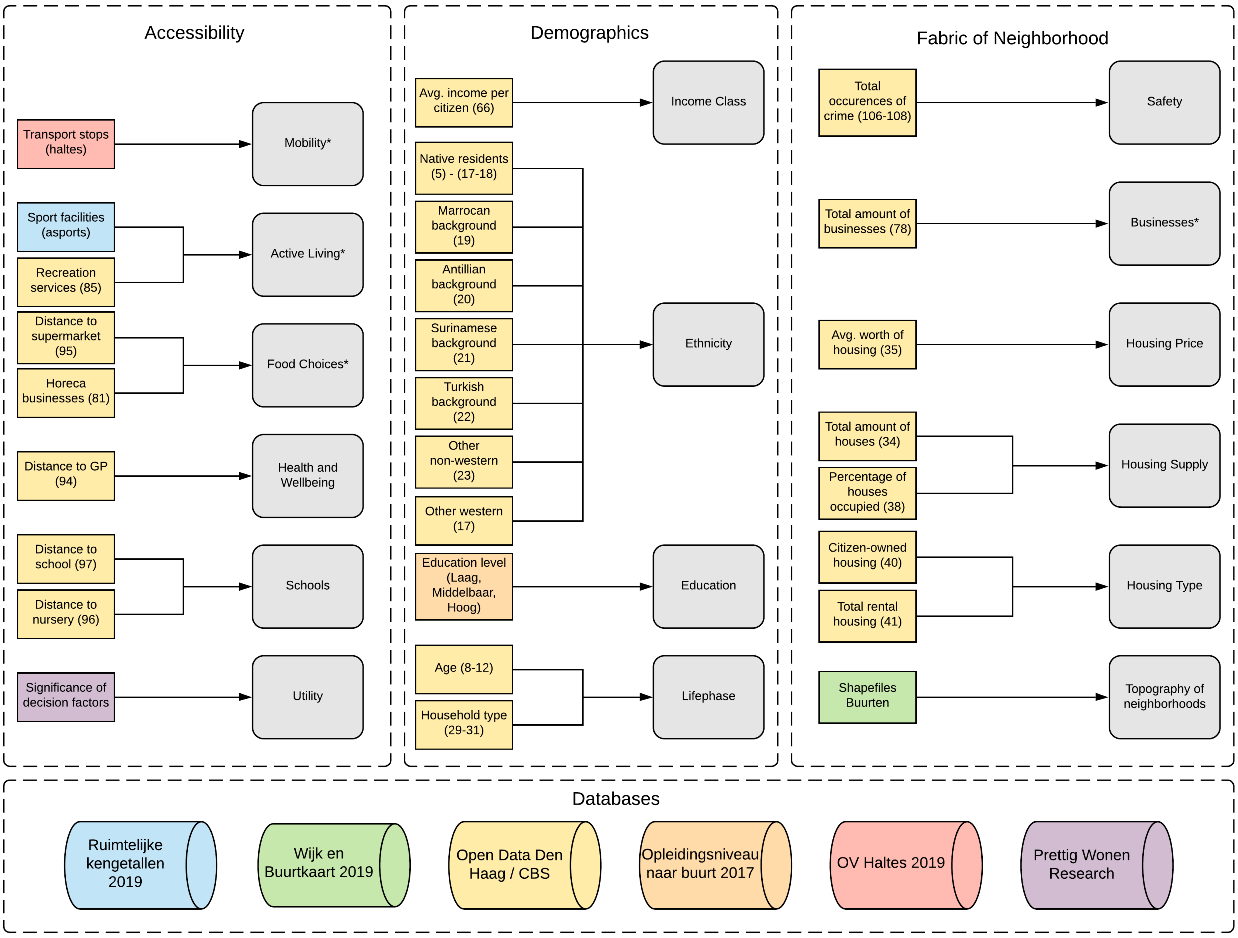


Figure 2.2: Data sources for the inputs and variables in the model. The color shows which dataset delivers each variable. Data sources with an asterisk are currently not used in the model, however can be insightful in further exploration of the data.

neighborhood ("buurt" in Dutch). These "buurten" have a unique number assigned to them which can be used as an identifier. There are in total 121 neighborhood, but some of which do not have residents so only 114 are observed in data. An overview of the neighborhoods that are excluded can be found in the Appendix (Section D).

The next step is to filter data on variables. Only those variables that are interesting to observe for this research will be included. Doing so massively slims down the datasets which in term makes it faster to calculate with. The list of variables chosen to be used are shown schematically in Figure 2.2.

For determining the significance of each factor in the decision-making utility, the results from the *Prettig Wonen* study are used (Center, 2020). To use the data, some calculations had to be performed. First, the data was spread over 5 different Excel sheets, and the information necessary had to first be clustered. After combining the relevant parts of data together into one Excel sheet, the data could be used.

The next step of the process was looking at the neighborhoods that were included in the study, and using available data from CBS, each neighborhood was categorized into one of four groups based on the same categorization of social groups which is introduced in Chapter 3. Doing so, the average of each of the four groups could be made. By this definition, each of the social groups has a unique set of significant factors which can be used in the calculation of utility in the model.

The last step is to translate some of the factors into use-able factors for the model. The most notable change from the study results that is made, is taking the average of all factors relating to safety perception. Doing so aggregates the complexity of this factor, however it does remove some information.

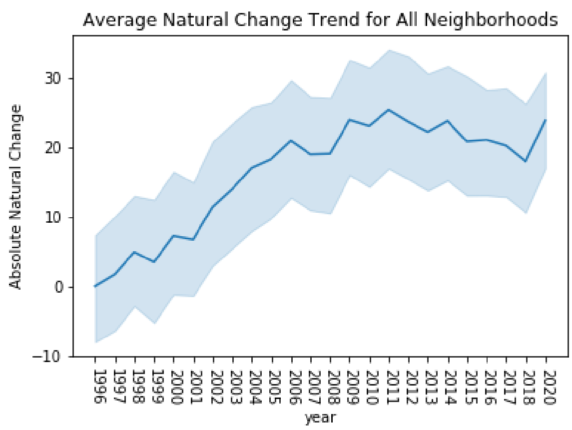
After all data has been prepared for use, the data is exported into separate *.csv* files for use in the model. The different topics and sources of data are separate for ease of use in Netlogo as well as to maintain an overview of what information is used where in the model.

2.3 Changes to the city

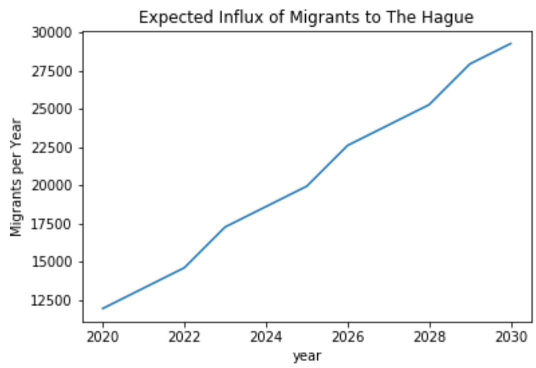
Data is used to model the changes to the city. These changes include housing price, housing supply, average income (inflation), size of influx of migrants, available amenities in a neighborhood, businesses present in a neighborhood, deaths and births in neighborhoods.

Using data from CBS and the open data from The Hague, these changes can be observed for the past ten years. Using these trends is then used to describe a possible future trend using linear regression. This does not necessarily reflect the actual future, and does not account for uncertainties. However, because this uncertainty is out of scope for the model, it is assumed that a linear regression is a fair assumption. Further research into the significance of uncertainty in these exogenous changes might be relevant for future work.

After calculating a trend for the coming years based on data from the past ten years, a new dataset is compiled which is used as an input for the model to define the changes to the city. An example is the natural change, the net result of births minus deaths for each neighborhood, which is shown in Figure 2.3a. Furthermore, Figure 2.3b shows the expected influx of migrants in the city of The Hague.



(a) The average trend of natural change for neighborhoods. To get a perspective on natural change for each neighborhood, the extensive list for all neighborhoods is available in the GitHub repository.



(b) The expected influx of migrants per year. This influx is based on the linear regression of current and historical influx data.

All data and relevant operations performed on the data are publicly available for review, on the Author's GitHub page (Vlug, 2020). On this repository, all data has been imported, calculated and processed using Python. The iPython notebooks are available and include comments for all of the decisions made in the processing of the data [1].

[1] https://github.com/Jochem285/OpenDataDenHaag/tree/master/Python

Chapter 3

Methods

This chapter highlights the methods used in the research. The research question proposed the aim and goal of the research and the sub questions show the missing information that the research aims to find. A research approach is made based on the research question and is part of the Appendix (Chapter B).

The research strategy of this research is the design science strategy, which allows for the creation of a conceptual framework (Bots, 2007; Teddlie & Tashakkori, 2003). Afterwards, this framework is used to make a simulation. Real-world data is used to simulate the behavior of citizens in The Hague.

3.1 Model Overview

After defining the research approach, the model overview is presented after which the conceptual model is constructed. This chapter highlights the different steps of the conceptualisation of the conceptual model, correlation study and creation of the Agent-Based model.

3.1.1 Scope of the System

Using data from open data sources as an input, a simulation model showing interactions between migrants entering a new city and the rest of the city can be made. Doing so will give the opportunity to see how the fabric changes as a result of the influx of migrants into the city. The behavior of citizens will be simulated using decision-making logic based on empirical evidence and scientific literature. Furthermore, the city's "reaction", or the way the fabric changes as a consequence of the inflow of new residents, is based on empirical data. Observing the changes to the fabric of the city is the aim of the research. First, however, making a conceptual model of the intended ABM shows the observed elements of the system.

A city is a place full of interactions and things happening over time. Because it is not possible (or feasible) to simulate everything that is happening in a city, a scope is chosen for defining the system at hand. This is done by creating a causal loop diagram (CLD), which shows the (causal) relation between factors observed in the system (Toole, 2005; Haan & de Heer, 2015). Figure 3.1 shows the CLD for the system observed. Because of the social nature of some factors, it is hard to quantify or even qualify the direction of factors. Because of this, changes to factors such as citizen composition do not have a qualitative direction. Therefore, the CLD does not contain indicators of direction.

When viewing the CLD, it becomes apparent that there is a feedback loop in the observed system. When the composition of citizens changes over time in a certain neighborhood, the composition of shops and amenities also changes. This change in fabric leads to a change in housing availability and pricing, which then leads to a shift of residents that can afford to live in the area. This in turn changes the needs and affordability of needs of citizens living in the neighborhood which in term then strengthens the feedback loop. Common, well studied examples of this phenomenon are gentrification (the change of character a neighborhood faces when citizens of a higher social class start moving in, raising the housing prices) and on the other end urban decay or "urban blight" (when a neighborhood falls behind or into disrepair and only attracts the lowest class of citizens causing

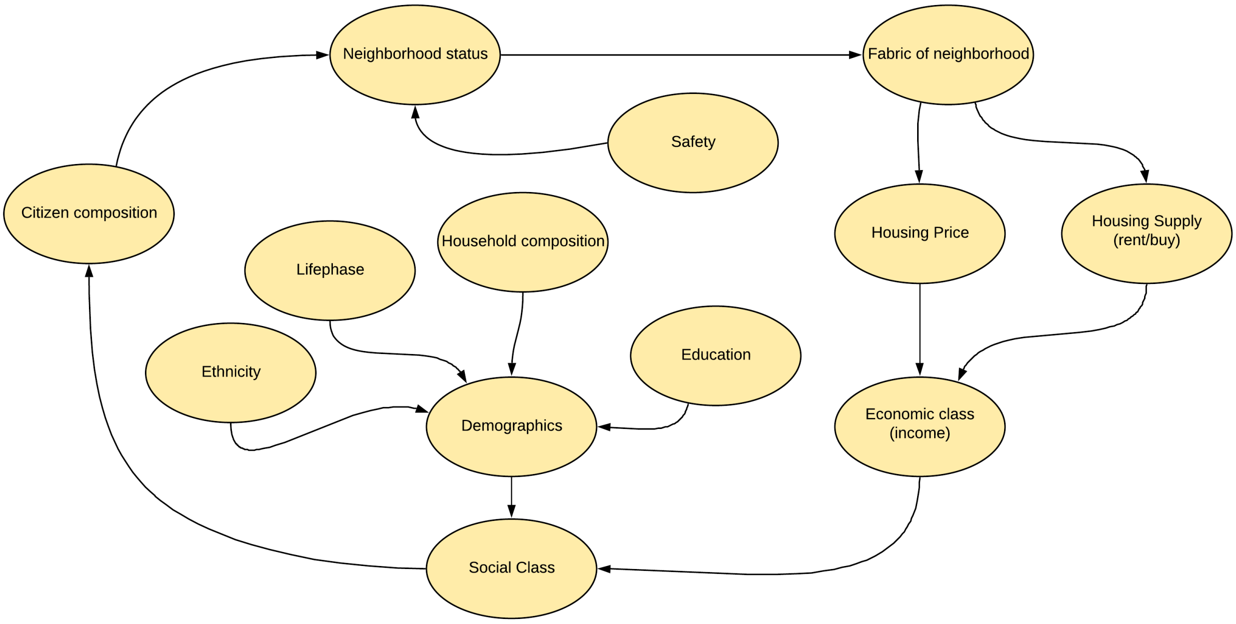


Figure 3.1: Causal Diagram

a "poverty sink") (Andersen, 2019; Schelling, 1978; Zuk et al., 2015; Betancur, 2014).

3.1.2 Conceptual Model

With the scope of the system defined, it is time to establish the key factors that are observed in a model. By creating a conceptual model, the interactions that will be simulated in an ABM are defined. This is done in such a way that it describe the interactions which can be observed in any city, given the scope of the system.

In their paper, Tomasiello et al. defined an Agent-Based model which describes accessibility to jobs of citizens of Sao Paolo, Brazil (Tomasiello et al., 2020). In their research, they defined a conceptual framework of citizens in the city. This framework is generic enough to be used in most cities where data is available. Furthermore, the focus of this framework is on defining the effect of migrations within a city on accessibility to work. This focus on migration can be useful for the research at hand, but lacks an explanation for the decision-making of migrants.

The model assumes citizens have all information and act fully rationally. In practice, migrants do not posses enough knowledge about the city to make a fully informed decision and tend to only look at possible housing options in the vicinity of other migrants (Perez et al., 2019; De Jong & Gardner, 2013). However, the decision-making logic of agents does not change given the rationality of agents. Furthermore, simulating bounded rationality of agents is too complex for the scope of this research.

A literature review by Klabunde and Willekens describes different approaches to building Agent-Based models regarding migration. From this review, it becomes apparent that the best approach for the research at hand is to define utilities for citizens on their moving behavior, since this will bring the most potential to the use of data for quantifying interactions and behavior (Klabunde & Willekens, 2016). However, this utility is not only based on economical maximization, but also takes into account psychological factors such as the inclination to live near people of the same social standings and ethnicity. A good example of these preferences is shown in the research of Perez et al., which defines the factors relevant for migrants moving into the city of Montreal, Canada (Perez et al., 2019).

An adaptation of scientific contributions made by Tomasiello et al., Perez et al. and Klabunde led to the creation of a new conceptual model. This model is presented in Figure 3.2.

The conceptual model is divided into two submodels, which define the agent archetypes of the model. These

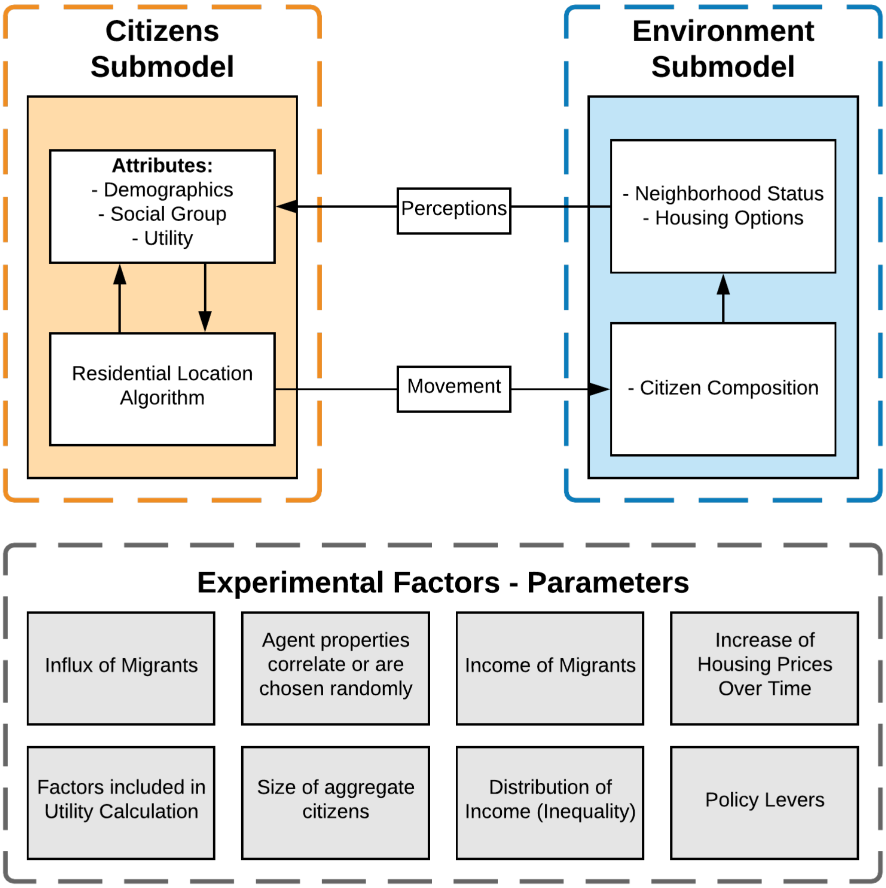


Figure 3.2: Conceptual Model

will be detailed in the following sections of this chapter. Afterwards, the logic of the input, initialization, simulation loop and output are examined.

Citizens submodel

The first of two submodels, is the citizens submodel, which defines the citizen agents of the ABM. Properties of the citizens are gathered using data from demographics research and used to define the population. Using this data, the perceptions of citizens (in a given neighborhood) are estimated. By categorizing citizens on the affordability of their needs, which is based on their ethnicity, economical class and age (for more details, see Section F), the perceptions of citizens can categorized. Given this perception, there are differences in the importance of certain factors for decision-making for each of the social groups. By looking at the factors that important for an agent, the personal utility can be created. The different factors included in the utility function for each social group are based on the pyramid of needs framework from Maslow (Poston, 2009; Maslow, 1943) and derived from research work from the Municipality of The Hague in conjunction with the Central Bureau of Statistics (CBS) (Center, 2020). By looking at the hierarchy of needs, it is assumed that the higher the social group, the more needs a citizen can afford which leads to more factors being included in their decision-making for housing options. The categorization for each social group is shown in Figure 3.3. The complete decision logic for agents per social group is explained in detail in Section H.

The utility function of each citizen agent compares different housing options to select the best alternative. By looking at all the neighborhoods of a city, a "best fit" for housing options can be made. Using this personal utility function is also used to check if the current housing option is (still) the best option. If this is not the case, citizens decide to move to a better housing option. This procedure is explained in more detail in Section 3.1.3.

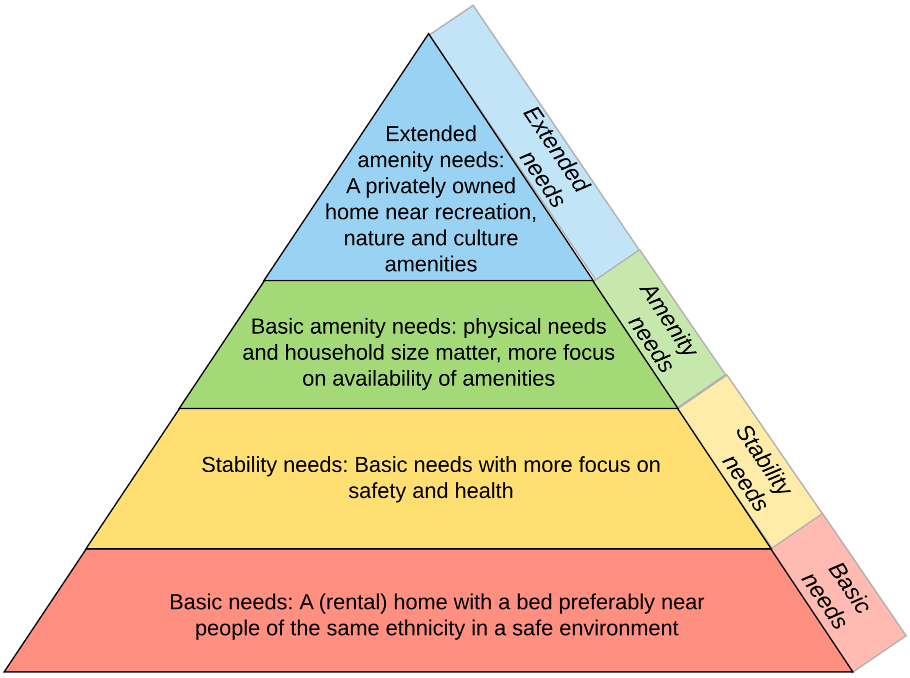


Figure 3.3: Social groups based on affordability of needs.

An important factor for many citizens are the people that live in an area. Research shows people are more inclined to live in a neighborhood with people from the same social group, education, age, background or ethnicity (Perez et al., 2019; Schelling, 1978; Lewis, Emerson, & Klineberg, 2011). Furthermore, research and surveys show that citizens tend to attribute social cohesion as the most important factor of describing livability in the city (Center, 2020; en Cultureel Planbureau SCP., Hart, Knol, Maas-de Waal, & Roes, 2002; Leidelmeijer & Van Kamp, 2003). To simulate this behavior, the average social group, presence of different ethnicities and ages are taken into account. Based on this prevalence, agents will decide whether the neighborhood fits their preferences, or in other words, decide if the neighborhood has the right social cohesion for them.

Research shows that people with low affordability of needs are more likely to live next to, and feel more social cohesion with, people of the same ethnicity or background (Perez et al., 2019; Center, 2020). However, citizens who can afford more needs care more for social cohesion by finding individuals of the same social group and cultural diversity has a positive feedback on social cohesion (Sturgis, Brunton-Smith, Kuha, & Jackson, 2014; Center, 2020). Because of this, it is assumed that only the lower social groups look for the same ethnicity and higher social groups look for neighbors of similar social grouping without looking for similar ethnicity.

Because of the ability to afford certain needs, the different social groups have different preferences for certain amenities and not all are taken into account when looking for housing options for all of the social groups. The higher the social group, the more important the role of the accessibility to amenities becomes. The accessibility to these amenities is measured using a function based on the work of Orozco, Deritei, Vancso, and Vasarhelyi and Nicoletti and is shown in Figure 3.3.

After calculating the accessibility to amenities in a neighborhood and the status of the neighborhood, the utility function is then used to to define the residential location algorithm, which calculates the best housing location the citizen agent can afford. Because of the great impact of life events, such as marriage, divorce, graduation or retirement, these life events define the decision-making for migration (Courgeau, 1985). To simulate this effect, agents in the model will only move when a life event occurs or when the phase of life they are in changes. Furthermore, statisical analysis has shown a correlation between age and probability of moving homes (CBS, 2019). This follows criterion 4 in the review paper of Klabunde and Willekens (2016). After deciding to move, agents that make these movements ultimately cause changes in the demographics and consistency of neighborhoods and thus change the fabric of the city. Over time, the availability of amenities and neighborhood status is observed as the main Key Performance Indicator (KPI), called the fabric of neighborhoods. By observing the movements (and especially the movements as a result of the influx of migration) and resulting changes to the fabric of the city, the impact of movement can be described. To observe changes to the city landscape, we define the second submodel, which looks at the city itself.

Environment submodel

The second submodel is the form of the city, or the environment. This includes housing prices, housing availability, types of housing, safety and occurrence of crime in a neighborhood and availability of urban amenities (Mobility, Active living, Food Choices, Nature, Schools , Health and Well-Being). Based on the work of Orozco et al. and Nicoletti, several important types of amenities have been identified. Surveys show the importance of safety (and the perception of safety), which is included as an attribute that is "present" to a certain degree (Center, 2020; en Cultureel Planbureau SCP. et al., 2002; Leidelmeijer & Van Kamp, 2003).

Each neighborhood observed in the model thus has a list of available amenities, presence of safety (or the lack of crime), availability of houses (per housing type), average housing prices, average income of its citizens, percentage prevalence of ethnicities and education.

This information is key to the decision-making of citizens, and results in movement behavior of agents. Furthermore, the mentioned factors can be measured to describe correlations between certain groups of people, certain type of behavior or decision-making and certain changes to the city fabric. This is done by experimenting with the input data and observing changes in the output.

Experimental factors

Lastly there is a set of parameters present in the model that can be varied to create alternative scenarios and experiment with the interactions and behavior of the model. Doing so will test the robustness of the model, as well as give a better indication of the possible explanations for observed interactions and outcomes (Nikolic et al., 2019; Tomasiello et al., 2020). It should be noted that the goal of making an Agent-Based Model is not to predict the outcomes of the future; it is to better understand interactions and observed behaviors (Feitosa, Le, & Vlek, 2011). The parameters are shown in the bottom part of Figure 3.2.

- The influx of migrants can be changed to simulate an increase flow of migrants (as a result of external effects such as civil war in a foreign country or changes to migration policies).
- The factors included in the decision-making process of citizens can be altered.
- The properties of agents can be either randomly distributed or dependent on each other.
- Citizens can each be modeled as individual agents, however, to reduce computational strain, agents can be aggregated into a group of similar citizens or households.
- The distribution of income of citizens can be adjusted by increasing or decreasing the standard deviation value of the normal curve.
- The prices of houses in the city (or a specific area) can be increased over time as a simulated result of market changes or price surges.
- The average income of migrants can be adjusted.
- Multiple policy levers can be tested in their effectiveness of guiding the change of fabric of neighborhoods.

The reasoning for altering these experimental factors is to discover the underlying factor that influences certain interactions or behavior. Furthermore, some of the experimental factors are bound to regional research; for example, case studies from The Hague show a big significance of social cohesion and a negligible significance for the prevalence of supermarkets and food choices whereas data from Chicago almost shows the opposite significance (Center, 2020; Nicoletti, 2020).

3.1.3 Model procedure

After defining the conceptual model, and defining the elements that are part of the model, it is time to define the model procedure. This procedure can be perceived as the "story" or narrative of the model and is used to give an overview of the interactions and mechanics involved in the model and simulation thereof (Nikolic et al., 2019). The procedure of the model can be divided into four parts; Input, Initialization, Simulation, and Output. An overview of the full procedure is shown in Figure 3.4, outlining the full model procedure.

Input

By using data from CBS and Open Data Den Haag (of The Hague, 2020), the inputs for the citizen agents are formulated. Furthermore, the input for the environment can also be generated. The process of gathering data and preparing it for model use is described in the Data Chapter (Chapter 2).

Initialization

After the data has been prepared and entered as inputs into the model, the agents and their initial location are defined. First, the map and layout of the observed areas are loaded into the model. This is done using GIS or shape files. Then, the neighborhoods are initialized as agents of the same class or "breed", with fixed locations based on data. The neighborhoods retrieve their respective properties such as average housing price, population size, availability of housing and many more from the data that has been initialized in the Input phase.

Next, the citizens of each respective neighborhood are imported into the model. This is done by checking the properties of each neighborhood and using the properties to define the amount of citizens that are spawned into each neighborhood. By looking at the aggregation size parameter in conjunction with the population data, the amount of citizens is determined. The properties of citizens is then determined in either a randomly distributed fashion or based on correlation logic. The first manner checks the properties of the neighborhood and uses them as distributions for a random draw. For example, if a certain neighborhood has a population size of 100 households and 5% of its inhabitants are of Turkish ethnicity, the random distribution assumes that 100 citizens will be created, each have a 5% chance of being Turkish. This is then done for all the relevant demographics of citizens.

The second manner in which citizens can be created at the start of the model is by assuming correlations between certain properties. If this approach is used, the initialization of citizens correlates to properties. For example, if an agent is created with a high income, the model assumes there is a higher chance of this citizen owning a home instead of renting. Similarly, the model assumes citizens of lower income and education to have a small chance of owning a home, and rather will be renting for a place to stay.

After the initialization of the city map, neighborhoods and its citizens is completed, the model is then ready to perform the last step of the initialization phase, which is to add an initial influx of migrants. The model assumes that at the start of each year, a new influx of migrants enter the city. For consistency's sake, this is also the case in year 0 (at the end of initialization).

To finalize the initialization, all citizens (including migrants) recalculate the social group they should be part of based on their age, income and education. After this, the model is ready to start the simulation phase.

Simulation

The main logic loop is presented in the simulation phase. The most important aspects of the simulation loop are discussed in this section. However, a complete and detailed overview of all the procedures and mechanics within the simulation can be found in the model narrative Appendix (Chapter H).

As Figure 3.1.3 shows, at the start of the simulation loop (which is every simulation *tick*) the citizen agents are asked if they are currently deciding about moving to another location. As discussed in Section 3.1.2, it is assumed that agents only move when a life event occurs in their life. Using a logit implementation (Klabunde & Willekens, 2016), it is calculated for each citizen if a life event is currently happening, and the agent decides to move.

If this is the case, the citizen agent reviews all locations which have housing availability and checks if they are within its price range. Using this short list of locations, the agent then checks if the locations provide higher utility than its current housing location. If there are no locations that can provide higher utility then their current location, the loop ends and the citizen agent does not move.

If there are locations that can provide a higher utility than the current location, the locations that meet this requirement are compared. The location that can provide the highest utility for this citizen agent will become the chosen location for the agent to move towards.

Now the citizen agent moves from one location to another. As a result, the status of both the location the agent leaves as well as the location the agent moves in to are recalculated since there is a change in citizen demographics for those locations. If the change in citizen composition leads to a change in average social group or income of the location, the average housing price is recalculated as a result of housing market dynamics.

After completing the simulation loop for one agent, the same is done for all other citizen agents within the model. Because the chance of a life event occurring are relatively low, this means that during the simulation of a *tick* only a small part of agents move.

After all citizens have decided to move or not move, the simulation *tick* is concluded by advancing time with 1 *tick*, or 3 months of simulation time. During this phase, the percentage of social groups and ethnicity are recalculated for each location based on the changes from the simulation. This is done after the movements to minimize computations.

Once every year (or 4 simulation *ticks*), a new influx of migrants is simulated entering the city. Based on migration data, the amount of migrants is determined (Statline, 2020a).After all agents have finished looking for suitable locations to move to, the simulation recalculates housing pricing and income of locations based on the changes of citizen composition. It is assumed that a rise in average income in a location results in a rise in average housing price as a result of housing market dynamics (Gaube & Remesch, 2013).

Output

The output of the model is generated at the end of a simulation loop, at which the citizen composition for each location is recalculated. Furthermore, once every 4 *ticks*, the housing price and average of citizen statistics is recalculated.

Using the citizen composition and demographic statistics such as income and social group are used to then define the location's neighborhood status. The fabric of the current location can be calculated using the housing price and availability, the safety of the neighborhood and availability of amenities. This output can be later analyse to see if there are correlations to be observed between changes to the citizen composition and city fabric. This is done in the data analysis phase of research, highlighted in Chapter 4.

3.2 Model Implementation

This chapter describes the implementation of the model. Using the framework of the conceptual model, the Agent-Based model can be built. First, the agents in the model are defined. Secondly, the behavior of the model is described using a model narrative in which the steps agents take and the interactions between agents and other agents or the environment are defined in a narrative style (Chapter H). Afterwards, the model implementation is documented with highlights to certain elements of behavior in the model. After the model has been implemented, it is tested using verification and (qualitative) validation in Section 3.4 and Chapter 4 (Nikolic et al., 2019).

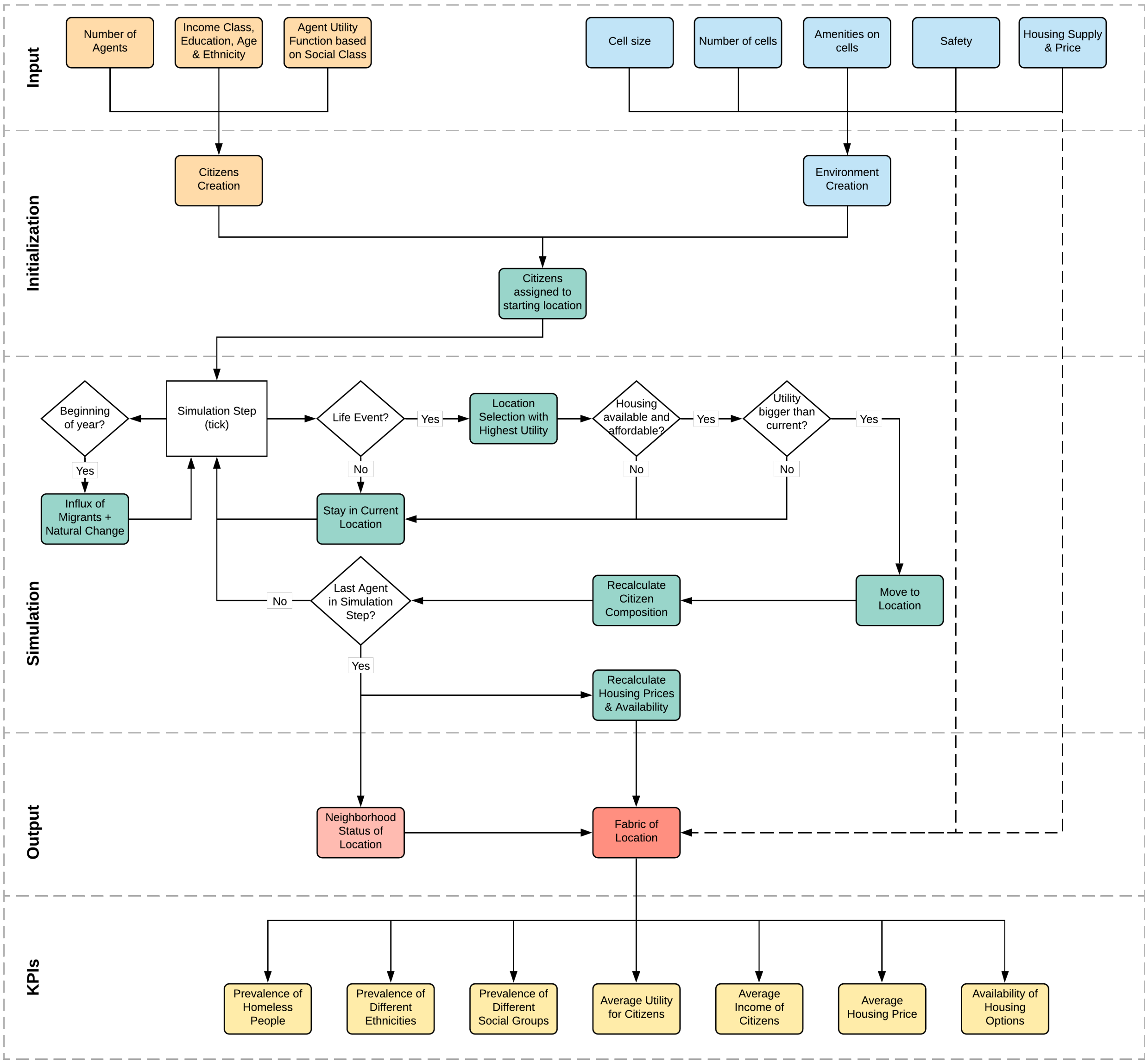


Figure 3.4: Procedure of the model.

3.2.1 Agent properties

First, the agents and their respective properties are defined. Doing so helps understand what information each agent has and which information or resources it needs to perform certain actions or interactions. Each agent has a list of properties, and the relation between each agent and their properties is shown using a Unified Modeling Language (UML) model, showing attributes of agents and enabling an easier translation of concept to model (Bauer & Odell, 2005). The UML diagram is shown in Figure E.

3.2.2 Agent Decision Logic

A complete overview of the complex mechanisms behind the interactions of agents in the model is described in the model narrative, in the Appendix (Chapter H). This section highlights the most important feature of the model narrative, the decision-making process of the citizen agent. Figure 3.5 gives an overview of the decision-making logic of citizen agents.

In the flowchart, the different factors included in the decision-making of citizen agents are shown. The factors that are included in the decision-making process are internal and external to the properties of the agent. In other words, to align with the framework as defined in the conceptual framework (Section 3.1.2), the agent assesses properties from himself (citizens submodel) as well as its observed location's environment (environment submodel).

Based on a citizen's affordability of needs (social group), certain factors are either included or excluded in the decision-making process. This is highlighted in more detail in Section F. Based on their affordability of needs, the agent might just only check neighborhoods for people of the same ethnicity but if the affordability of needs allows for it, might also include a wide range of needs such as amenities, safety and social grouping needs.

By checking neighborhoods for the availability of these factors, the utility of each neighborhood is calculated. This is done by multiplying the different factors (which are normalized) with a "significance factor", based on the relevance of each factor calculated based on survey work (Center, 2020). Then, each factor (with its multiplication) is then summed to make up for a total utility score. This is done for every potential location. Afterwards, all possible locations are compared and the location with the highest utility is chosen as the location to move to. In the case the highest possible utility is in the location the citizen already lives in, it is assumed the agent does not need to move at this moment and no movement occurs. An overview of the decision-making is shown in Figure 3.5.

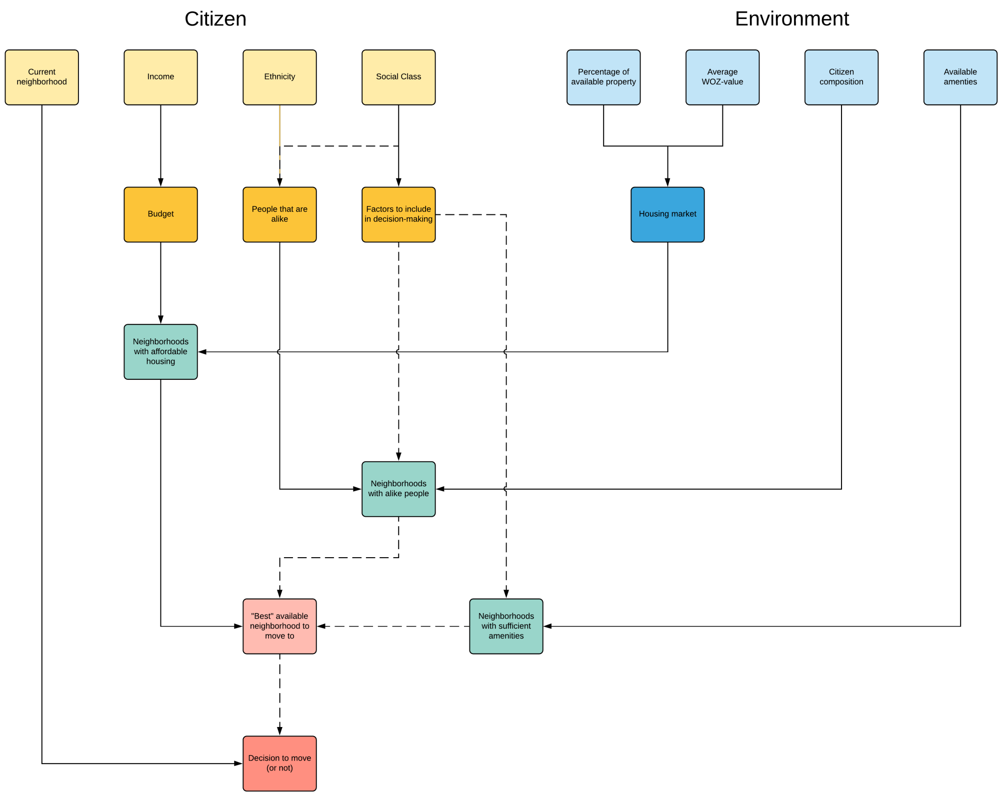


Figure 3.5: Decision Logic for citizens deciding to consider moving.

3.2.3 The Agent-Based Model

After defining the framework of the model using the conceptual model, the narrative and procedures of the model in the model narrative and decision logic, the last step of the model implementation is to translate these concepts to working simulations using coding language in software.

The chosen computer program to code and simulate the ABM is Netlogo. This fully open-source software runs on all Operating Systems and has little to no requirements to operate (Wilensky, 1999). Research and modeling in complex scientific context (such as modeling human behavior) should take into account the "5 Rs": Re-run, Repeat, Reproduce, Reuse, Replicate (Benureau & Rougier, 2018). By making sure both the data as well as the software are free to use and open access, the "5 Rs" can be approached.

The choice for Netlogo to construct the ABM is threefold: the ease of use and overview the user interface brings, is useful for explaining the complex concepts and behavior observed in the model to policy makers. Secondly, past experience of the author has led to a skill set using this software which is greater than the capabilities in other software packages. Lastly, Netlogo offers a very robust experimentation toolkit with regards to the stochasticity and reproducibility of *logit* calculations. Because of the well defined and robust stochastic engine, it is easier to validate if observed behavior in the model can be addressed to randomness/stochasticity or to interaction of the agents in the model (Wilensky & Rand, 2015).

When building the model and code, one should be aware of the limitations of simulation software. The goal in making a model is not to reproduce a clone of the real-world city, but to simulate certain characteristics and interactions observed in the real-world. As Box puts it: "All models are wrong, but some can be useful". In other words, models do not have to replicate the exact real-world, as long as the modeler is aware of the goal of the model.

Furthermore, there are computational limits to what is feasible in modeling. Modeling all citizens of a city as individual agents is not always necessary (and sometimes even downright impossible), and can slow down computations a lot. When working in Netlogo, one should be aware of certain filter operations causing more strain on the simulation, and should try to avoid over-using filters whenever possible (Railsback et al., 2017).

An overview of the implementation of the model is shown in Figure 3.6, where the *dashboard* user interface is shown during a run. All the graphs and plots alongside the model simulation are used to monitor the interactions of agents and to verify that the intended behavior can be observed. This process is called the verification of the model and is explained briefly in Section 3.4. It is explained in more detail in Section 3.4.

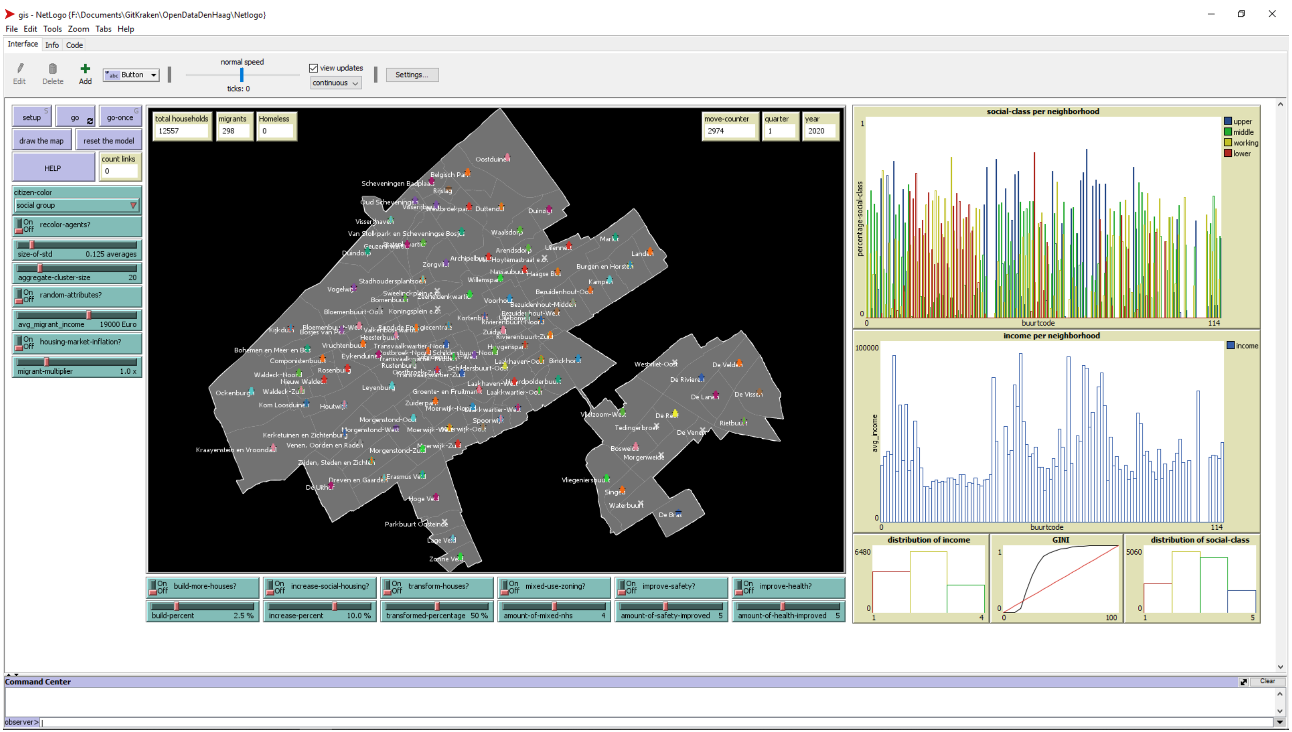


Figure 3.6: Overview of the dashboard view of the Agent-Based Model in Netlogo. An explanation of all buttons, graphs, plots and levers can be found in the Appendix. The run simulated in this overview has an aggregation of 1:20, hence the number of citizen is scaled down by a factor 20.

3.3 Policy Implementation

The first goal of the model is to explore the dynamics of the city and the impact of migrants on the fabric of the city. The second goal is to provide policy makers with insights in the behavior of the system with regards to policy analysis. By defining policy levers that are relevant for policy makers, the impact of policy interventions can be observed. In the model, several policy scenarios have been included as policy levers which are tested in the experimentation phase of the research. The outcomes of these scenarios are discussed in the results Section (Chapter 4).

It should be noted that the focus of experimentation on policy scenarios is to better understand the dynamics and describe the possible changes to the fabric of a city given certain policy alternatives. Because of this, the proposed alternatives that the municipality might undertake are only modeled in their implementation. In other words, the financial costs, feasibility and time constraints of the proposed policies are out of scope of the model.

3.3.1 Current alternatives

The municipality of The Hague is aware of the expected influx of new citizens and is preparing the city to be able to cope with the growth of citizen-count. Three alternatives have been presented as possible "routes" to take to cope with the increasing number of people moving to the city: 1) increase the amount of available housing options (build new houses), 2) transform empty/dilapidated housing properties into modern (smaller) housing properties, 3) mixed-use or re-purpose of business and industrial zones (van Amsterdam et al., 2015). All three alternatives will be simulated in the model to test their respective effectiveness. Apart from the alternatives presented by the municipality, two extra alternatives are added that try to improve the fabric of the city by increasing the availability of amenities that provide health and safety respectively.

3.3.2 Building More Houses

Building more houses is the most straight-forward alternative for increasing the amount of available housing options in a city. However, this alternative comes with multiple downsides. Primarily, building is expensive. To make sure new housing options live up to modern standards, fit in the current building style, are sustainable enough to be future-proof and are big enough to accommodate family growth the costs of building new houses are very high. On top of that, with new regulation on the emission of certain poly- and perfluor-alkyl substances (PFAS), the amount of new buildings allowed to be constructed is limited (van Deen, Voskamp, & Bezuijen, 2018). Furthermore, a second limiting factor is space. Because of laws and regulations regarding the public space, there are very little areas left which can be used to build new houses (Den Haag, 2016). Because of these limiting factors, and the fact that building new houses takes time, this alternative is modeled in such a way that its impact is relatively small.

Build New Housing Options

In the model, when this policy scenario is observed, every year the top 5 of neighborhoods with the least amount of available housing are determined and they build 2,5% new houses in that year. This assumes that building is always possible within the neighborhoods with the least available homes and thus disregards building, space, PFAS, budget and regulatory restrictions. Because this policy ignores these restrictions, the impact of building new homes is limited. An increase of 2,5% might entail an increase of 0-200 housing options, depending on the size of the neighborhood.

Increase Social Housing Availability

People that are most affected by a scarcity of housing options are the citizens with the lowest incomes. By increasing the availability of social housing, more people with lower incomes will be able to find a suitable home and divert from a life living on the streets. The municipality can increase social housing options by either enforcing regulations or buying suitable properties and using them for social rent. In the model, turning on this policy lever will increase the availability of social housing by 5% whilst decreasing private rent by 5%.

3.3.3 Transform Existing Housing Into Smaller Properties

Another option that has been mentioned by the municipality as a possible alternative is the transformation of current houses to smaller studio or apartment housing (van Amsterdam et al., 2015). This would mean that the municipality would buy houses and transform them into multiple (smaller) rental properties. In the model, this scenario assumes that the neighborhood with the least amount of available rental housing space will "lose" 50% of available housing that is for sale, whilst gaining the same amount in both private rental and social rental properties (by assuming each bought house can be divided into 2 rental properties).

3.3.4 Re-purpose Industry/Business Zoning

The last proposed alternative for providing more housing options to citizens, is to re-purpose industrial and business zoning into mixed-use (van Amsterdam et al., 2015). In short, this means areas where currently only businesses are located, will allow for business properties to be transformed into housing. To implement this in the model, we simplify the real-world issues of space, location and environment issues that could arise when looking for suitable places to transform into housing. This means, that the modelled alternative assumes that when this policy is active, half of the "neighborhoods" that currently do not have housing options will start building houses. This will increase with respectively 60 houses for sale, 140 for rent and 200 for social rent per year.

3.3.5 Increasing Health Amenities

The fabric of the city might also be improved by changing other factors that influence the livability of a neighborhood. Two alternatives for policies are proposed and modeled: increasing health amenities and improving safety in the neighborhood.

First, the overall perceived health score can be improved by increasing the amount of General Practitioner's offices in the area. A higher availability of healthcare in the local area, and thus a higher amount of GP offices per citizen, has a positive effect on health and perceived health of the neighborhood (Groenewegen, Kerssens, Sixma, van der Eijk, & Boerma, 2005). This policy attempts to improve health in neighborhoods by increasing the amount of GP offices in a neighborhood. Each year, this policy provides the 5 neighborhoods with the worst healthcare ratings with an increase of doctor's offices of 10%.

3.3.6 Improving Safety

Similarly to the improvement of the healthcare, the perceived safety of neighborhoods is something that can be improved with governmental spending. By increasing the presence of police, neighborhood managers or community workers, the (perceived) safety of a neighborhood can be increased (Pitner, Yu, & Brown, 2012). Furthermore, community projects can help increase the perceived safety such as neighborhood watch. All these initiatives require municipal funding and are thus aggregated into the same policy alternative.

In the model, the policy alternative is modeled as a way to decrease the average amount of crimes occurring in a neighborhood. When this policy is active, every year, the worst 5 neighborhoods implement a reduction to the amount of crimes occurring. Because the implementation can vary in effectiveness, this is modeled with a random effectiveness between 0-15%.

3.3.7 Change Social Housing Prices

As a control variable, the social housing prices of all neighborhoods can be altered to see how it influences the poorest citizens of the city. By parameterizing the social housing price (either lowering or raising the price), the effects of the social housing price on the decision-making of citizens can be observed. The most relevant KPI to observe in this instance, is the amount of homeless people in the city.

3.4 Model Verification

To make sure the model performs as intended, verification of the model is necessary. Since modeling a complex system as observed in the real-world is prone to mistakes and errors, it is important to verify the model (Nikolic et al., 2019). Verification of the model is needed to make sure the model was built as intended and to make sure no unintended behavior may explain certain results. In short, the verification process is a helpful tool to verify the integrity of the model.

The full verification steps that were taken to make sure the model runs as intended, are shown in the Appendix (Chapter I). Here, we can see that all programmed behavior and inputs of data and decision logic are a product of intended model implementation.

Chapter 4

Results

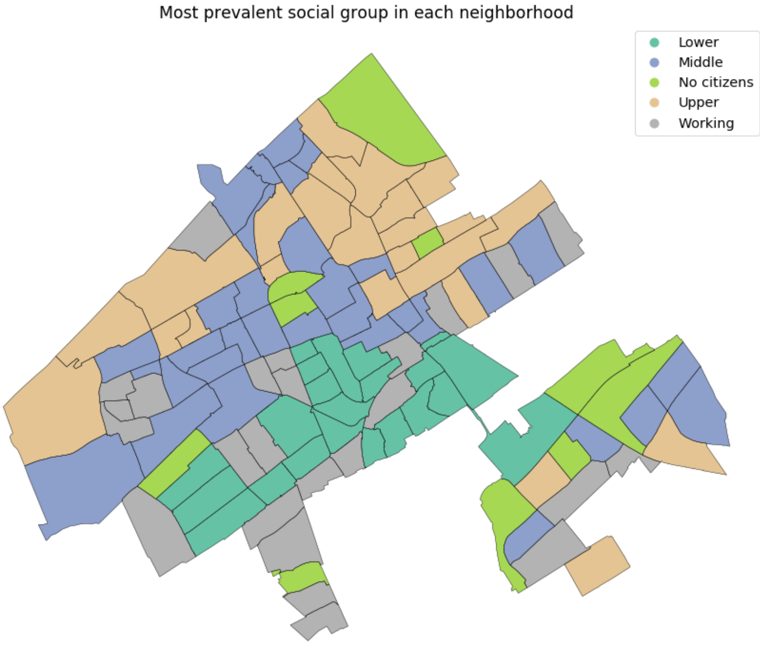
After designing the conceptual model (Section 3.1.2), gathering data (Chapter 2), defining the model narrative (Chapter H), designing the Agent-Based model (Section 3.2.3), verification of building the right thing (Chapter I) and running the simulation (Chapter J), the outcomes of the Agent-Based Model are ready to be observed. From the resulting output data, behaviors within the ABM simulation can be described and used to describe changes to the fabric of the city.

When running simulation experiments, repetition of experiments is needed to verify integrity of results and make sure it is not caused by a random seed (Wilensky & Rand, 2007). Because of this, the model (and all of its subsequent policy levers and experimental parameters) were simulated 32-fold. More details of this process and the resulting outcomes are shown in Chapter J. Because of the computational strain of simulating the model many times, time constraints and the occurrence of issues during the first iterations of running the simulation, aggregation was used to simulated agents in the system. This means that each agent in the model represents 10 households of the real-world equivalent in The Hague. To make sure the resulting output was not affected by this simulation decision, a sensitivity analysis shows its (negligible) impact on the results (Chapter L).

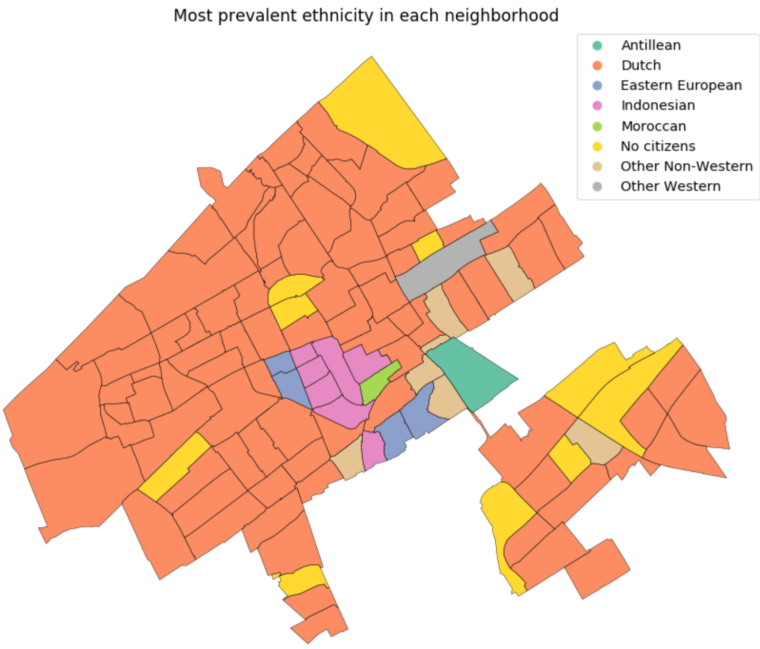
This chapter highlights the most important results that were observed from the data, and attempts to explain or describe the behavior that results in the observations. Furthermore, those areas of the model that showed the most significant behaviour (changes) are observed in more detail in the sensitivity analysis (Chapter L). The underlying structure of analysis follows the proposed research flow described in Section 1.9 (Figure B.1 outlines this structure). First however, to get an idea of the output space, the data is categorized and an overview of the data is made.

4.1 Result Data Overview

Since the model simulates over 260.000 agents as citizen households, it is not feasible to show results of each individual agent. Furthermore, to describe the behavior of the model (and the changes to the fabric of the city), aggregation is needed to better define what it is that is changing. There are several aggregation methods, such as aggregation on district level or geolocation. Furthermore, it is also possible to aggregate by defining neighborhood clusters of citizen composition using the highest prevalence of social group for each neighborhood. The map in Figure 4.1a show the different types of neighborhood based on the most prevalent social groups in that neighborhood. Similarly, Figure 4.1b shows a map of most prevalent ethnicity of each neighborhood.



(a) Map showing the categories of neighborhoods, which are defined using the most prevalent social group in each neighborhood at the end of the baseline simulation.



(b) Map showing the categories of neighborhoods, which are defined using the most prevalent ethnicity in each neighborhood at the end of the baseline simulation.

The model simulation runs cover 40 time steps, simulating 10 years of changes to the city (2020-2030) whilst gathering information on all 114 neighborhoods in The Hague and the over 260.000 households living there (more information on the neighborhoods is given in Chapter D). Furthermore, different policies and parameters have been tested to see how interactions change given certain inputs. Finally, each unique simulation scenario of the model is tested 32 times to make sure the stochasticity of the model does not interfere with the results. Lastly, there are 22 Key Performance Indicators that are observed to check the performance of a simulation run (for more information, see Chapter K). In total, this means there are over 5 million data points resulting from running the simulation.

4.2 Relation Between Migration Patterns and Changes to the City

As the XLMR framework shows, the input of the model with regards to exogenous or external factors can have a significant influence on the behavior of a model (R. Lempert, 2002; R. J. Lempert, 2003). To describe the relation between migration patterns and changes to the fabric of the city, the influence of these external inputs needs to be analyzed.

To analyze the input space, multiple experimentation parameters (as proposed in Section 3.1.2) are present in the model which are varied in value to see how the system behaves given a change in parameter space. Such variables include the amount of migrants arriving in The Hague, the average amount of income of migrants, the average amount of deviation in citizen income, the assignment mechanism of citizen properties and lastly the aggregation size of households into a single agent in the model.

4.2.1 The Parameter Space

To test the influence of the variability of values for migrant properties on changes to the fabric of the city, multiple parameters have been tested. Table 4.1 overviews the parameter space, the variation of values for each of the parameters tested within the ABM. The normal values used in the model are based on statistics from CBS and the municipality of The Hague (Koot, van Elk, & Jongen, 2019; Huijnk, 2016). The parameter space observed in the simulation run is broad and bidirectional, since the expected sensitivity of parameters is unclear. By testing each of the variables separately, we make sure to rule out interference of effects from multiple parameters (Saltelli & Annoni, 2010). The most significant observations from the exploration of the parameter space are part of the results chapter, details within the parameter space are explained in Chapter L.

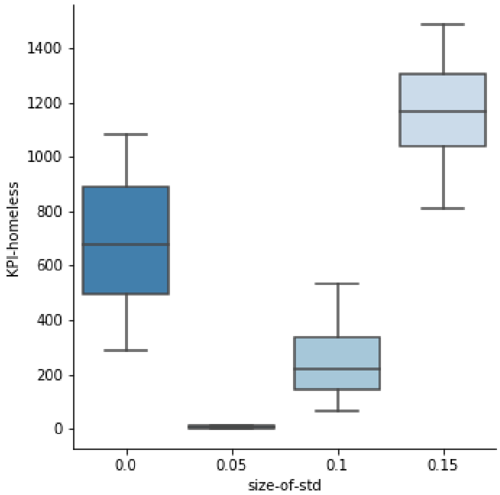
Parameter	Variable	Metric	Values	Default Valu
P1	Standard deviation of income distribution	Multiple of average	0.00, 0.05, 0.10, 0.15	0.05
P2	Multiplier of expected influx of migrants	Multiplier expected influx	0.5, 1.0, 1.5, 2.0	1
P3	Average spendable income of migrant	Thousands of Euros / year	15, 17, 19, 21	19
P4	Randomly-assigned attributes	Boolean	yes, no	no
P5	Housing market and income inflation	Boolean	yes, no	no
P6	Size of aggregation of citizen agents	# Households / agent	1, 5, 10, 100	10

Table 4.1: Parameter space of experimental parameters for testing sensitivity of agent properties.

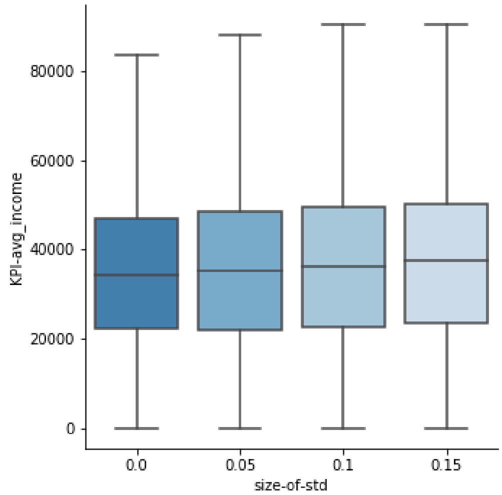
For each of the identified parameters an analysis is done to see which variables have an impact on the model behavior, and thus can say something about the changes of the city fabric. This analysis tries to show the implications of parameter value changes by looking at statistical differences, impact on a geospatial level, comparing changes in citizen composition and looking at change over time.

4.2.2 Outcome Differences from Income Inequality

Data on the average income of citizens are imported to the model for each neighborhood. However, no real-world data is available for the distribution of this income. Thus, it is assumed to follow a random distribution. Because the value for the standard deviation of income is unknown, the impact of varying this value has been tested. The value for the deviation of income can be interpreted as a proxy for wealth (in)equality, as a higher standard deviation means the spread of incomes from the average is bigger.



(a) Homeless citizens given the spread of income distribution. Both a perfectly equal income distribution as well as high inequality lead to homelessness.



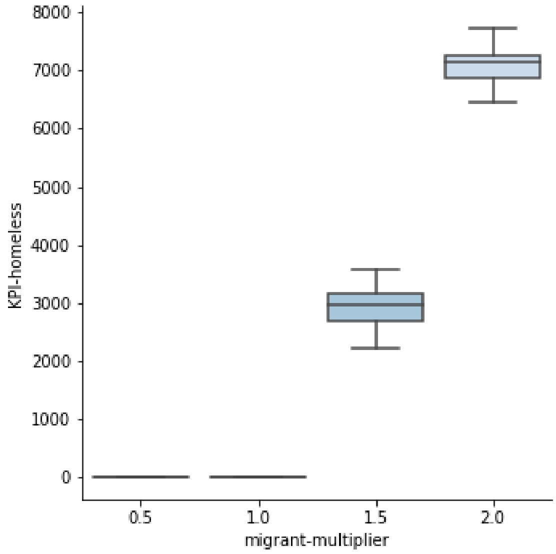
(b) Average income in neighborhoods given the spread of income distribution. More inequality leads to the disappearing of poorer citizens, thus rising the average income.

When all citizens have exactly the same income as the average of each neighborhood, an interesting phenomenon occurs. In this instance shown in Figure 4.2a, more people become homeless because all citizens compete for the exact same houses with the exact same budget. When there is some (minor) differences in budget, these housing problems seem to disappear. However, when inequality gets bigger, more citizens struggle to find a suitable home. At the biggest inequality, many citizens at the low end of the income spectrum have such a low income they cannot afford any housing option and thus turn homeless. This leads to a misrepresentation of citizens in neighborhoods, as poor people are no longer living in neighborhoods. This effect becomes visible when looking at the average income measured in neighborhoods (Figure 4.2b).

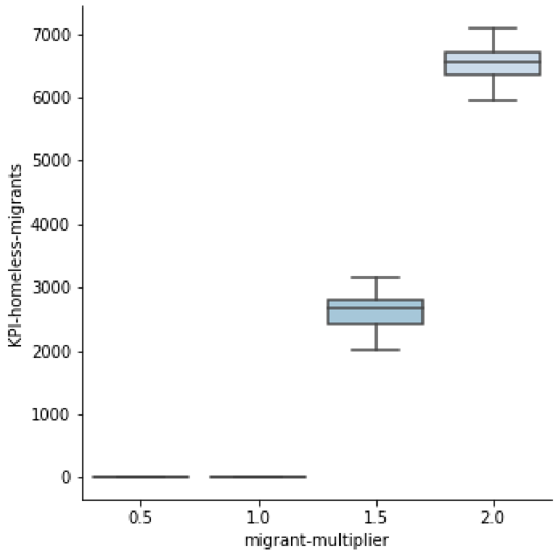
4.2.3 Outcome Differences from Influx of Migrants

The goal of the model is to get insights into the changing of the city fabric given the influx of migrants. There are data sources for the influx of migrants of the past 10 years which are used as a basis for describing the expected influx of migrants in the future (by using linear regression). However, as discussed in Chapter 1, there is much uncertainty in the expected amount of migrants entering the city in the future. To cope with this uncertainty, the changes to the city fabric are observed for different amounts of migrants entering the city, to get a better understanding of the impact of different amounts of incoming migrants. This is done by multiplying the expected number of migrants by a value between 0.5 (to assume less migrants will come in the future) and 2 (to assume double the amount of expected migrants will arrive in the near future.

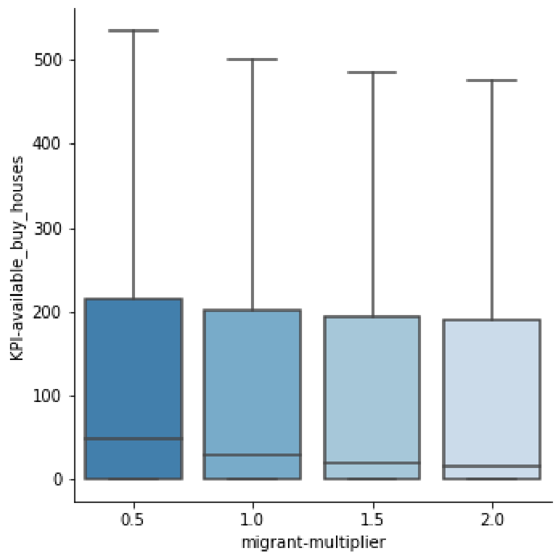
Figure 4.3 shows the significant changes to the model output given differences in the value for the migrant influx multiplier. As expected, an increase in migrants arriving in the city leads to an increase in homeless migrants, since there are not enough houses to accommodate all migrants (Figure 4.3a and 4.3b). As the amount of migrants entering the city increases, the supply of housing options decreases (as shown in Figure 4.3c, but similar behavior can be observed for all three sectors). When more migrants enter the city, it also influences the average composition of neighborhoods. As there are proportionally more migrants from a lower social group, the prevalence of this group increases when more migrants enter the city (Figure 4.3d).



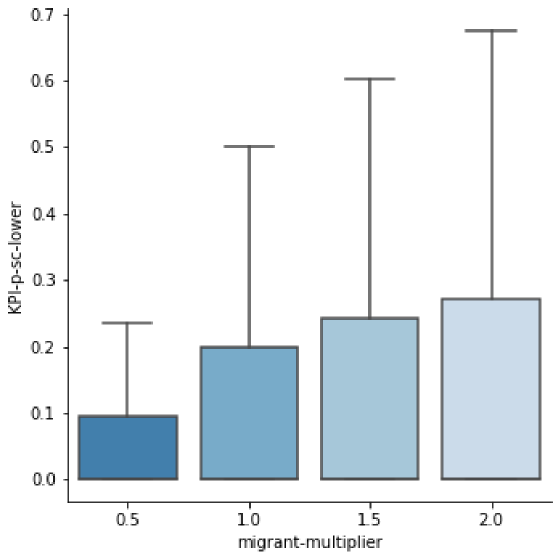
(a) Total amount of homeless people



(b) Amount of homeless migrants



(c) Available houses for sale

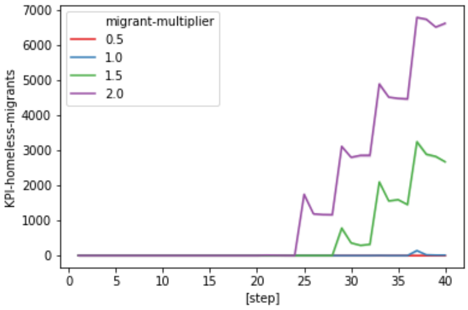


(d) Percentage prevalence of citizens in lower social group

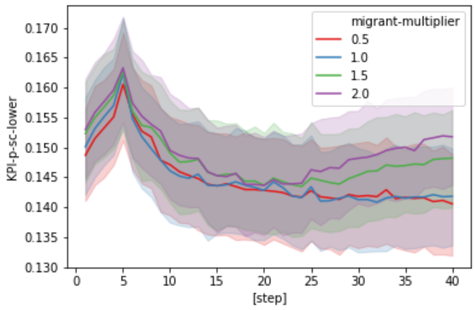
Figure 4.3: Boxplots showing significant differences in outcomes of KPIs when testing the sensitivity of the multiplier for expected migrant influx.

Time-wise, the effect of a bigger influx of migrants can be observed in a timeline plot. Figure 4.4a shows the increase in homelessness over time, the cyclical pattern that can be observed here is due to the fact that the model simulates the influx of new migrant on the first time step of each year which creates a spike in the graph. The decrease in homelessness after each spike can be attributed to some of the migrants succeeding in finding a home after multiple tries (which is possible due to the movement of other citizens in the system and due to changes in hte availability of houses due to natural change within neighborhoods).

Figure 4.4b shows the development of average prevalence of citizens of the lower social group in neighborhoods. When looking at the graph from time step 20 on wards, there is a clear correlation between the percentile rise in lower social group and the amount of influx of migrants. Because most migrants have lower education and income, they are more present in the lower social group. The rise in the first 5 time steps correlates to a fall in prevalence of the middle social group, and might be the result of citizens shifting towards other neighborhoods and therefore changing the overall average of prevalence for all neighborhoods. In other words, the composition of many neighborhoods is more homogeneous in time step 5, and becomes more heterogeneous again over time. This effect seems to be unrelated to the influx of migrants however.



(a) Homeless citizens given the influx of migrants over time.



(b) Percentage prevalence of lower social group citizens over time.

4.2.4 Outcome Differences from Income of Migrants

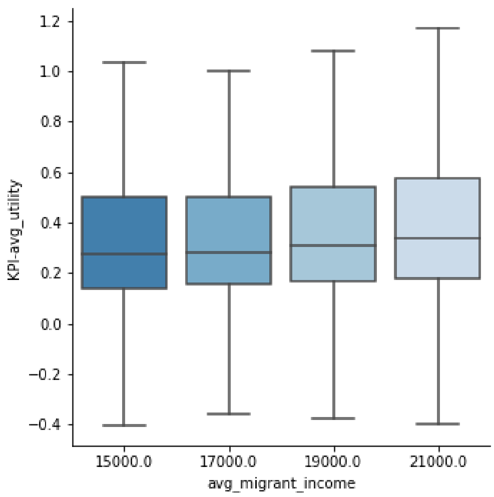
Apart from looking at the expected amount of migrants entering the city, the properties of migrants is also of importance. The most important factor is the average spendable income of migrants, as this will determine the possible housing options for migrants.

Unfortunately, there is little data available on the distribution of income of migrants arriving in the city. Because of this, assumptions have to be made with regards to the distribution. There are statistics on distribution averages in relation to country of origin which is used as an indicator for the default value for income of migrants (Koot et al., 2019; CBS, 2019). The assumed default income value is set to 19.000 Euros. To test the sensitivity of income, three more values are tested. The resulting outcomes are shown in Figure 4.5.

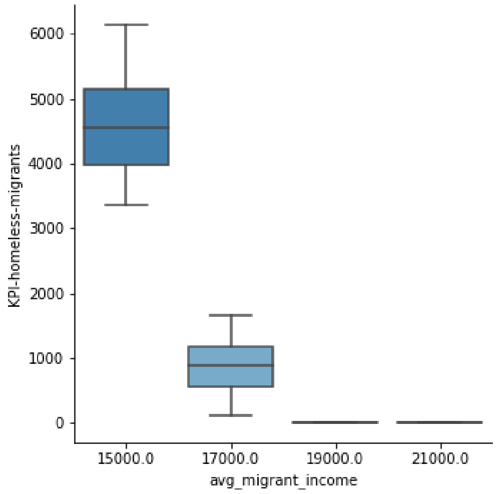
As Figure 4.5b shows, there seems to be a break-off point for migrant homelessness. When the average income of migrants is higher than 17.000 Euros per year, migrants are almost always able to find a home and are not forced to live on the streets. Furthermore, an increase in average income directly correlates to a decrease in availability of rental options in the private sector (Figure 4.5c), which is due to the affordability of migrants to rent more expensive houses. Similar outcomes can be observed when looking at the social rental housing options (Figure 4.5d), but when average income surpasses 19.000 Euros, migrants start looking more into private rent instead of social rent options.

Another interesting observation is the increase of average utility (Figure 4.5a). When migrants have more income to spend, they can afford more suitable (or nicer) housing options which increases the overall utility of all citizens.

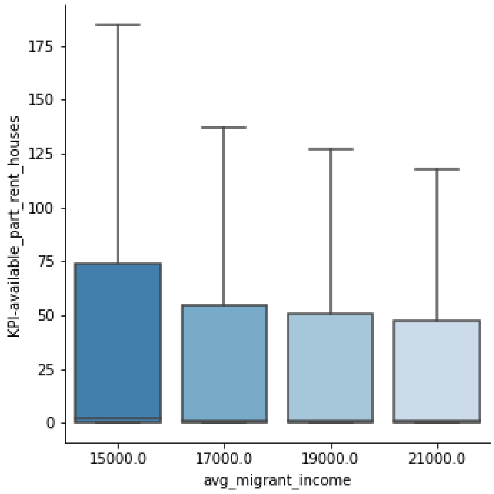
Lastly, an increase in average income among migrants leads to a decrease in the prevalence of citizens in the lower social group (Figure 4.5f). This can be addressed to the calculation of social groups, where higher income groups tend to group in working, middle or upper social groups. Thus, less migrants with low income leads to less people in the lower social group.



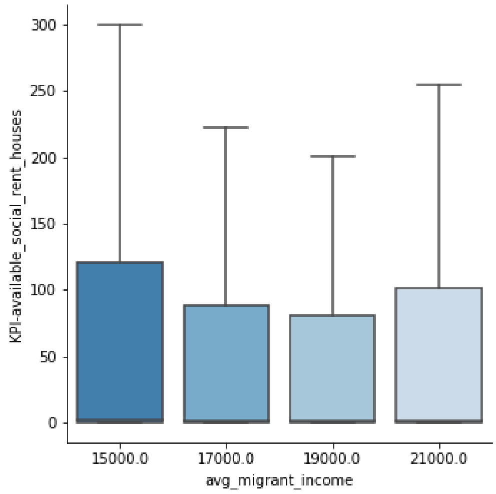
(a) Neighborhood average utility of citizens



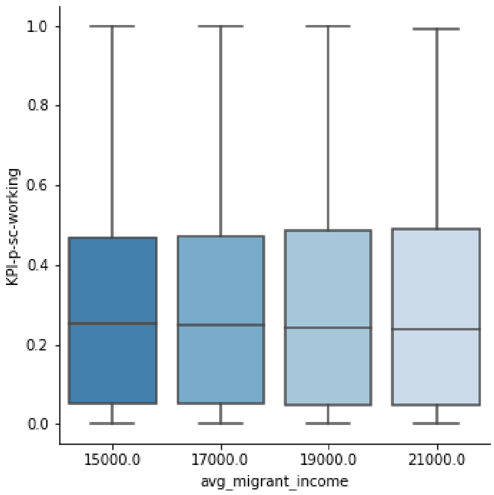
(b) Amount of homeless migrants



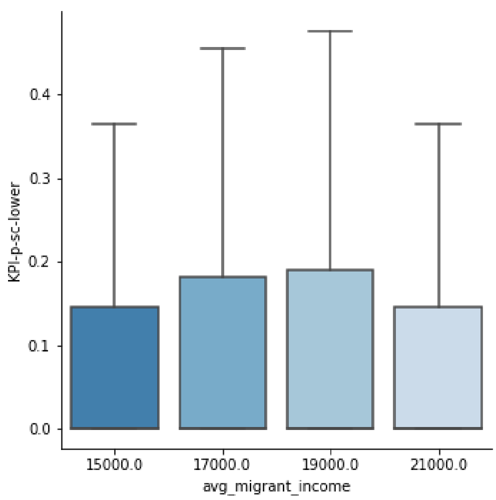
(c) Availability of houses for rent in the private sector



(d) Availability of houses for rent in the social sector



(e) Percentage prevalence of citizens from working social group

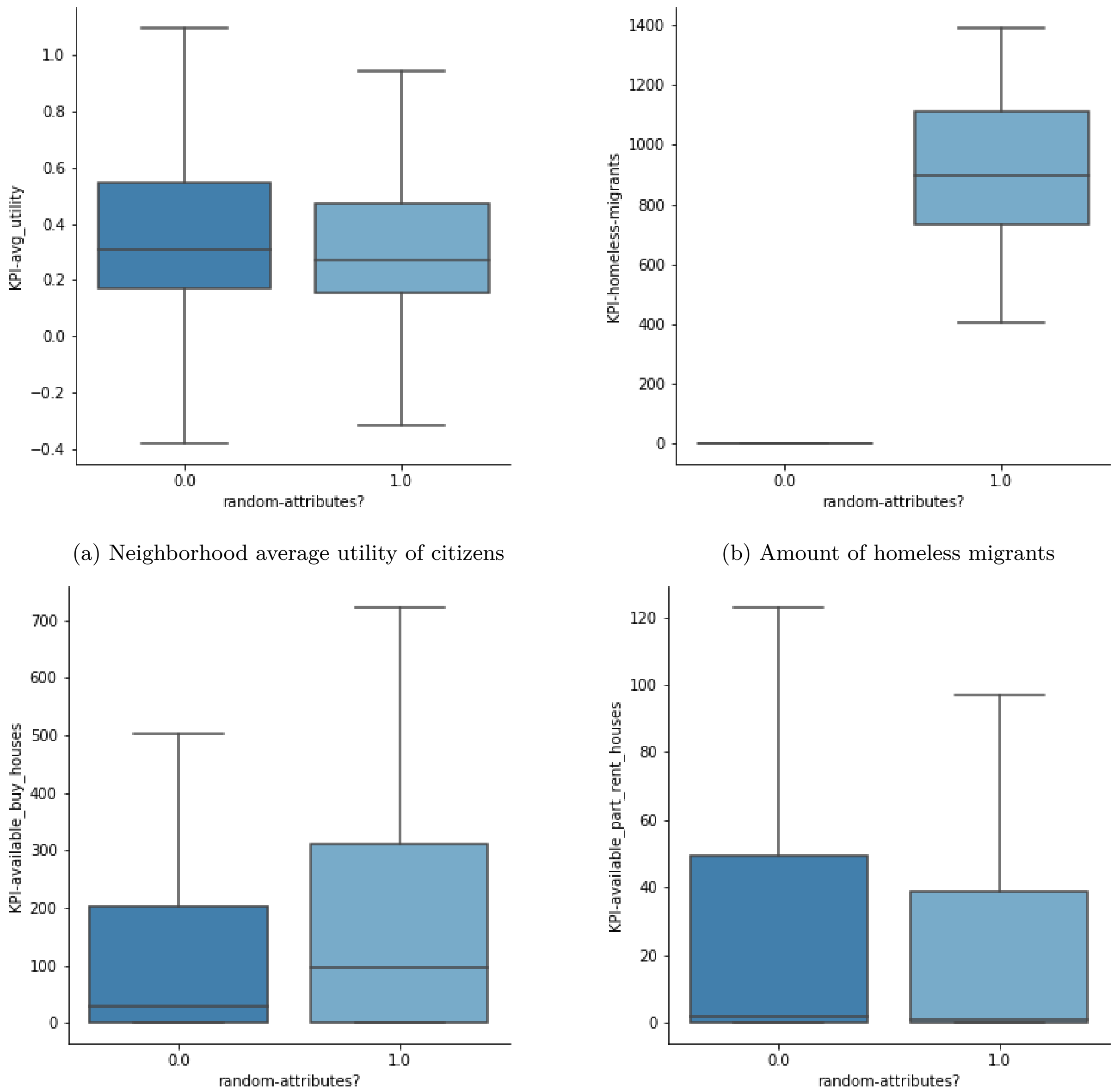


(f) Percentage prevalence of citizens from lower social group

Figure 4.5: Boxplots showing significant differences in outcomes of KPIs when testing the sensitivity of the multiplier for average income of arriving migrants.

4.2.5 Outcome Differences from Randomized Agent Properties

As discussed in Chapter H and Chapter F, some properties of citizens in the model are defined based on other properties of the agent. An example of this is the relation between income and education. Because the correlation between properties relies on assumptions of citizen composition (see also Chapter G), the influence of this modeling decision is measured. The way the model simulates agent properties (dependently or randomly) can only be "switched on" or "switched off", and thus the experimentation space for this parameter is a Boolean. Since there are a lot of differences observed from changing this parameter, the plots showing the differences in outcomes are drawn in both Figure 4.6 and Figure 4.7.



(a) Neighborhood average utility of citizens

(b) Amount of homeless migrants

(c) Amount of available houses for sale

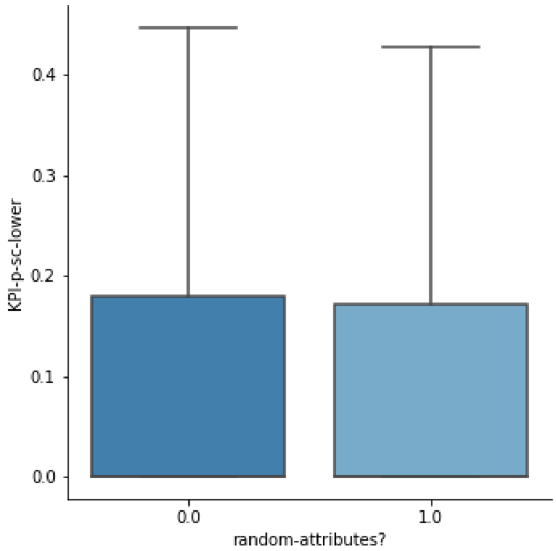
(d) Amount of available private sector housing rent options

Figure 4.6: Part 1 of Boxplots showing significant differences in outcomes of KPIs when testing the sensitivity of the random assignment of agent attributes.

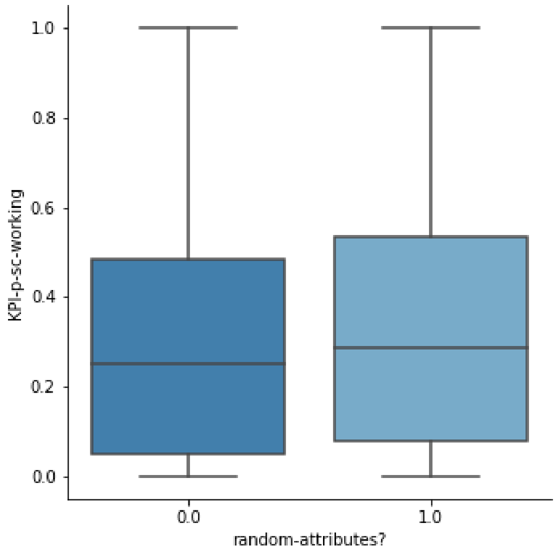
The average utility of citizens seems to be negatively impacted by the random distribution of properties. When properties are not correlated, citizens have a harder time adjusting their needs to their budget. For example, in the real-world, there is a correlation between income and education. Educated people tend to have more needs (which they can afford with a higher income). However, as Figure 4.6a shows, if the attributes are assigned randomly, needs are met less often.

Looking at Figure 4.6b, a positive trend can be observed when attributes are randomized in the case of homelessness. When agent properties are randomized, more "unfortunate" combinations of properties might occur.

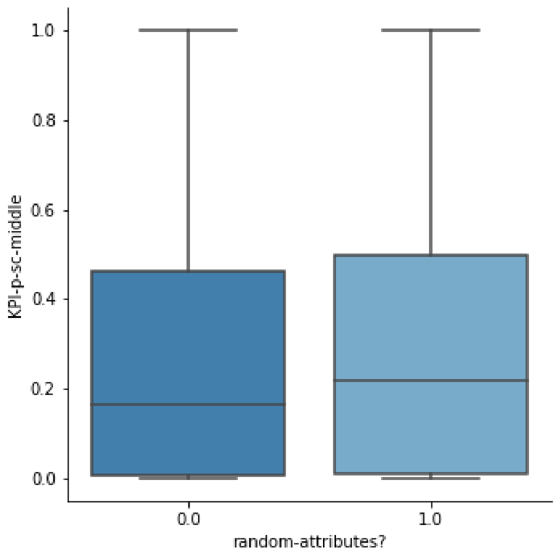
Furthermore, when properties are assigned randomly, more houses that are for sale are left available (Figure 4.6c) whilst the amount of available rent houses decreases (Figure 4.6d). This can be explained due to the fact that higher educated citizens tend to look for houses to buy, but because of the random assignment of properties do not always have an income to support this wish. At the same time, because of this unfortunate selection of properties, more citizens tend to resort to rental housing options as an alternative.



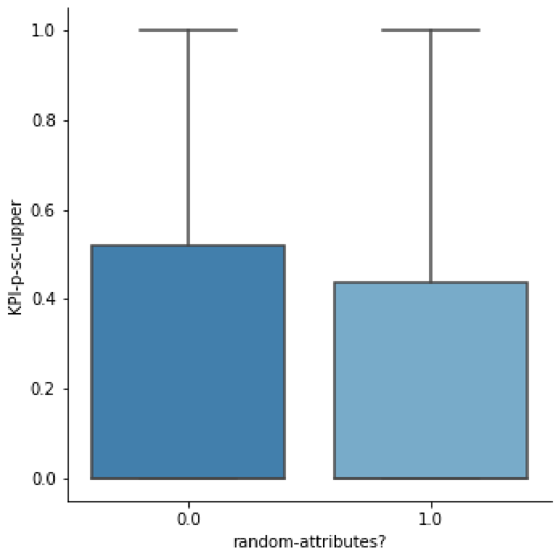
(a) Percentage of people in the lower social group.



(b) Percentage of people in the working social group.



(c) Percentage of people in the middle social group.



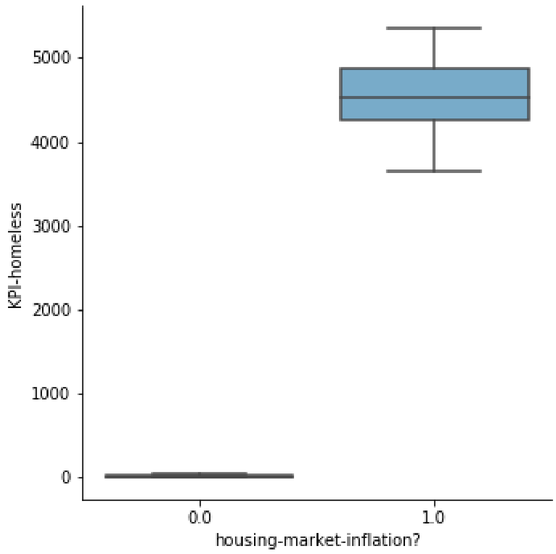
(d) Percentage of people in the upper social group.

Figure 4.7: Part 2 of Boxplots showing significant differences in outcomes of KPIs when testing the sensitivity of the random assignment of agent attributes.

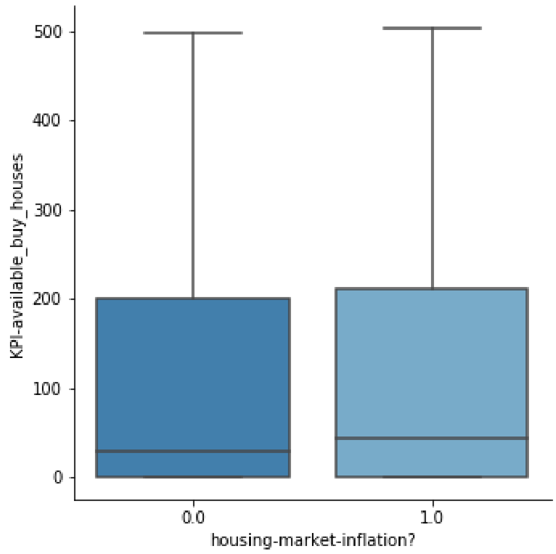
There is also a significant change in citizen composition when looking at the percentage prevalence of each social group when the attributes of citizens are selected randomly (as seen in Figure 4.7). This can easily be explained when looking at the way the social group of citizens is calculated (Chapter F). Because education and income no longer relate to each other, a rise in working group and middle social group can be observed, whilst at the same time a decrease in upper social group can be seen, as the combination of both high income and high education is now more rare.

4.2.6 Outcome Differences from Inflation and Price Change

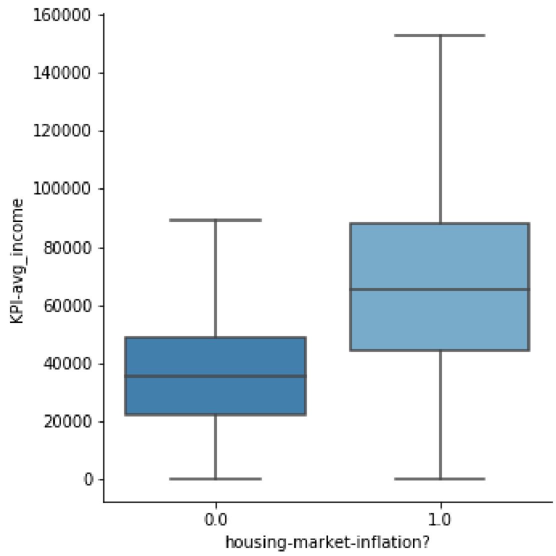
An exogenous factor that is not included in the model by default, but might play a role in real-life and thus in the model, is inflation. Income increases over time as a result of inflation. However, when looking at the trend of average housing market prices, an upward increase in average housing price can be observed as a "market reaction" to a shortage of supply (van Amsterdam et al., 2015; Määttänen & Terviö, 2014). Since there is a shortage of housing options in the *Randstad*, the supply and demand are not in equilibrium (R. C. Kloosterman & Lambregts, 2001; van Amsterdam et al., 2015). To add this effect to the simulation, a parameter was added that models an increase in housing price whilst at the same time modeling an increase in income (due to inflation). The increase in housing price is simulated as a steady increase of 7% per year (based on the current trend in housing sale and rental price increase (van de Statistiek, 2020b, 2020a)), where inflation increases income by 3-10% (with an average increase of 6%). The resulting changes to the outcomes are shown in Figure 4.8.



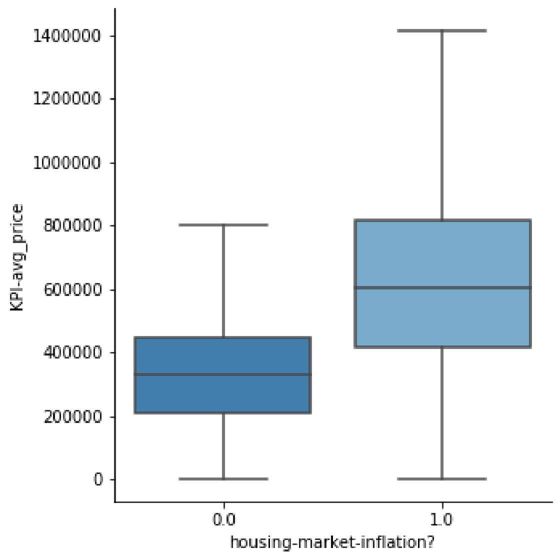
(a) Total amount of homeless people



(b) Availability of houses for sale



(c) Income distribution



(d) Housing price distribution

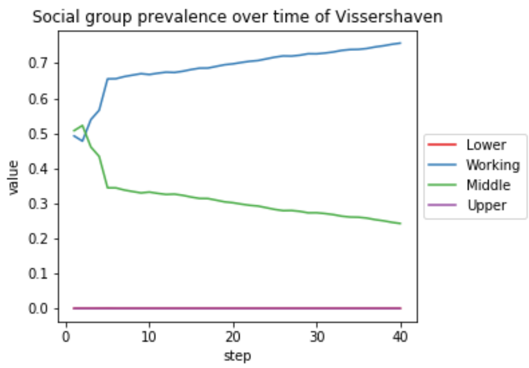
Figure 4.8: Boxplots showing significant differences in outcomes of KPIs when testing the sensitivity of the inflation mechanism in the model.

As expected, Figure 4.8c shows an increase in average income since the inflation increases income over time. Similarly, the average price of housing options also increases as shown in Figure 4.8d. More surprisingly, Figure 4.8a shows that the increased housing prices and income also lead to significantly more homelessness. This might be explained by the fact that the range of affordable housing options for migrants is already small (as shown in Section 4.2.4) and any change in housing market prices causes them to lose the ability to afford housing. This hypothesis is further confirmed when looking at the availability of houses for sale (Figure 4.8b), which rises because less people are able to afford buying a home.

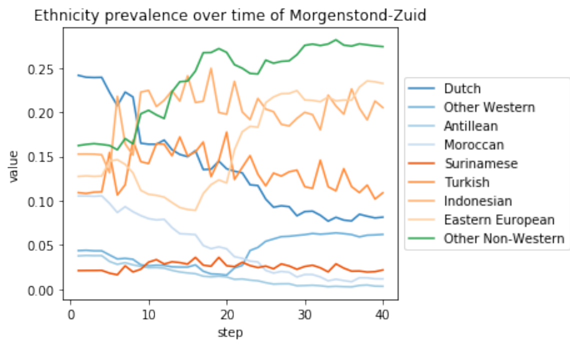
4.2.7 Spatial Differences

The outcome data from the parameter space can also be observed in a spatial sense. Looking at differences in the outcomes between neighborhoods given the parameter space, can give a better indication of not only how the fabric of the city changes, but where it happens. To get a grasp of the differences between the outcomes, maps are used to show the outcomes. Furthermore, because of the amount of neighborhoods, clustering of neighborhoods is used to get an indication of a group of neighborhoods at the same time.

Two trends can be observed in many neighborhoods which can be explained by the decision-making of citizens. First, some neighborhoods which have many citizens of the same social group, tend to push out neighbors of other social groups. An example of this behavior can be seen in Figure 4.9a. Secondly, the prevalence of ethnicity over time seems to follow one of two patterns: either the prevalence of different ethinicities is more or less stable and does not change (significantly) over time, this seems to be the case often when the biggest ethnicity present is Dutch. The other possible pattern that is observed in data, is foreign ethnicity "competing" on prevalence in a neighborhood. This behavior shows season-like movement behavior of citizens of the same ethnicity moving in and out of a neighborhood, which leads to a constant change of the most prevalent ethnicity. An example of this pattern becomes apparent when looking at the neighborhood "Morgenstond-Zuid", in Figure 4.9b.



(a) Dynamics of social group prevalence over time for neighborhood 2: Vissershaven.



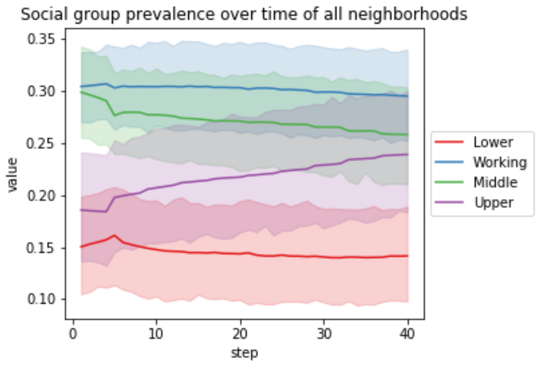
(b) Dynamics of ethnicity prevalence over time for the neighborhood of "Morgenstond-Zuid".

Figure 4.9: Dynamics of two neighborhoods observed over time in the baseline scenario.

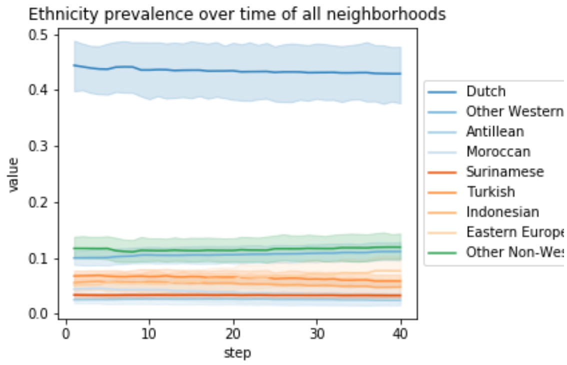
4.2.8 Temporal Differences

Lastly, the differences in the parameter space can be observed from a temporal perspective. By looking at the change in city fabric over time, a better notion of the interactions within the model given certain parameters can be described. By using time frame statistics and graphs, the temporal aspect is shown.

For instance, the composition of citizens living in neighborhoods can be tracked over time. This can give a rough indication of the prevalence of citizen group types, but can also reveal a better explanation for the observed dynamics. Figure 4.10a shows the prevalence of social groups over time. Here, we can see a clear "reaction" or correlation between the rise of one group in relation to the decline of another. An example is the increase of lower and upper social groups whilst at the same time a decrease in working and middle social groups can be observed, at time step 5. This "pushing out" behavior can explain why the fabric of certain neighborhoods tends to change over time.

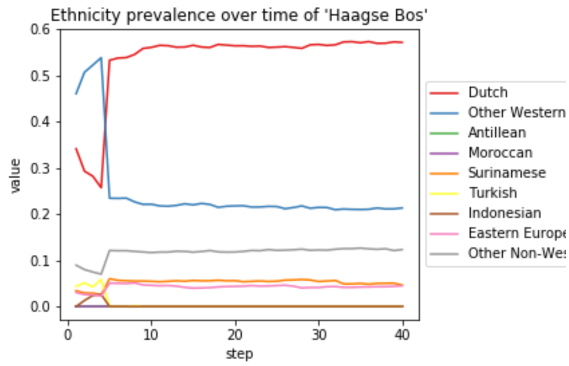


(a) Dynamics of social group prevalence over time for all neighborhoods in The Hague.

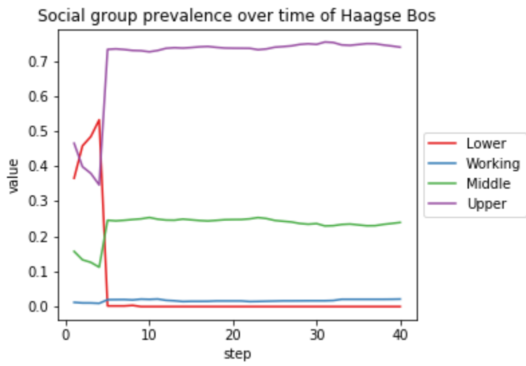


(b) Dynamics of ethnicity prevalence over time for all neighborhoods in The Hague.

Another example of these dynamics can be observed when looking at the prevalence of ethnicities within neighborhoods. Figure 4.10b shows this development over time. Interestingly, the "pushing out" of one ethnicity over another is barely present when looking at the aggregated data of all neighborhoods at once. However, when looking at one single neighborhood, this effect can still be observed, for example for the neighborhood "Haagse Bos" in Figure 4.11a and 4.11b.



(a) Ethnicity prevelance dynamics over time for neighborhood 49: "Haagse Bos". Here, a wave of new Dutch citizens "push out" those of Other Western ethnicity.



(b) Prevelance dynamics of social group over time for neighborhood 49: "Haagse Bos". Here, by rapidly increasing housing prices, the lower social group gets pushes out by the upper social group.

4.2.9 Exogenous Factors Changing the City Fabric

The impact of exogenous factors is the system is sizeable and very significant. As this section shows, the exogenous effects that change the fabric of the city are: the amount of migrants coming to The Hague, the amount of spendable income migrants have, the distribution of income of citizens (and thus the amount of inequality between citizens) and the rise in housing prices as a result of high demand or low supply of housing options. Many of these exogenous factors are not only uncontrollable by intervention, but also have a fair share of uncertainty in their future status. To cope with this uncertainty, the sensitivity analysis shows a range of possible outcomes when looking at an array of values for the aforementioned parameters. Some key observations are listed.

Looking at the influx of migrants, the expected amount of new people migrating when extrapolating current and historical data seems to have an effect on neighborhood composition. Over time, the citizen composition of neighborhoods with many (former) migrants, or people from lower- and working social groups, tend to attract migrants. Over time, this can lead to the "push out" or "pull out" of other citizens, which causes a neighborhood to become homogeneous in citizen composition. The benefit of homogeneity is an increase in social cohesion, the downside is an increase in equality between different neighborhoods which can ultimately lead to an increase in crime or neglect of a neighborhood.

Furthermore, the analysis shows that migrants need a spendable income of at least 17.000 Euros per year to be able to afford housing, and major homelessness starts to appear when income is lower than this level.

The inequality or distribution of income of citizens has an impact on the fabric of the city as well. When inequality increases, more "extremes" start to appear on both sides of the income spectrum, leading to more homelessness and at the same time a rise in average housing price and upper social class prevalence in neighborhoods with houses that are for sale.

If the current trend of increase in housing prices keeps on going up, this will lead to such high housing prices that more people will become homeless. Furthermore, because the prices increase, people are forced to live in smaller houses or in neighborhoods that do not meet their standards, which causes the overall utility to drop (and thus, implying a drop of well-being of citizens). By countering the increase in housing prices, these adverse effects can be slowed down or negated.

Overall the model shows that without intervention, exogenous factors will lead to a greater divide between people in the long term. Bigger social group and income inequality will be observed and will lead to bigger inequalities between neighborhoods. This might lead to a divide between citizens, losing a shared identity as a citizen of The Hague.

To prevent the adverse effects of the exogenous factors that are described in this section, a range of possible policy interventions have been created and simulated in the model. The usefulness and effects of each policy lever are discussed in Section 4.4.

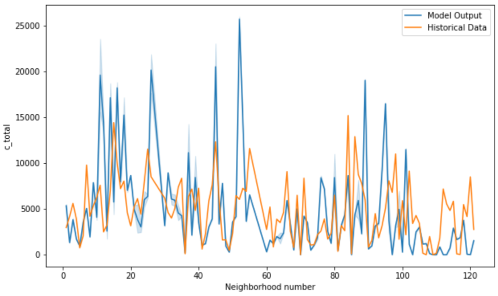
4.3 Validity of the Model

To get a sense of the performance of the model and to better understand how accurate the outcomes of the model are in comparison to the real-world, validation is needed. Because of the complex nature of the observed interactions and behavior of agents in the model, validation of the decision-making logic is not possible (within the scope and time constraints of this research). However, a second option that can give an indication to the correlation between model outputs and real-world observations, is a historical validation.

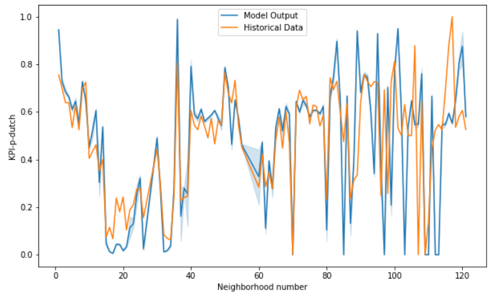
4.3.1 Historical Validation

Data can be very useful to describe patterns in a system, such as the dynamics of a system over time. Such patterns can be analyzed and compared to similar patterns observed as outputs of the model, which in term gives a sense of the performance of the model (Windrum, Fagiolo, & Moneta, 2007; Pullum & Cui, 2012). By using data which has been gathered every year in the past, running the model from a prior moment and comparing the end of the simulation with the current real-world data of 2020, a good indication of the validity of the model can be gained.

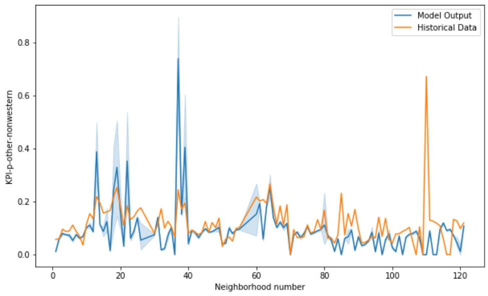
To do so, all necessary data sources to run the model have to implement data from a prior moment and run the model up to the most recent available data source in time. Currently, the most recent data available is from the beginning of 2020. Unfortunately, the model relies on an extensive dataset for modeling households which has only been available starting in 2017. This is therefore chosen as the starting point of the model for historical verification. With a run time of 3.5 years, the outcomes should give a good indication of the performance of the model in comparison to the real-world observations.



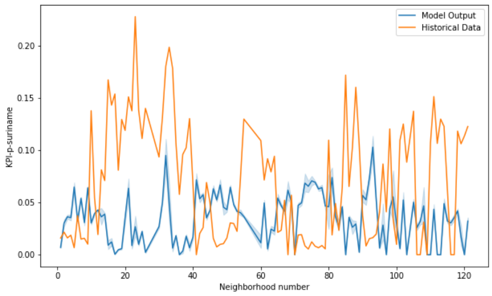
(a) Total citizens per neighborhood comparison.



(b) Dutch ethnicity percentage prevalence comparison.



(c) Other Non-Western ethnicity percentage prevalence comparison.



(d) Surinamese ethnicity percentage prevalence comparison.

Figure 4.12: Comparison of citizen prevalence for each neighborhood between model simulation outcomes and real-world observations.

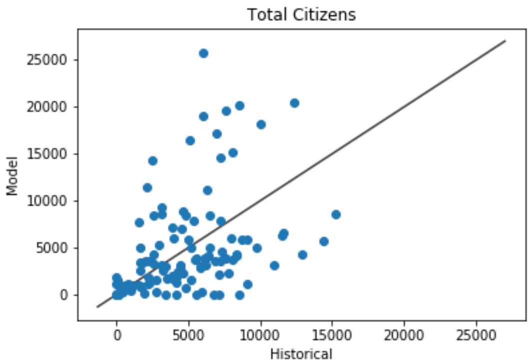
Figure 4.12 highlights the comparison of four of the most significant KPIs between model output and real-world observations. Figure 4.12a shows the total amount of residents per neighborhood at the beginning of 2020. Here, we see the model assumes more people will be able to live in certain neighborhoods than is the case in real-life. This can be attributed to the fact that the model looks at the household level, and assumes a uniform distribution of household size. In practice, there is a correlation between household size and the price of housing options. Since this correlation is not present in the model, some neighborhoods have much more residents than the real-world observations show.

Next, Figure 4.12b shows the percentage prevalence of Dutch citizens, here we can see the model has a decent "fit" with the real-world observations. The exceptions are neighborhoods with 0 citizens, which, depending on data show no residents where there should be, or citizens moving into neighborhoods where no homes are built. Apart from this anomaly, the model seems to show similar "trends" in prevalence of the Dutch ethnicity when compared to the real-world observations, ergo, the model seems to be able to simulate the decision-making of Dutch citizens in the modeled 3.5 years pretty accurately.

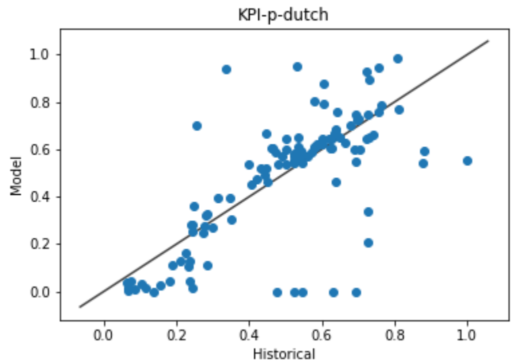
The third graph shows the performance of simulating the decision-making behavior of non-western ethnicities (Figure 4.12c). Apart from two major miscalculations, the model seems to simulate the decision-making accurately. The two exceptions show the model was unable to identify 1 neighborhood as interesting for migrants, whilst overestimating the likeability of another.

Lastly, Figure 4.12d shows the ethnicity that performed the worst when comparing real-world observations with the decision-making simulated in the model. Here, the plot shows the model was completely unable to explain the decision-making of citizens of Surinamese background. This might mean the assumptions for the decision-making of citizens do not apply to Surinamese citizens, or factors are missing which are critical to describe the moving behavior of Surinamese citizens.

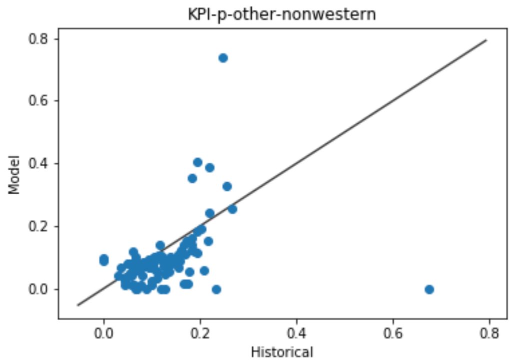
By comparing the observed values from the real-world and the model simulation, and plotting the outcomes in relation to one another, a better view of the performance is gained. Figure 4.13 shows the relation by plotting the real-world value for each neighborhood on the x-axis, whilst plotting the outcomes from the simulation model on the y-axis. The black line in the middle of the graph shows the line where $y = x$, in other words, where the observed real-world values matches the outcomes of the model for a neighborhood. By defining the R^2 value of each of the value pairs, the correlation between the pairs can be calculated. Using R^2 assumes a linear regression model, which looks similar to the plotted data.



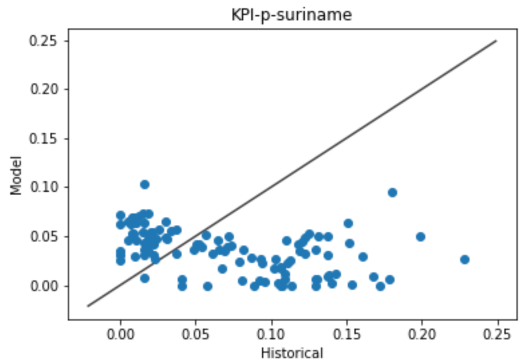
(a) Total citizens. R^2 value: 0.11



(b) Dutch percentage prevalence.R^2 value: 0.48



(c) Other Non-Western percentage prevalence. R^2 value: -0.17



(d) Surinamese percentage prevalence. R^2 value: -9.98

Figure 4.13: Comparison of citizen prevalence for each neighborhood between model simulation outcomes and real-world observations.

As Figure 4.13a shows, the model both under- and overshoots the real-world data. Since the regression line is in the middle of the data points, the model can be assumed accurate but not precise. This can indicate that different ethnicities have different decision-making approaches, since the model does not skew into one direction but both directions. The overall trend seems to be accurate, which means the model can (to some degree) describe the decision-making of citizens.

Figure 4.13b shows the performance for the Dutch ethnicity which seems to performing pretty well. As the results show, the most significant model "misses" are neighborhoods which have no citizens in the model, but do have values in the real-world.

Furthermore, Figure 4.13c shows the performance of the model for simulating the prevalence of non-western ethnicities. Here, the model seems pretty accurate, with the exception of 2 neighborhoods. Just as Figure 4.12c showed, the missing of these neighborhoods significantly decreases the precision of the model.

Lastly, Figure 4.13d shows that the model is unable to fully capture the real-world decision-making behavior of Surinamese citizens. Not only is the model imprecise, the model outcomes are skewed in comparison to the real-world observations, showing a low accuracy as well.

4.3.2 Validity Findings

As the historical validation has showed, the model is able to capture some of the decision-making behavior of citizens in such a way that it represents similar outcomes as observed in the real-world. This means that the model is able to explain (some) of the decision-making of citizens and therefore simulate how the composition of citizens in the city changes over time. However, there are certain situations where the model is unable to capture the complexity of the real-world, and cannot replicate the outcomes as observed in the city. For instance, the model is unable to explain the decision-making behavior and moving pattern of Surinamese citizens. Furthermore, because the model simulates agents as households, which all have the same size, the model is unable to correctly represent the population sizes of neighborhoods.

Because we now know that in its core, the model is able to explain (some) of the decision-making logic and moving patterns of citizens, it can be used to simulate the impact of migration on cities. Furthermore, the model can be used to test the effectiveness of certain policy interventions. It should be noted however, that not all complexity can be captured in the model, and more research is needed to explain the big dissonance between observed prevalence of Surinamese citizens in the real-world versus the outcomes of the model.

4.4 Opportunities and Adverse Effects for The Hague

To cope with the uncertainty of the ever-changing fabric of the city of The Hague, and to prevent adverse effects of these changes, the municipality is able to influence the outcomes of the changes of the city by applying policies. The policies (as defined in Section 3.3), can influence the outcome of the model and thus give an indication of how certain alternatives can influence the shaping of the city. Furthermore, this chapter shows the relation between certain interactions appearing and the corresponding policy levers being used. To conclude, the sensitivity of each lever is discussed to get a better grasp on the effect of value sizes of each policy.

First, to get a better understanding of the expected future when no policy intervention takes place, Figure 4.14 show the average citizen utility for each of the neighborhoods. Here we can see, there are a few "problem areas" which show severe utility problems whilst there are also a few areas that outperform over the rest of the city. The average utility seems to correlate to the several other factors, as Figure 4.15a shows. For instance, neighborhoods where more private sector rents houses and houses for sale are still available, tend to have a low utility. This could be a good predictor for poor well-being, as neighborhoods with high utility tend to have full occupancy in the private rent sector and houses sell quickly. Similarly, more people living in an area correlates with higher utility, which can be explained by the effects of social cohesion increasing well-being. Lastly, there is a negative correlation between prevalence of citizens in the middle social group and utility, indicating that in the baseline scenario there are issues for this social group in finding suitable homes.

On a positive note, when looking at the changes in average utility over time, there is a clear upward trend visible in the baseline scenario. Figure 4.15b show the increasing utility, which might be explained by citizens finding more suitable homes for their needs over time. The slight drop in utility at the end of the simulation runs can be related to the massive influx of migrants, which occurs at year 9 (time step 36).

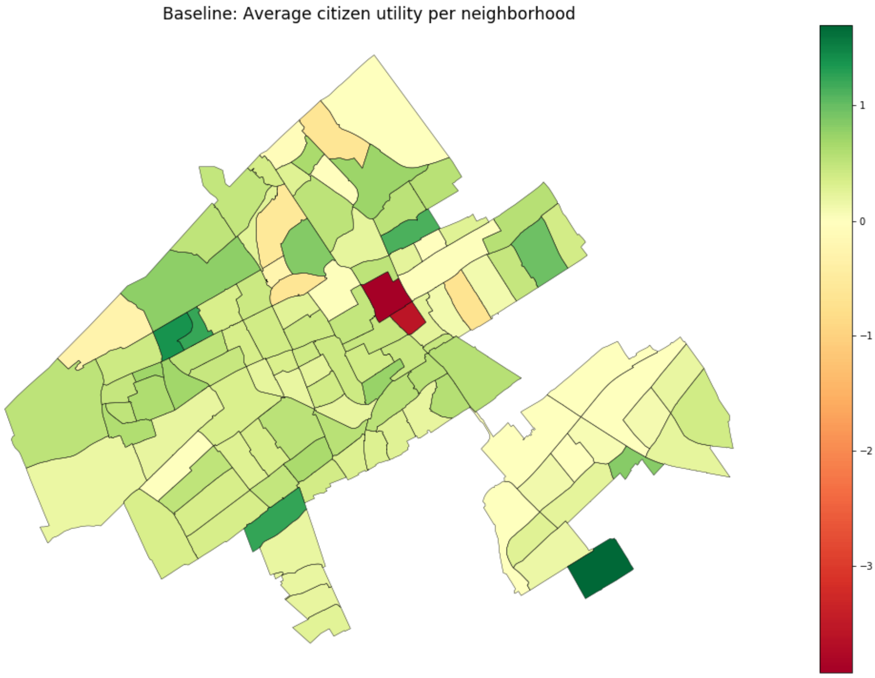
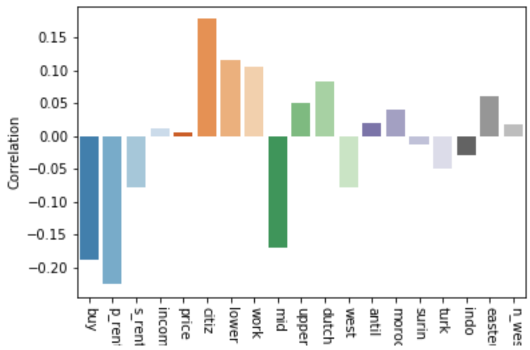
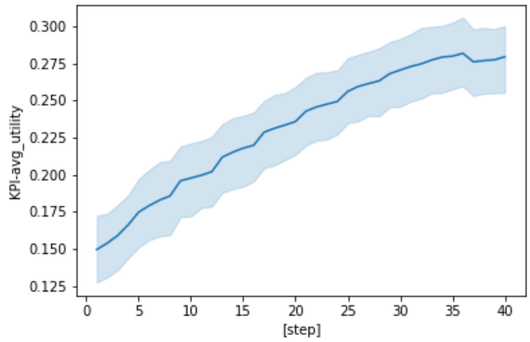


Figure 4.14: Average performance of citizen utility at the end of simulations for each neighborhood when no policy intervention takes place.



(a) Correlation between citizen utility and other KPI values.



(b) Development of citizen utility over time.

4.4.1 Policy Levers

The actions the municipality (with or without help of the government) can take to influence the system are defined in policy levers. This levers describe certain actions that can either be "turned on" or "turned off" (Section 3.3 explains all levers). This means, that the policies within the scope of the research are either present and actively performed, or not present. It could be interesting to see the use of certain policies over certain time frames instead of assuming the policy to run for the whole duration of the model. However, this is unfortunately out of scope for this research. To provide a quick overview of the possible policy interventions, Table 4.2 sums each of the policy levers and a summary of its effects. All policy interventions happen once a year, at the end of the year. It should be noted that the feasibility and budget-wise possibility of performing policies is out of scope of the research, and the implications of policy interventions are described to provide insight to possible outcomes.

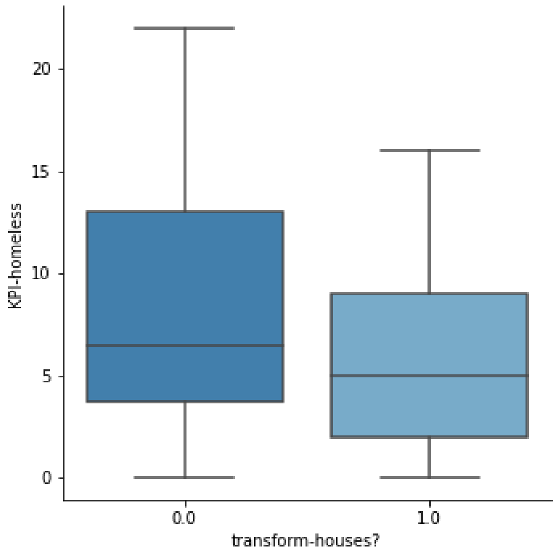
Lever	Name	Description
L1	transform-houses?	By the municipality buying properties, the 5 neighborhoods with the least amount of rental properties transform 50% of houses for sale into rental properties (both private and social sector).
L2	improve-health?	By building more healthcare facilities, the 5 neighborhoods with the worst healthcare get a 10% increase in the availability of healthcare facilities.
L3	increase-social-housing?	At the beginning of the model simulation, 10% of all rental properties in the private sector are bought and transformed to social rent housing options.
L4	build-more-houses?	By funding projects and granting construction permissions, the 5 neighborhoods with the least amount of housing availability will increase the availability in all sectors by 2,5%.
L5	mixed-use-zoning?	By allowing mixed-use of commercial or industrial zones, new housing is created in non-residential areas. This equates to 400 new housing options per year in each mixed-use zone.
L6	improve-safety?	By increasing the budget of the police and social workers, the amount of crimes in the worst 5 neighborhoods drops with 0-15%. Because of uncertainty, this drop is not a fixed value but randomly drawn.

Table 4.2: Table showing all policy levers in the model simulation. Values of parameters in the simulation are fixed but tested for different values in the sensitivity analysis.

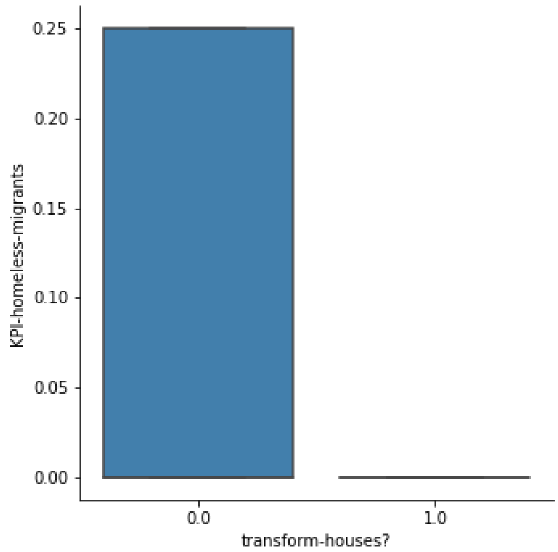
The model is simulated with all policy levers to see how the model behaves, and interactions change, given the policy levers. Each policy has been tested in isolation, to make sure no interference of other factors are measured when observing the output. Furthermore, using sensitivity analysis, the impact of each single policy lever and its values are tested. More details on the process of testing the sensitivity are explained in Chapter L. The impact of each policy lever is discussed below.

4.4.2 Transforming Houses to Rental Options

The first policy intervention option for the municipality of The Hague that is simulated is transforming current housing properties by buying houses that are for sale, and turning them into (smaller) homes which can be rented. In practice this means that for every house that is removed from the housing market, one housing option for both private sector rent and social rent are added. This is possible because most houses for sale can be divided into smaller apartments, which can then be put up for rent. The default policy lever assumes that this process takes place once a year for the 5 neighborhoods with the least available rental properties, and transforms 50% of all houses that are currently for sale. The robustness of these values is further analyzed in the sensitivity analysis in Chapter L.

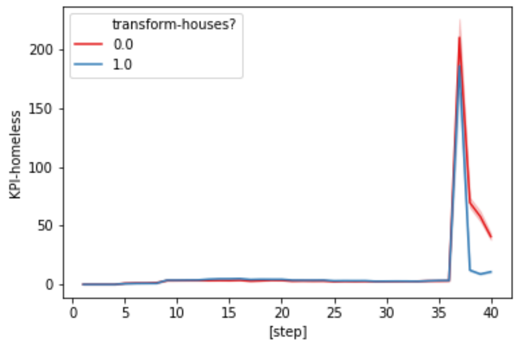


(a) Distribution of homelessness at the end of simulations when comparing transforming houses into rental property versus doing nothing.

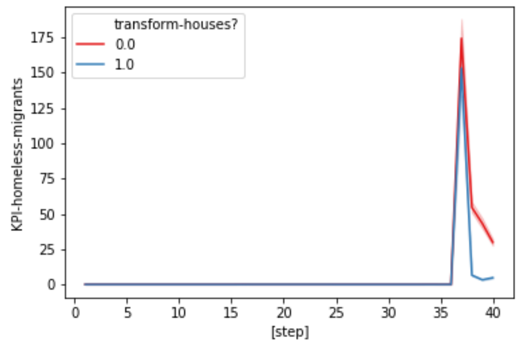


(b) Distribution of homeless migrants at the end of simulations when comparing transforming houses into rental property versus doing nothing.

As Figure 4.16a shows, the most significant change that can be observed from putting this policy into place is the decrease in homelessness, and especially among homeless migrants (Figure 4.16b). This can be explained by the fact that most migrants can only afford to live in rental housing options, and maybe can even only afford social rental options. By increasing the amount of rental properties, more migrants are able to find a home and thus homelessness drops.



(a) Average homelessness over time when comparing transforming houses into rental property versus doing nothing.



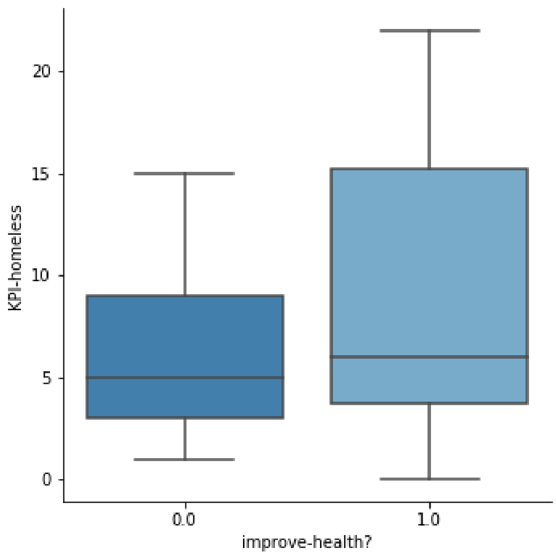
(b) Average homelessness of migrants over time when comparing transforming houses into rental property versus doing nothing.

More interestingly, the dynamics of implementing this policy lever can be observed when looking at a timeline of the implications. As Figure 4.17a shows, housing problems only start to occur at the end of the simulation. This is also the case for migrants, as shown in Figure 4.17b. Here, we can see that the spike in homeless migrants is solved almost completely within 4 time steps (ergo, 1 year). However, an ever-increasing influx of migrants might indicate housing problems in further futures.

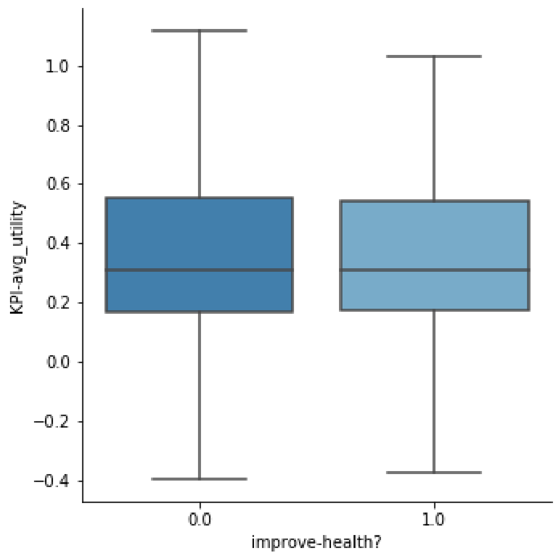
4.4.3 Improving Healthcare Facilities

The second policy intervention possibility that is tested in the model is improving health care. Since the average healthcare in neighborhoods contributes to the well-being of citizens, this policy should address the average utility of citizens. By increasing the amount of doctor's offices or other healthcare facilities, this policy models an increase in healthcare availability of 10% for the 5 worst-performing neighborhoods.

The outcomes that can be observed from testing this policy in the model are quite unexpected. Looking at the distribution of utility, Figure 4.18b shows little to no increase in utility is observed from improving healthcare. However, a surprise adverse effect of increasing healthcare is the increase in homelessness (Figure 4.18a). A probable cause of this adverse effect might be the fact that the healthcare improvements focus on neighborhoods with the worst current healthcare, which coincidentally are the neighborhoods with the lowest housing prices. By improving healthcare, these neighborhoods become more attractive for a different type of citizen, which pushes out the poorest citizens by increasing housing prices over time. This ultimately leads to more homelessness. It should be noted that many of the benefits of increasing healthcare in poorer neighborhoods are not observed in this model, such as an increase in lifespan or a decrease in sick leave.



(a) Distribution of homelessness at the end of simulations when comparing improving healthcare versus doing nothing.

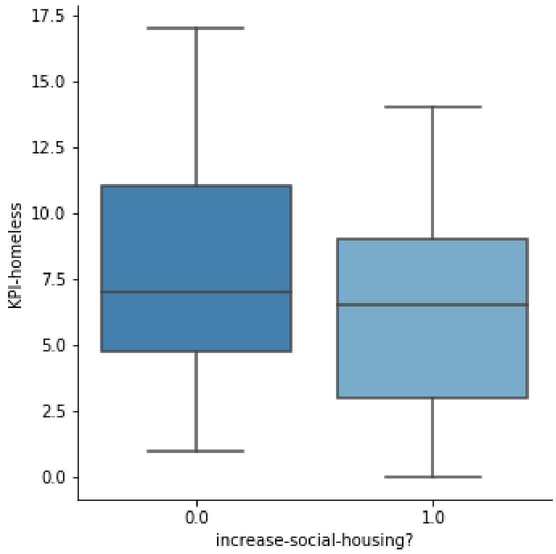


(b) Distribution of homelessness at the end of simulations when comparing improving healthcare versus doing nothing.

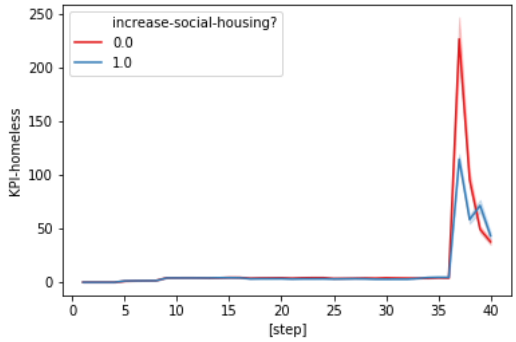
4.4.4 Increase Social Housing Availability

By increasing the amount of available social housing options in neighborhoods, more (poorer) citizens are able to find a suitable, affordable homes. This policy lever implies an intervention where the municipality forces 10% of all rental properties in the private sector to transform into social rental housing. This will happen only once, at the start of the model.

Looking at the results, the impact of this policy is only visible when looking at the amount of homeless people. Figure 4.19a less people tend to be homeless at the end of the simulation when compared to not applying this policy lever. The dynamics become more apperent when looking at Figure 4.19b, which shows the dampening effect for the expected homelessness spike around year 9 (tick 36). This can be prevented by a bigger supply of social housing options to provide for the poorest in society, which correlates to the influx of migrants. It should be noted that the impact of this policy might be smaller then other policies as this is the only policy which intervenes only once, at the start of the model.



(a) Distribution of homelessness at the end of simulations when comparing increasing social rent housing versus doing nothing.

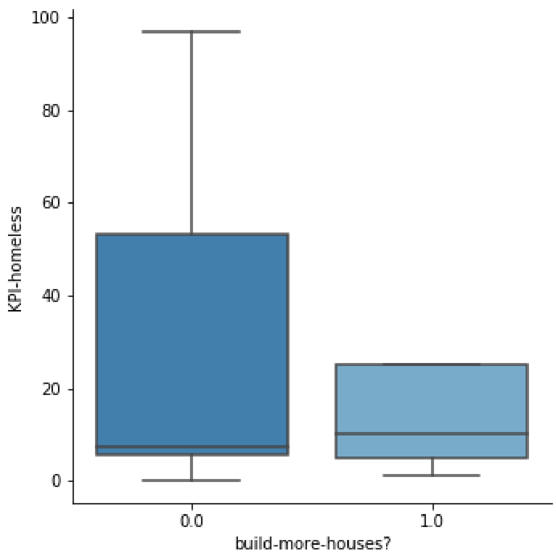


(b) Average amount of homeless people over time when comparing increasing the amount of social rent housing versus doing nothing.

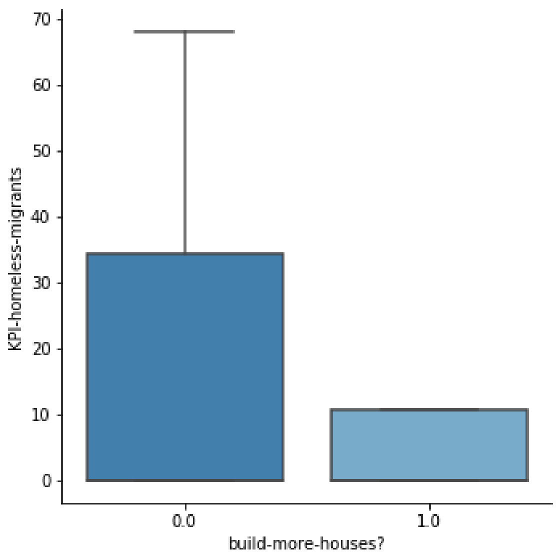
4.4.5 Build More Houses

By providing more building permits or subsidising new building projects, the municipality is able to promote the construction of new housing options in the city. In the model, this policy lever assumes an increase of 2,5% in housing availability (for all three sectors) in the 5 neighborhoods with the least amount of free housing options. This happens once every year.

Looking at the results, the most significant change in outcomes can be observed when looking at homelessness (Figure 4.20a) and in particular, homelessness amongst migrants (Figure 4.20b). As the boxplot shows, building extra houses provides more possibilities for citizens to be able to find a suitable home to live in, therefore decreasing homelessness. Interestingly, because migrant homelessness is impacted the most, it can be assumed that most new housing options are constructed in the neighborhoods with low housing prices.

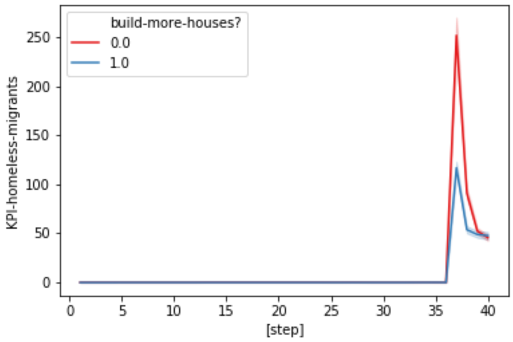


(a) Distribution of homelessness at the end of simulations when comparing building new housing options versus doing nothing.

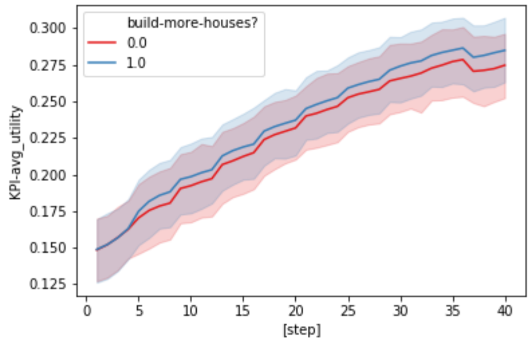


(b) Distribution of homeless migrants at the end of simulations when comparing building new housing options versus doing nothing.

When looking at the dynamics of implementing this policy over time, the effect of building more housing options becomes more clear (Figure 4.21a). Because more housing options are available already in year 9 (tick 36), the spike in arriving migrants can mostly be handled by the surplus in housing created by the construction of new homes. Another interesting development can be observed when looking at the dynamics of utility (Figure 4.21b). After year 1 (tick 4), when the first construction of new homes occurs, the average utility development is significantly higher than in the baseline scenario, which means more homes also increases well-being of citizens.



(a) Average migrant homelessness over time when comparing building new housing options versus doing nothing.



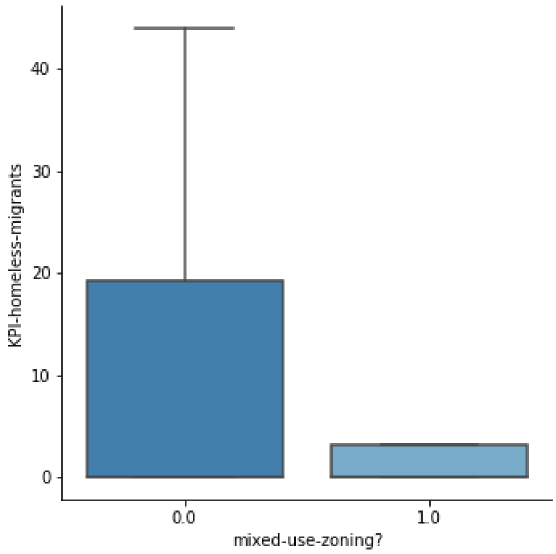
(b) Development of citizen utility over time when comparing building new housing options versus doing nothing.

4.4.6 Allow Mixed-Use Zones

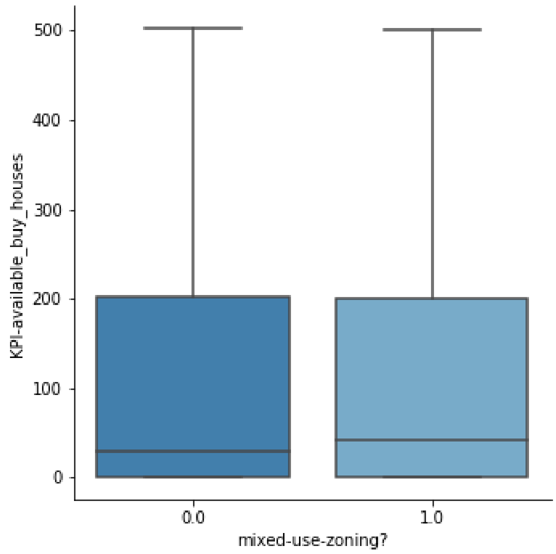
To cope with the shortage of housing options, another possible intervention for the municipality is allow mixed-used zoning. This means areas that are currently designated as industrial or commercial zone will transform some of the building or land into suitable housing options. This policy lever models the designation of 4 zones which will allow for the building of residential homes, which will create 400 new houses each year.

Looking at the outcome space for this policy lever, there are many changes to be observed. A selection of observable differences between baseline scenario and policy implementation are shown to give an indication of the impact of this policy lever. First, Figure 4.22a shows the amount of homeless migrants at the end of simulation runs. Here we can see that increasing the availability of housing options using mixed-use zoning decreases homelessness. Figure 4.22b shows that the availability of houses for sale is significantly bigger at the end of the run compared to the baseline scenario. This difference can also be observed for private sector and social rent housing availability.

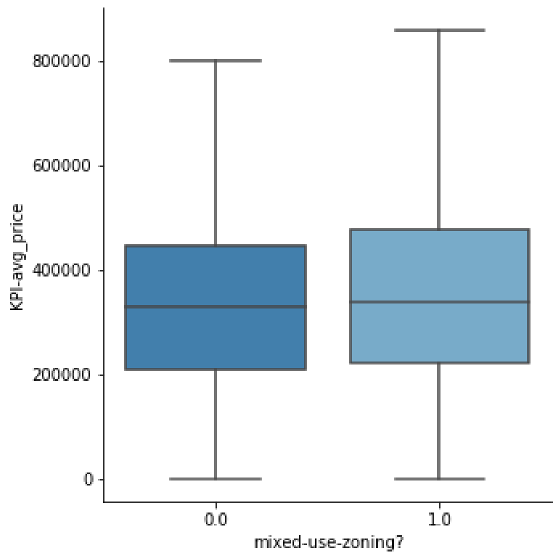
Furthermore, Figure 4.22c shows that increasing housing availability in "new" neighborhoods increases the average housing prices. The way to model interpolates housing price for new areas with mix-used zoning, is by taking the average housing prices of the 4 nearest neighborhoods. An explanation for the increase in average housing price as a result of mixed-use zoning can be that many of the new neighborhoods are situated in the more expensive parts of the city. Lastly, Figure 4.22d shows an increase in the prevalence of upper social group in neighborhoods, which is also the case for middle social group (but not shown in this plot). An explanation could be that the upper and middle social groups now have a more diverse selection of neighborhoods which fill the requirements of their needs, thus increasing the overall prevalence of these social groups.



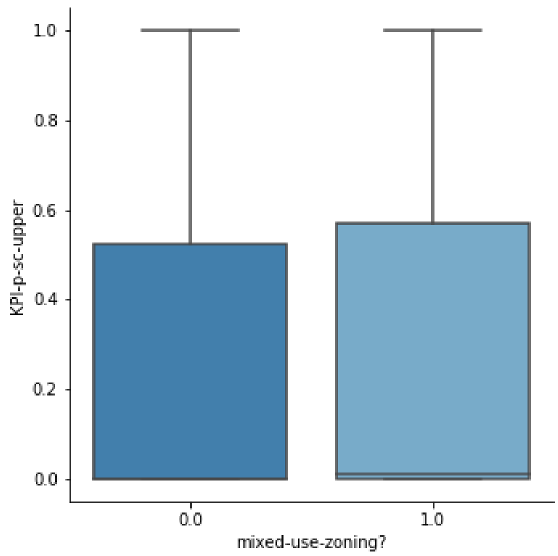
(a) Migrant homelessness when comparing housing in mixed-use zones versus doing nothing.



(b) Availability of houses for sale when comparing housing in mixed-use zones versus doing nothing.



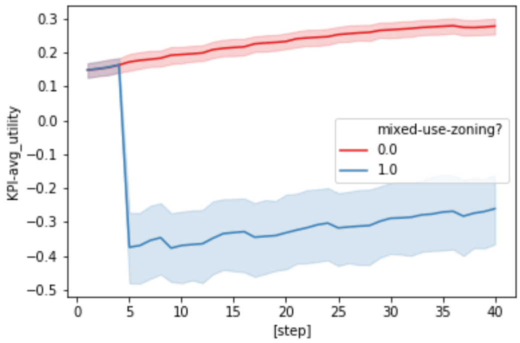
(c) Average housing value across neighborhoods when comparing housing in mixed-use zones versus doing nothing.



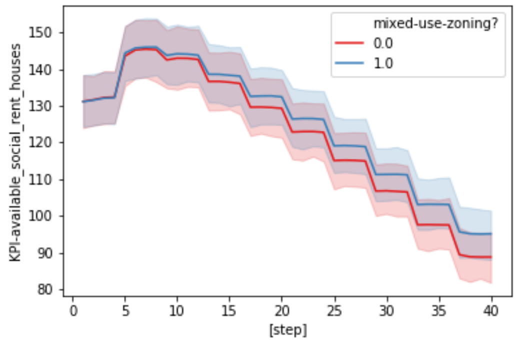
(d) Prevalence of upper social group in neighborhoods when comparing housing in mixed-use zones versus doing nothing.

When looking at the dynamics of this policy lever over time, two novel and interesting interactions can be observed. Figure 4.23a shows the development of utility over time. As the graph shows, the construction of mixed-use housing options has a significant negative impact on the average utility of citizens. This might be caused by the fact that these new neighborhoods are located in commercial or industrial zones, which lack amenities and services. Furthermore, the perceived safety of these areas might be lower, which also lower utility.

The second plot (Figure 4.23b) shows the trend for the availability of social rent housing options over time. As the graph shows, the extra housing options help combat the dwindling supply of housing options (which are caused by the influx of migrants, as can be observed at the cyclical decrease in availability). A similar trend can be seen when looking at the housing options for private sector rent homes. Surprisingly, the availability of houses for sale only seems to grow over time, confirming that migrants are barely ever able to afford to buy a house and thus only impact the rental housing availability.



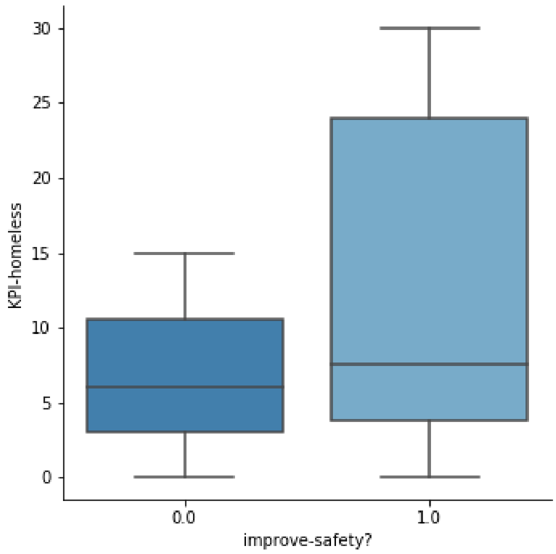
(a) Citizen utility over time when comparing allowing mixed-use housing versus doing nothing.



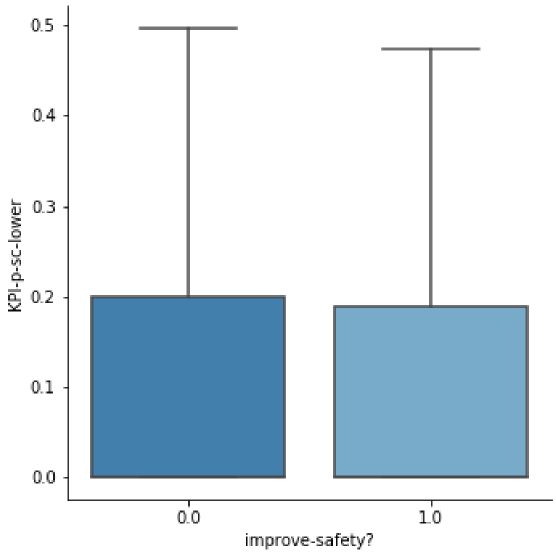
(b) Availability of social housing over time when comparing allowing mixed-use housing versus doing nothing.

4.4.7 Improve Safety, Reduce Crime

To improve citizen's well-being, the municipality can invest in a safer environment. By increasing the amount of police, social worker and neighborhood care funding, the municipality can try to decrease crime and increase the notion of safety. In the model, this policy lever assumes a decrease in crimes occurring in the 5 neighborhoods with the most amount of crime. This decrease ranges between 0-15% for each year. Because the effectiveness of increasing policy presence or social workings in decreasing crime is not absolute, a random distribution between 0-15% models the decrease.

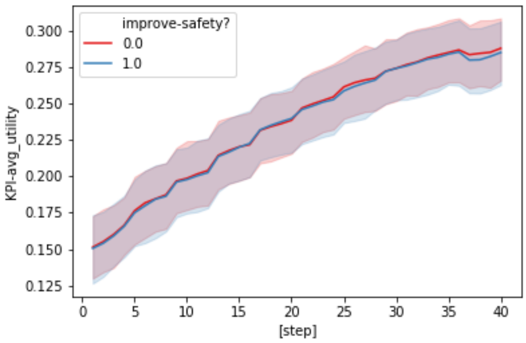


(a) Distribution of homelessness at the end of simulations when comparing improving safety versus doing nothing.

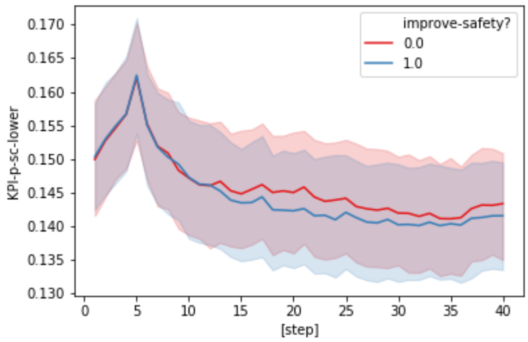


(b) Prevalence of lower social group at the end of simulations when comparing improving safety versus doing nothing.

The goal of improving safety in neighborhoods is to increase well-being and utility of citizens, and make unsafe neighborhoods more attractive. However, the data shows little significant change in utility. Figure 4.24a shows an increase in homelessness at the end of simulations where safety is increased. This is a surprising and unexpected outcome, and might be caused by an increase in competitiveness in the poorest neighborhoods. In other words, when the poorest neighborhoods become safer, they also become more attractive for citizens that have more money. This pushes out the poorest of citizens, therefore becoming homeless. Similarly, Figure 4.24b shows a decline in the presence of lower social group citizens at the end of the simulation when safety is increased.



(a) Citizen utility over time when comparing increasing safety versus doing nothing.



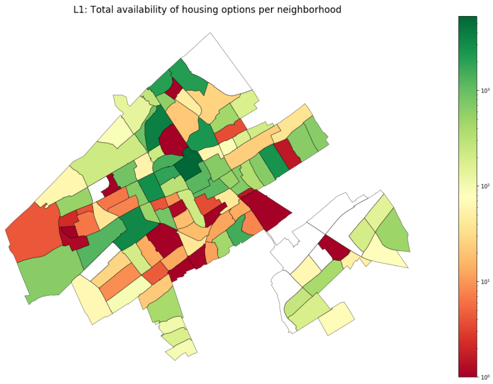
(b) Prevalence of lower social group citizens over time when comparing increasing safety versus doing nothing.

When observing the dynamics from a temporal perspective, the implementation of safety increasing policies show to mostly have adverse effects. Figure 4.25a shows a minor decline in utility over time, which might be explained by the pushing out of poor citizens and being replaced by those who require more needs to be fulfilled than the neighborhood has to offer. Figure 4.25b shows the decrease of the prevalence of citizens from the lower social group, which can be attributed to the pushing out effect. The increase in prevalence at the end of the simulation can be explained by the influx of migrants, which peaks at year 9 (tick 36) of the simulation.

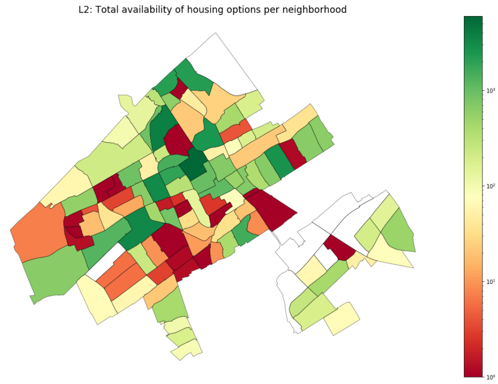
In short, increasing the safety in neighborhoods tends to have more negative adverse effects than positive outcomes. Currently, the selection of neighborhoods that get an increase in safety is based on reducing crime in the most crime-ridden neighborhoods first, however, this tactic seems to be inefficient.

4.4.8 Spatial Observations

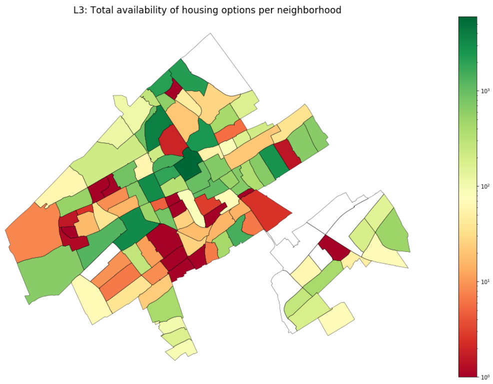
When looking at the effectiveness of the simulated policy levers, it can be fruitful to observe changes from a spatial perspective. By plotting the overall performance of policy levers on a map of all neighborhoods, a sense of performance can quickly be obtained. The most relevant KPIs to use for plotting effectiveness are availability of housing and utility, as most policies strive to better these statistics. Figure 4.26 plots the summed availability of housing options of all three sectors, and projects maps for each of the policy levers. Here we can see that the distinction between different policies is not significantly great. In other words, problematic availability of housing is tied to specific neighborhoods, and a universal policy approach seems ineffective in tackling the problem.



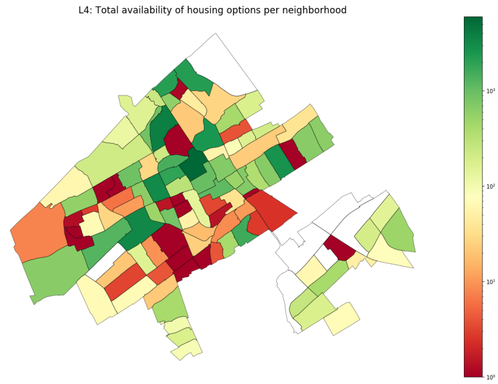
(a) Policy Lever 1: Transforming Houses.



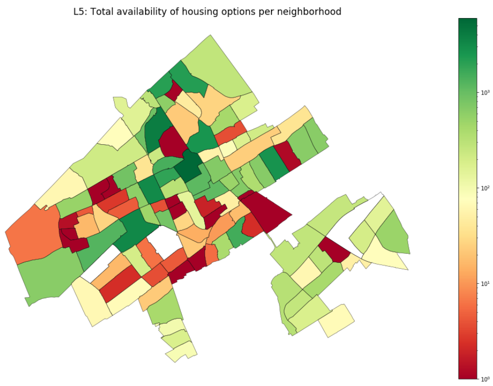
(b) Policy Lever 2: Increase Healthcare.



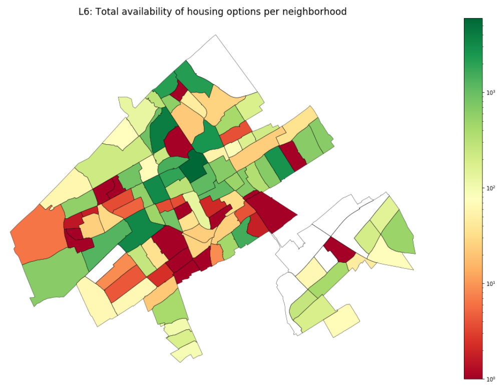
(c) Policy Lever 3: Increase Social Housing.



(d) Policy Lever 4: Build More Housing.

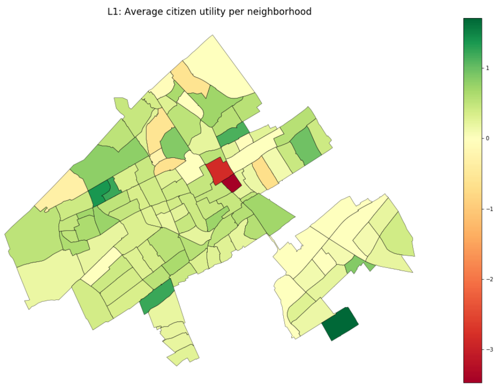


(e) Policy Lever 5: Allow Mixed-Use Zoning.

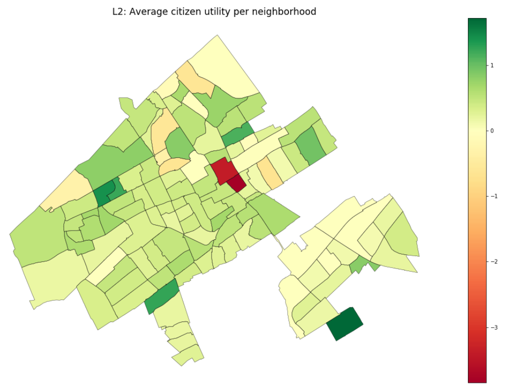


(f) Policy Lever 6: Increase Safety.

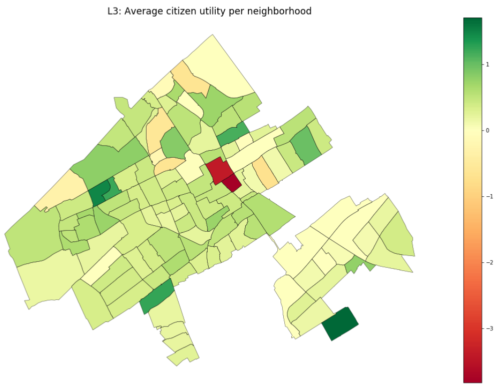
Figure 4.26: Average housing availability for each of the neighborhoods in The Hague at the end of the simulation, showing the effectiveness of policies in prevented a housing shortage. White indicates no citizens.



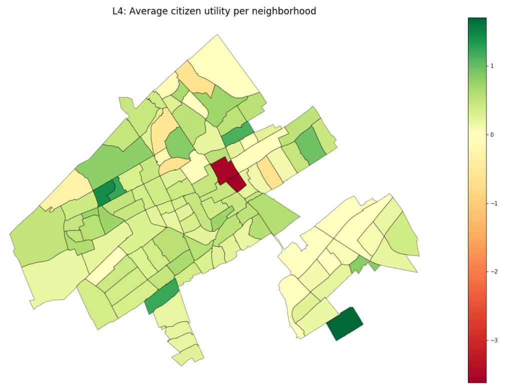
(a) Policy Lever 1: Transforming Houses.



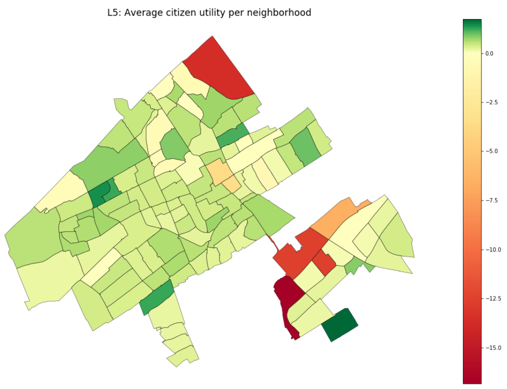
(b) Policy Lever 2: Increase Healthcare.



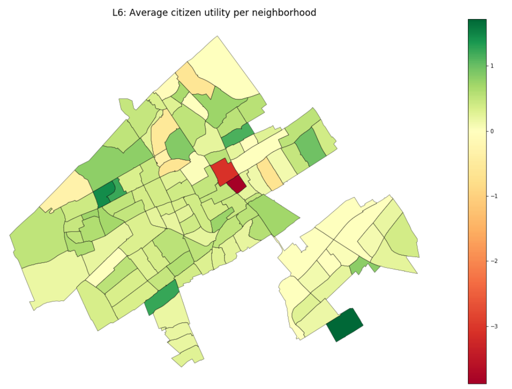
(c) Policy Lever 3: Increase Social Housing.



(d) Policy Lever 4: Build More Housing.



(e) Policy Lever 5: Allow Mixed-Use Zoning.



(f) Policy Lever 6: Increase Safety.

Figure 4.27: Average citizen utility for each of the neighborhoods in The Hague at the end of the simulation, showing the effectiveness of policies in increasing citizen well-being. Neighborhoods with no citizens are show with value 0.

The second set of maps shows the overall utility, or well-being, of citizens at the end of simulation runs given their respective policy input. Figure 4.27 shows the problematic and great performing neighborhoods quickly, as they stand out in color. The pattern between all policy levers seems very similar, with the exception of mixed-use zoning. Allowing citizens to live in areas which do not have amenities (yet), bring severe utility issues, as the new red zones indicate.

4.4.9 Temporal Observations

By looking at the development of outcomes over time, more understanding of the effects of certain policies can be gathered. For example, a rise in a certain KPI can be more easily observed to its cause when looking at the moment in time where the change start happening.

To get a better sense of the impact of policy interventions in relation to the average well-being of citizens, the average utility is plotted over time. Figure 4.28 show the relation between the development of utility and policy levers. The major observation that immediately stands out is the drop in utility for policy lever L5. In this intervention, new housing options are available in mixed-use zones. The drop in utility corresponds with the moment the first houses are available and can be explained by the lack of amenities and services in this region. To make sure the well-being of citizens that move to mixed-use zoning location can be guaranteed, mixed-use zoning should only be available in those areas which are close to amenities and services to prevent this decline in utility.

A second observation from Figure 4.28 is that policy lever L1 seems to slightly outperform the other alternatives when comparing utility. In this scenario, more houses that are for sale are transformed into rental properties. A plausible reasoning for improving utility is the fact that many citizens cannot afford to buy a house, and an increase in rental options thus increases overall utility since there are more suitable options for more citizens around the city.

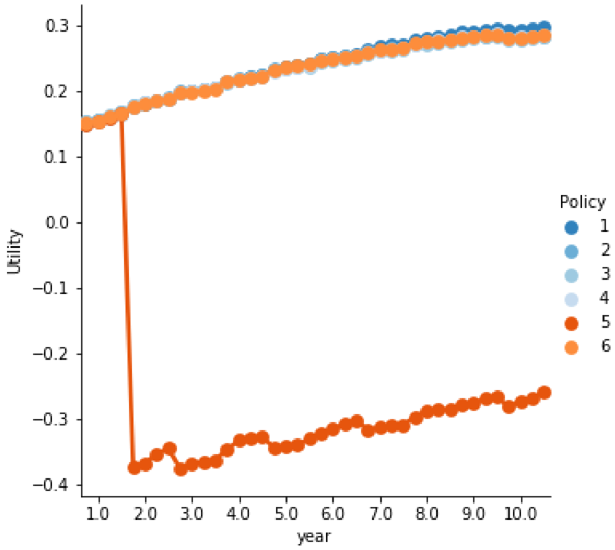


Figure 4.28: Time-wise development of citizen utility given different policy interventions. The numbers correspond with the policy lever numbers.

4.4.10 Changing the Fabric with Policies

As this section has shown, 6 possible policy interventions have been tested for their viability in improving the city fabric of The Hague. All of the policy levers show significant changes in outcomes of the simulation model, but not all changes are intended or positive. A summary of observations is listed in Table 4.3, which denotes all policies and their most interesting outcomes.

Lever	Description	Outcome
L1	By the municipality buying properties, the 5 neighborhoods with the least amount of rental properties transform 50% of houses for sale into rental properties (both private and social sector).	An increase of rental housing options leads to an overall increase in citizen utility and a decrease in homelessness, especially amongst migrants. Overall, this is one of the best performing policies.
L2	By building more healthcare facilities, the 5 neighborhoods with the worst healthcare get a 10% increase in the availability of healthcare facilities.	Improving healthcare in the worst-performing neighborhoods leads to the pushing out of poor citizens, increasing homelessness. The positive change in utility is at the same time negligible. Many benefits of this policy are not measured in the model, such as an increase in health of citizens and life expectancy.
L3	At the beginning of the model simulation, 10% of all rental properties in the private sector are bought and transformed to social rent housing options.	Transforming private sector rent homes to social rent housing has a significant impact on the decrease of homelessness. Especially at the end of the simulation, when the influx of migrants peaks, the extra social rent housing options help with providing homes for the poorest in the city.
L4	By funding projects and granting construction permissions, the 5 neighborhoods with the least amount of housing availability will increase the availability in all sectors by 2,5%.	Building more houses in neighborhoods with the lowest availability has two positive effects, first homelessness decreases as more people are able to find a suitable home in their price range. Secondly, the neighborhoods with the biggest shortage in housing supply have the best utility averages, thus increasing housing options in these locations has a net benefit effect on average utility.
L5	By allowing mixed-use of commercial or industrial zones, new housing is created in non-residential areas. This equates to 400 new housing options per year in each mixed-use zone.	Although this policy lever can supply the most new housing options of all policies, and thus decreasing homelessness the most, this policy also causes major drops in average citizen utility. The new housing locations lack amenities and services, which has a negative effect on citizen utility.
L6	By increasing the budget of the police and social workers, the amount of crimes in the worst 5 neighborhoods drops with 0-15%. Because of uncertainty, this drop is not a fixed value but randomly drawn.	Similarly to increasing healthcare facilities in the poorest performing neighborhood, decreasing crime has little to no benefit to the outcome of the simulation. Because crime is reduced in the poorest neighborhoods, poor citizens get pushed out of their homes which causes a rise in housing price, a rise in homelessness and a drop in average utility. The positive effects of reducing crime do not reflect in the model outcome data, such as an improvement in perception of safety.

Table 4.3: Table showing all policy levers in the model simulation, and their respective expected outcomes.

As the table shows, some policies have outcomes which are along the lines of expectation, however some policies show adverse effects which were not anticipated. Furthermore, the improving of neighborhood problems such as health or safety do not reflect a positive change to the fabric in the model. This can be attributed to adverse effects but also a lack of measurement of some of the factors that these policies try to improve such as life expectancy, perception of safety and other less quantifiable factors.

The most important factor that can give an indication of the performance of a policy lever, is the average citizen utility of residents of neighborhoods. As Figure 4.27 shows, some neighborhoods are reluctant to change given the current policy levers. These "star" neighborhoods and "problem" neighborhoods are performing very well or very bad under all conditions, and require extra attention. Furthermore, a more local policy intervention approach might be more suitable in these areas to improve the situation.

The second set of factors that gives a good indication of the performance of policy levers, as well as indicating which neighborhoods are favoured by citizens, is the availability of housing options (as shown in Figure 4.26). The neighborhoods with the lowest availability tend to have the highest utility, or otherwise have the lowest average costs (which thus harbours the ability for the poorest citizens to be able to afford living there). The areas with the least amount of housing option availability thus require extra attention from the municipality, as these neighborhoods are most often a predictor of factors that citizens are looking for when considering to move someplace else.

Lastly, it should be noted that changing the values or frequency of intervention in the sensitivity analysis might show if policies can perform better when different values are chosen for the parameters. This can also show the performance of each of the policies with regards to the assigned parameter values. This would be an interesting thing to do in the sensitivity analysis (Chapter L). However, because of time limitations, this has not yet been done.

Chapter 5

Conclusion

The research that was done was two-fold. First, literature on city dynamics and changes to the city fabric and the behavior of citizens within a city was consulted as a means to formulate a framework. This framework, the conceptual model of the research, can be used to define the interactions of citizens within the city to describe the changes to the city fabric. The current use of this model has only been tested to describe the relation between the influx of migrants to urban areas and the resulting changes to the city fabric. However, the conceptual model can be applied to different areas of research as well with regards to describing changes to the city fabric. For instance, with little adjustments, the conceptual model can be used to describe traffic and congestion in a city, (in)equality of income, education or amenities and services within the city or other phenomena.

The second part of the research focused on applying the conceptual model to describe the relation between the influx of migrants and the resulting changes to the fabric of the city. To do so, a case study was used using real-world data. Using this data, an Agent-Based Model was made that simulates the decision-making behavior of migrants (and citizens) of a city and is capable of observing the resulting changes to the city. Using real-world data can then also be used to validate the model by comparing the output of the simulation to real-world data using a historical validation analysis.

As a case-study, data from the city of The Hague, The Netherlands was used. The municipality of The Hague wants to get a better understanding of the dynamics of changes in the city fabric as a result of the influx of migrants. By better understanding the dynamics, and knowing what interventions the municipality can take to either combat a certain change or promote another, the municipality will be able to make better decisions on which steps to take.

By first exploring the parameter space of exogenous factors, a better understanding of what changes the fabric has been established. An increase in the expected influx of migrants will lead to mass homelessness in the long-term. Furthermore, the average spendable income of migrants is an important determinant for homelessness; at least 17.000 Euros as an average income is needed for migrants to be able to find a home.

The income distribution of citizens also has an impact on the system, and the changing of the fabric of the city. As inequality in income distribution increases, more adverse effects start to show such as homelessness, drop in average well-being of citizens, vacant properties which are for sale, an increase in average housing price and a bigger gap between those able to afford utilities, amenities and services; and those who do not.

Lastly, the observed rise in average housing pricing will ultimately lead to many citizens no longer being able to afford their current living standard, and thus a shift in citizen composition start to occur. The upper social class will more or less remain the same, but citizens in the middle and working social class will no longer be able to afford their needs, which results in a drop in well-being and even an increase in homelessness.

To combat the negative future outlooks or promote the positive outcomes, the municipality is able to use policies to steer the changes to the fabric of the city. The model describes 6 possible policy interventions which are tested in simulation. In short, all policies which increase the availability of housing (and especially increasing social rent housing) show a benefit to the city. By increasing the supply of housing options, an average increase in well-being can be observed as well as a decrease in homelessness. However, increasing the availability of housing options by allowing mixed-use zoning (residential in commercial or industrial areas) has some negative side-effects. Since these new neighborhoods within commercial or industrial areas lack proper amenities and services, it leads to a decrease in well-being of citizens.

The municipality is also able to improve the situation in the worst performing neighborhoods with respect to healthcare and safety. However, the outcomes of the simulation show a negative impact when enacting policies in the improvement of either healthcare or safety. Improving the situation in the worst-performing neighborhoods leads to an increase in attractiveness of these neighborhoods, which leads to an increase in housing prices, thus leading to the pushing out of poor citizens by people that can afford more needs. This in term leads to the poor citizens being forced to live in worse conditions, or even becoming homeless, which in the long-term decreases the overall well-being of citizens. It should be noted that the positive effects of improving both healthcare and safety are immeasurable in the model, as many factors that get improved by such policies are not in the scope of the model and thus not included in the outcomes.

It should be noted that for all policy levers, the viability, feasibility and operationability have not been studied in detail. This means that the costs for both implementing as well as operating a certain policy lever have not been addressed, nor has the political or policy-wise possibility of implementation be studied. Considering the building of new homes, the renovation or transformation of current buildings and the buying of property by the municipality, this means further research is needed into the feasibility and viability of these options. For example, not every neighborhood will have enough (safe, feasible) space to expand the current capacity of housing options. Furthermore, laws and regulations might prohibit the building or transformation of homes. Lastly, the budget of the municipality is not unlimited and should be considered carefully to prioritize the neighborhoods that have the biggest need increases in housing availability.

Chapter 6

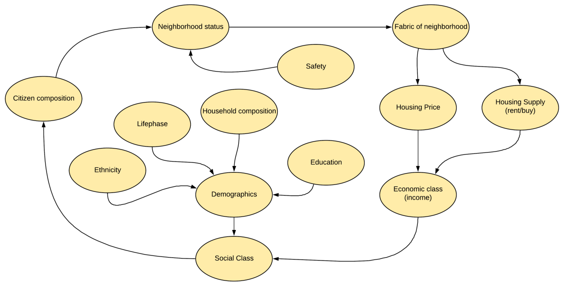
Discussion

An increase in the influx of migrants is expected to change the way West-European cities look in the future (United-Nations, 2019). By exploring the decision-making behavior of citizens and migrants with regards to moving patterns, a better understanding of changes to the city fabric can be obtained. The city fabric, the combination of citizen composition and properties that describe the state of neighborhoods in a city, is an ever-changing dynamic part of describing a city and is a good proxy for describing citizen well-being and the state of a neighborhood.

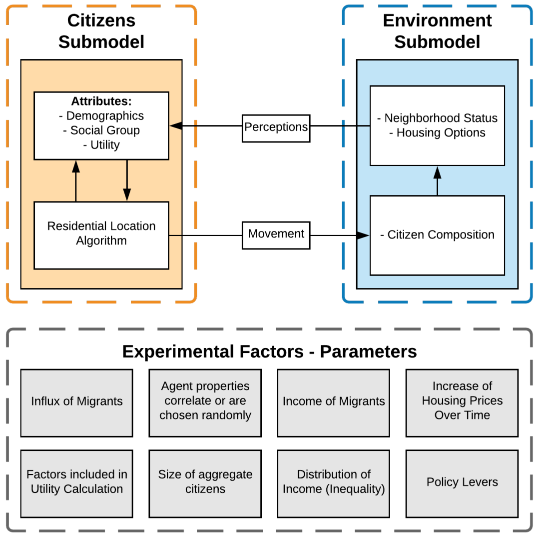
The goal of this research was to develop a model that can capture the complex nature of decision-making of citizens in the city and try to describe the changes to the city fabric when looking at the influx of migrants. This research has built a conceptual model as a framework for describing the system which defines the dynamics of changes to the fabric of the city given the influx of migrants. After the development of the conceptual model, data from the city of The Hague, The Netherlands was used to model the system. Using open data sources (of The Hague, 2020; CBS, 2019; Koot et al., 2019; Statline, 2020b), an Agent-Based Model (ABM) was constructed in the software package *Netlogo* (Wilensky, 1999). After the ABM was made, experimentation of simulation parameters was used to find answers to gaps in current knowledge regarding the development of the city fabric over time. After running the simulation in different ranges of parameter values, the results were analyzed using data analysis in Python (Downey, 2018).

First, a general basis of the current state of scientific research and literature was established by means of a literature study. Interesting models with regards to simulation of migration flows using ABMs were found and used as a basis for defining the knowledge gap of the research (Chen, 2012; Klabunde & Willekens, 2016; Bettencourt et al., 2010; Suleimenova et al., 2017; Tomasiello et al., 2020; Perez et al., 2019). Using the frameworks, models and literature, a conceptual model was made that encapsulates the dynamics of migrants entering the city and their decision-making with regards to finding a suitable location to live (Section 3.1.2).
To test the conceptual model and show the interactions observed in a real-world situation, an Agent-Based Model was constructed on the case study of the city of The Hague. Using data from readily available (open) data sources, the city was simulated in the ABM. Using literature on social behavior patterns, in combination with survey data from citizens of The Hague, the decision-making logic and preferences of citizens were established (Center, 2020; Argyle, 1994; Forrester, 1970; Maslow, 1943; Koot et al., 2019; van Amsterdam et al., 2015).
The resulting outcomes that were generated from simulating the model using scenarios and parameter settings, are analyzed to answer the research question and to get a better understanding of the interactions of citizens within the model. To better interpret the results of the analysis, the data is observed using correlation comparison, temporal observations and spatial observations (Chapter 4).

The first gap in knowledge was an absence of a clear definition of the components and interactions that form the basis of a model to describe the impact of migration on the city fabric. First, a system diagram was made based on literature to show the causal relations between factors influencing the city fabric (Figure 6.1a). Using this system as a basis, the inputs and outputs were then formalized into a conceptual model (Figure 6.1b).



(a) Causal diagram of the system describing the impact of migrants on the changes to the city fabric.



(b) Conceptual model of the city, describing the relation between citizens and their decision-making resulting in changes to the city.

To summarize the findings from the making of the conceptual model, there is a clear relation between the well-being of citizens in a neighborhood and the average housing price. When a neighborhood performs well, prices go up, leading to a change in citizen composition. This might lead to changes in external factors as well, such as a decrease in crime and degradation of the neighborhood. Furthermore, the conceptual model shows that citizens do not only look for the availability of amenities or a suitable environment, a more important factor in decision-making is the citizens that live in a neighborhood of choice. People tend to be more prone to live with other citizens of the same social standing or background, and thus clustering of similar people can be observed in the city (Perez et al., 2019; Forrester, 1970).

The second knowledge gap that has been researched is translating the definitions of the conceptual model into a case-study which operationalizes real-world data into a simulation model. Because of the heavy focus on the behavior of citizens, and the interaction between citizen and migrants, an Agent-Based Model was chosen as the approach to model the system (Klabunde & Willekens, 2016). For the ABM, *Netlogo* was used as a tool to model the system for its ease of use, interpretability, and prior experience of the Author (Wilensky & Payette, 1998).

An overview of the model in "action" is shown in Figure 3.1, but the whole model is open and free to use by cloning the *GitHub* repository of the Author [1]. Here, you can also find all the data used for the project, as well as all data analyses, results, graphs and documentation. By building the model, further gaps in knowledge can be answered using the model as a basis for experimentation and to observe agent behavior and interactions given parameter inputs.

[1] https://github.com/Jochem285/OpenDataDenHaag/tree/master/Netlogo

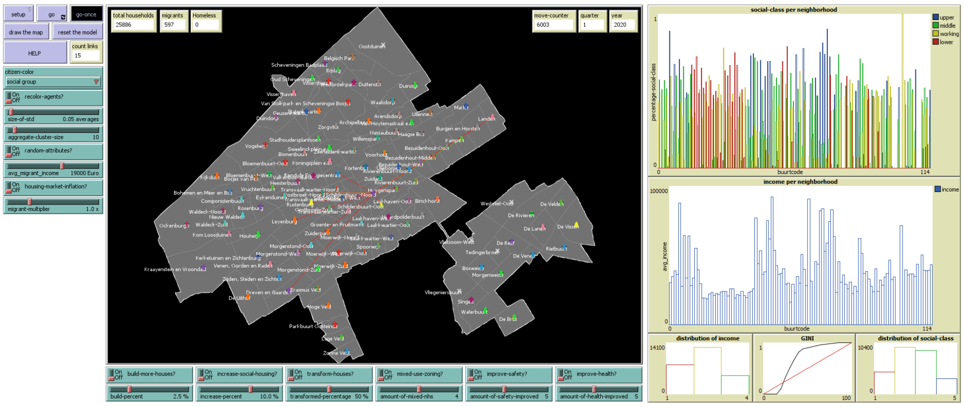
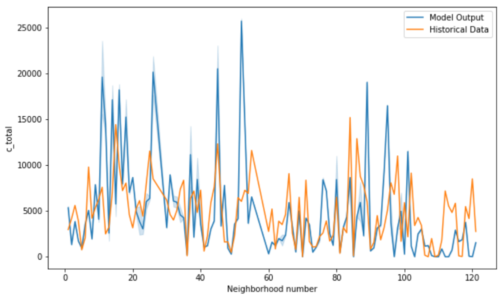
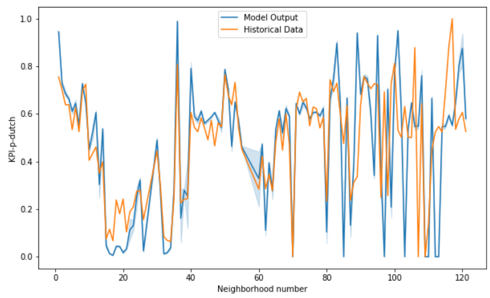


Figure 6.2: Overview of the model dashboard for the Agent-Based Model, made in Netlogo

After creating the Agent-Based Model, the validity of the model has been tested. Using historical data from The Hague, the output of the simulation is compared to data from the real-world. Doing so gives an indication of the performance of the model to capture the (complexity of) real-world interactions and phenomena. By running the model for three years of simulation time and comparing the output with real-world data of the same period (2017-2020), the overall behavior of the model can be observed. As Figure 6.3a shows, the overall output of the model in comparison to real-world data shows similar trends but leaves room for improvement. Looking at the most prevalent ethnicity of citizens, Dutch citizens, the model is better at simulating the behavior of these residents (Figure 6.3b). An explanation for this result can be that the social theories and surveys used to model the decision-making behavior of agents in the model leans heavily on inputs of Dutch citizens, and therefore is more robust in aggregation their behavior. The model is less accurate in predicting the behavior from for example citizens of the Surinamese ethnicity.



(a) Total citizens per neighborhood comparison.



(b) Dutch ethnicity percentage prevalence comparison.

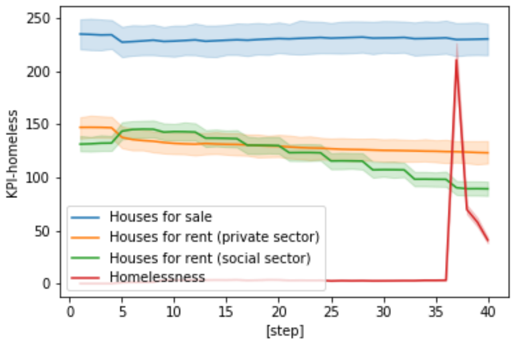
The ABM can be used to fill next gap in knowledge, which is finding the relationship between migration patterns and decision-making behavior of citizens in relation to changes to the city fabric. By observing the outcomes of the model simulations, and performing a sensitivity analysis on the inputs, a better understanding of the relation between the migration patterns and changes to fabric is found.

Looking at migrants and their behavior, an increase in the expected influx of migrants in the (near) future will lead to mass homelessness in the long-term. Furthermore, the average spendable income of migrants is an important determinant for homelessness; the lower the average income, the greater the amount of homeless migrants that can be observed in the model output. This relation is also found in literature, where the most heavily affected citizens are those with a migration background (Koot et al., 2019).

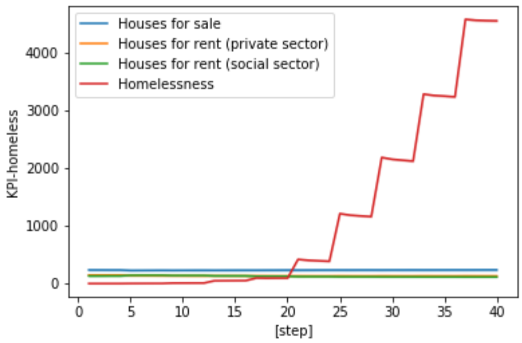
The income distribution of citizens also has an impact on the system, and the changing of the fabric of the city. As inequality in income distribution increases, more adverse effects start to show such as homelessness, drop in average well-being of citizens, vacant properties which are for sale, an increase in average housing price and a bigger gap between those able to afford utilities, amenities and services; and those who do not. Furthermore, a rise in average housing pricing will ultimately lead to many citizens no longer being able to afford their current living standard, an thus a shift in citizen composition start to occur in neighborhoods when housing prices increase. This shift in citizen composition as a reaction to changes in housing prices is observed in literature and coined with the term *gentrification* (Nara & Torrens, 2005; Betancur, 2014; Zuk et al., 2015).

Looking at the baseline scenario runs, the simulation runs where no parameters are changed from default and no policy interventions take place, the "business as usual" outcomes show that the expected influx of migrants will start being problematic around year 9 of the simulation, or in other words, around 2028. If no intervention takes place, migrants will start being unable to find affordable housing options, as can be seen in Figure 6.4a.

The expected continuing increase of housing prices as shown in recent statistical analysis proves to be problematic for the city as well (van de Statistiek, 2020b). In the model, simulating a steady increase of housing price of 7% per year causes bigger housing shortages in certain neighborhoods, decreases overall well-being and forces more homelessness. As Figure 6.4b shows, the impact of rising housing prices changes almost nothing to the availability (since citizens are not able to afford housing), whilst rapidly increasing homelessness.



(a) Timeline of availability of housing options compared to homelessness of citizens in the city of The Hague. The spike in homelessness at the end of the simulation is caused by a big influx of migrants, which are not able to find a suitable home in the first quarter after arriving in the city.



(b) Timeline of availability of housing options compared to homelessness of citizens in the city of The Hague given a steady increase of housing prices. After 5 years the rise in housing prices becomes problematic as citizens (and migrants) are no longer able to afford new homes, thus homelessness starts to rise.

Furthermore, there are a few neighborhoods which have significantly different citizen composition and fabric compared to other neighborhoods, as shown in Figure 6.5. These neighborhood might require intervention to steer the changing of the neighborhood fabric over time. Thus, the municipality can intervene the "business as usual" scenario, by means of policy interventions.

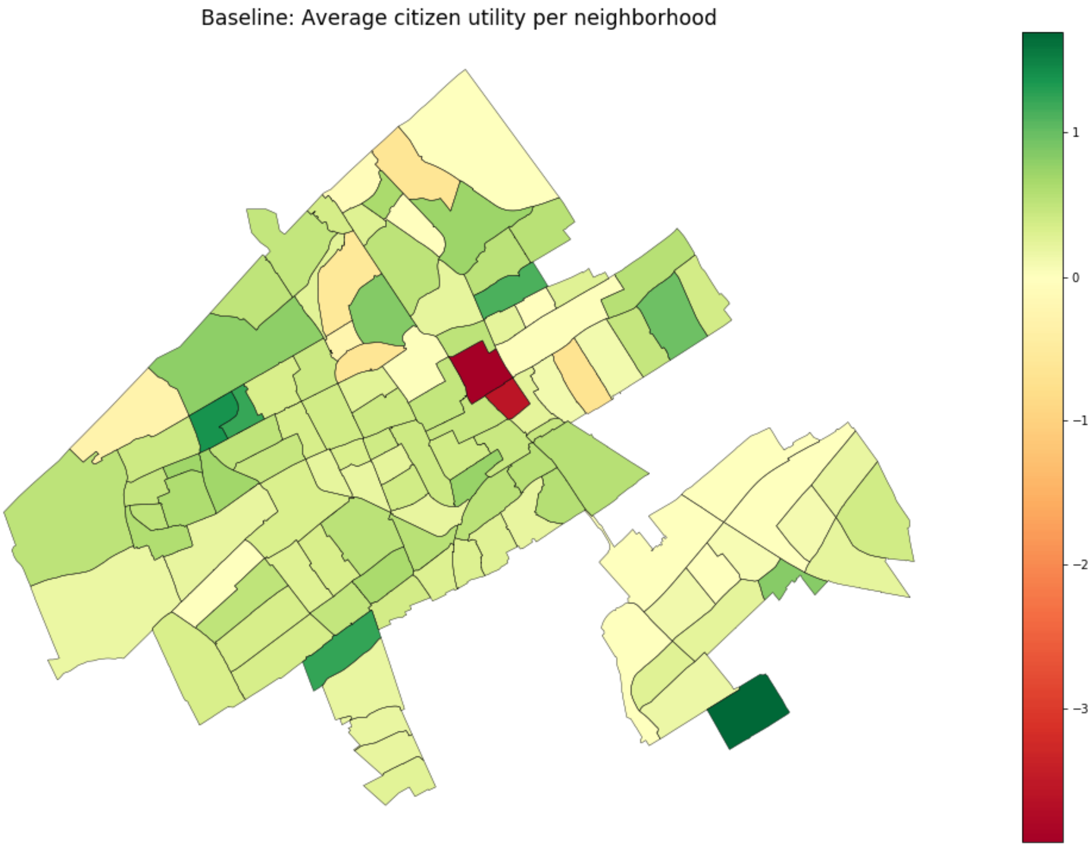


Figure 6.5: Map of The Hague showing the average "utility" of citizens living in that neighborhood at the end of the simulation of the baseline runs. The utility is a good indicator of citizen well-being, and a negative value indicates citizens have more negative factors influencing their well-being than positive factors.

To get an indication of what the municipality of The Hague can do to change the fabric of the city, or "steer" certain negative changes into positives, policy interventions have been modeled. Using policy levers in the model, the effectiveness of interventions by the municipality on the outcomes of the simulation model have been tested. In the model, 6 different types of policy interventions have been simulated to see how they impact the outcomes of the model. From literature, most of these interventions were proposed as possibilities to combat housing shortages and problems regarding healthcare and safety (van Amsterdam et al., 2015; Center, 2020).

By increasing the supply of housing options, an average increase in well-being can be observed as well as a decrease in homelessness. However, increasing the availability of housing options by allowing mixed-use zoning (residential housing in commercial or industrial areas) has some negative side-effects. Since these new neighborhoods within commercial or industrial areas lack proper amenities and services, it leads to a decrease in well-being of citizens (Center, 2020; van Amsterdam et al., 2015).

The municipality is also able to improve the situation in the worst performing neighborhoods with respect to healthcare and safety. However, the outcomes of the simulation show a negative impact when enacting policies in the improvement of either healthcare or safety. Improving the situation in the worst-performing neighborhoods leads to an increase in attractiveness of these neighborhoods, which leads to an increase in housing prices, thus leading to the pushing out of poor citizens by people that can afford more needs. This in term leads to the poor citizens being forced to live in worse conditions, or even becoming homeless, which in the long-term decreases the overall well-being of citizens. Although some of these observations can be seen in real-life (Zuk et al., 2015; Betancur, 2014), it could be assumed that the absence of positive impacts in the outcome space of these policy levers is caused by a shortcoming in the utility calculation logic, as improving the neighborhood status should lead to positive changes in utility. It should also be noted that the positive effects of improving both healthcare and safety are immeasurable in the model, as many factors that get improved by such policies are not in the scope of the model and thus not included in the outcomes.

To conclude the findings of the research and answer the main research question, the explored behavior, interactions and outcomes of the model can be interpreted to give a policy recommendation to the municipality of The Hague. By looking at the impact of migrants on the shaping of the city, and the changing of the fabric of the city, one can conclude that the impact of migrants is significant. Over time, the (increasing) influx of migrants in an environment that is already plagued by a shortage in housing and an increasing amount of citizens staying instead of leaving, housing shortages will become a major issue and focal point for the municipality (van Amsterdam et al., 2015; CBS, 2019; Koot et al., 2019). Not only will the shortage of housing lead to homelessness in the future, the scarcity of the housing market leads to price inflation and poor people being forced out of their current homes due to no longer being able to afford the costs.

As the model shows, building new housing options will be necessary in the future to make sure the city can cope with the influx. Furthermore, looking at the outcomes of the model, a universal approach for the whole city seems to be ineffective. The municipality will need to focus their attention on looking at the requirements of citizens for each individual neighborhood, and look at the current shortcomings of the neighborhoods and try to accommodate for those needs, while being aware of the effects such as *gentrification* pushing out vulnerable citizen groups. To prevent inequality arising within and between neighborhoods, it is important to strengthen the social cohesion between citizens. As surveys show, this cohesion is attributed as the most important positive factor for describing well-being of living (Center, 2020). As the model shows, citizens tend to base most of their decision-making when evaluating locations to move to based on their perception of the citizen composition in that area. Finding a healthy balance of citizen composition with a shared identity for neighborhoods should be the focus of policy interventions.

From a scientific point of view, this research, the conceptual model and the Agent-Based Model that were made in the process of the research can be used as tools to better understand the dynamics and impact of migration in cities. Looking at the case-study for The Hague, and comparing with literature from cities in the United States, some similarities can be observed in the decision-making process. However, many differences with regards to citizen composition, housing supply, pricing and market and the fabric of cities are observable, and thus further research using this framework might prove relevant as scientific contributions. Furthermore, some of the assumptions made in the decision-making logic of migrants choosing a location to move to might be validated in the real-world. For example, it is unknown how much influence the limited rationale of actors plays a role in decision-making (Gigerenzer & Selten, 2002).

6.1 Further research

As the discussion concluded, many questions arise from the results of this research. Moreover, further study might confirm the validity of the model as well as compare the case study of The Hague with other case studies. A list of interesting fields of study for future research are listed in this section.

As Chapter G indicates, a lot of mechanics and inner workings of the model are based on assumptions. Although data is readily available for the case of The Hague, the decision-making of citizens is a black box which requires assumptions on its functioning. Furthermore, because of a lack of data, some properties such as income distribution within neighborhoods, income distribution of migrants, educational background and other properties of migrants and the effectiveness of enacting policies have all been based on assumptions. Further research into the validity of these assumptions might prove fruitful.

The model that was made for the research project is an agent-based model, which focuses on the interactions between migrants and the city, and looking at the resulting changes as an outcome of interactions. However, another approach of modeling the system is looking from a System Dynamics' perspective. Here, the system would be interpreted as a set of stocks and flows. Looking at the influx of migrants could then be interpreted as a series of "goods" moving in and out of the city, and the problem would be considered much more operational. In other words, instead of looking at the behavior and the interaction *trends*, SD would look at the precise stocks and flows and better show which neighborhoods are failing in supplying enough housing availability to keep up with the demand (Andris, Halverson, & Hardisty, 2011).

During the writing of this Thesis, the pandemic of COVID-19 occurred, having severe impact on daily life and the economy. As of writing, the long-term impact on the economy, migration, housing and changes to the fabric of the city are unknown. If a recession occurs, it can result in citizens losing the ability to be able to pay their upkeep or rent for their homes. Which might even lead to evictions (Desmond, 2016). Where do people go after getting evicted or forced to move? Will the pricing of houses change as a result of the pandemic? What will be the result of that in terms of the composition of the fabric of the city? More research is could be interesting in this topic.

6.2 Limitations

In the model, agents represent households and not a single individual citizen. Because of limited time for the project, and with the risk of making a too complex model, the design decision was made to not include the complexity of simulating agents on an individual level. Because of this, agent properties do not cover all aspects of the real-world complexity. For example, literature shows one of the reasons for housing shortages is the ever-decreasing average household size, caused by young people living on their own (and staying single for longer), which results in a decrease in houses-to-citizens ratio (van Amsterdam et al., 2015). Another example of a phenomenon that is observed in literature but out of scope of the current model, is the increase of people of middle age staying in the city center instead of moving to the suburban part of the city (van Amsterdam et al., 2015). Modeling agents on the individual level might help understanding these factors better.

Because of the way the decision-making of citizens was modeled, social cohesion proves to be the biggest factor for defining well-being of citizens and also a major factor for the decision-making of citizens when looking for new housing options. Firstly, the mechanism for defining social cohesion in the model is very limited, it only looks for similarities of citizen properties, and only identifies a single property at a time. An extension to this would be identifying multiple properties, or including properties such as education or job type. Secondly, studies show social cohesion is dependent on an individual's social network, and thus, defining social cohesion as a factor for all citizens in a neighborhood is a simplification of real-world processes and neglects factors such as a shared identity. Lastly, because the influence of the social cohesion on the well-being of citizens is so big, sensitivity analysis is required to make sure the factorization of values is correct.

The second biggest factor that describes utility, and the biggest negative factor, is crime and perception of safety. This is based on the survey work of (Center, 2020), where multiple different factors regarding safety are addressed. These factors are bundled into 1 factor describing safety, based on the average importance of all the separate factors. Furthermore, the calculation of safety within the model is only based on reported crimes in a neighborhood, whilst there is no data on the perception of safety. There, to make sure the utility calculation of safety is rigorous, these problems have to be addressed before the utility calculation can be interpreted as robust.

In the model, the average price of housing is adjusted at the end of a simulation year to reflect changes in the market. This is based on the average income of the citizens that live there. However, this is a simplification of the "invisible hand" of the housing market, and many studies show that the inflation of market prices does not always correlate with demand or income of citizens. Furthermore, the model assumes all agents live in the house that they buy/rent, and that a lease or mortgage can be cancelled at any moment. This is of course a simplification of the real-world and could influence the model behavior.

The role of uncertainty in the housing pricing market is very significant, as current price keep on increasing (especially in the area of The Hague) and will become unsustainable in the near future. More research into the price increase trend and its resulting impact on the housing market is therefore needed to better explain how this will impact the whole fabric of the city (van de Statistiek, 2020a, 2020b; van Amsterdam et al., 2015; CBS, 2019; Koot et al., 2019).

The model has no mechanism for coping with homeless citizens, or the introduction of new migrants to the city. This means that in the model, if someone enters the city they immediately start looking for housing options and if they are unable to find affordable, available housing, they will become homeless. There is no instance of governmental help or organizations that seek to bridge the housing gap for migrants or poor citizens. Because of this, the severity of certain policy levers with regards to homelessness might be overstated and dramatized.

Time constraints and limited resources have put constraints on the scope and granularity of the research. For example, the pandemic of COVID-19 has changed some of the work of the project, especially with regards of the working environment and the possibility of physical meetings and interviews.

Secondly, the focus of the project has changed over time causing the goal of research to slowly adjust to its current form. This process was iterative but required some additional work to adjust what had already been done, and fit it to the new approach. Thirdly, because of poor management of in-vitro testing of the model in its infancy state, some major issues within the model were only discovered when the model was already completed, haltering the progress of the research.

A problem that is out of scope of this research, but very prevalent in surveys, is the lack of parking space for cars. Almost all neighborhoods present this problem. Possible solutions could be increasing the available parking space, however, since the city is lacking the building area, a better approach would be to encourage citizens to replace their car for other forms of transport (Verkade & Brommelstroet, 2020). Furthermore, an increase in public transportation options might incentivize citizens to get rid of their car.

Public transportation is one of the amenities that was initially part of the research, but literature showed no relation between the decision-making of citizens looking for suitable homes and the availability of public transport. However, anecdotal evidence shows some citizens do take into account public transport when deciding on a new home. Furthermore, observing changes to the city fabric might include public transport as a means of tracking inequality in amenities and services.

In the model, amenities are present in data and only used to model the decision-making logic of citizens in the middle and upper social class. The influence of these amenities is almost negligible in the decision-making, and the amenities present in the model are static. However, when observing the changes to the city fabric, it might be interesting to model changes to amenities and services in neighborhoods as well, and describe a more significant role to the presence of such instances.

Writing this Thesis, the research project, making the model and process from A to Z has been an incredible learning experience. I have learned to be more practical in defining the in- and outputs of my research, be more concise in scoping down the research and research goal, be pragmatic about the steps I want to take and how the research questions come in, I learned to prevent becoming a "fallen-over bookcase" (Nikolic, 2020) by reading too much literature, I learned to make many mistakes in modeling (although I thought I was already pretty good at it), and learned the hard way that models only crash when you test them on clusters of computers whilst working fine on your own computer, I have learned to be more rigorous when looking at the input and output data of my model, but most importantly I have learned to know when to scope down a project and say enough is enough. Finishing this report is a feat that makes me realize how far I have come and I am thankful for being able to learn so many things is such a small time. I can't wait for what the future might bring.

References

Alfeld, L. E. (1995). Urban dynamics—the first fifty years. *System Dynamics Review*, *11*(3), 199–217.

Andersen, H. S. (2019). *Urban sores: On the interaction between segregation, urban decay and deprived neighbourhoods*. Routledge.

Andris, C., Halverson, S., & Hardisty, F. (2011). Predicting migration system dynamics with conditional and posterior probabilities. In *Proceedings 2011 ieee international conference on spatial data mining and geographical knowledge services* (pp. 192–197).

Argyle, M. (1994). *The psychology of social class*. Psychology Press.

Atkinson, R. (2002). *Does gentrification help or harm urban neighbourhoods?: An assessment of the evidence-base in the context of new urban agenda* (Vol. 5). ESRC Centre for Neighbourhood Research Glasgow.

Balcilar, M., & Nugent, J. (2019). The migration of fear: An analysis of migration choices of syrian refugees. *Quarterly Review of Economics and Finance*, *73*, 95-110. doi: 10.1016/j.qref.2018.09.007

Bao-xing, D. Q. (2003). Challenges faced by china in its rapid urbanization process in the near future [j]. *Urban Studies*, *6*.

Barbosa, H., Barthelemy, M., Ghoshal, G., James, C. R., Lenormand, M., Louail, T., ... Tomasini, M. (2018). Human mobility: Models and applications. *Physics Reports*, *734*, 1–74.

Batty, M. (2013). *The new science of cities*. MIT press.

Bauer, B., & Odell, J. (2005). Uml 2.0 and agents: how to build agent-based systems with the new uml standard. *Engineering applications of artificial intelligence*, *18*(2), 141–157.

Bazzan, A. L., & Klügl, F. (2014). A review on agent-based technology for traffic and transportation. *The Knowledge Engineering Review*, *29*(3), 375–403.

Benureau, F. C. Y., & Rougier, N. P. (2018, Jan). Re-run, repeat, reproduce, reuse, replicate: Transforming code into scientific contributions. *Frontiers in Neuroinformatics*, *11*. Retrieved from http://dx.doi.org/10.3389/fninf.2017.00069 doi: 10.3389/fninf.2017.00069

Betancur, J. J. (2014). Gentrification in latin america: Overview and critical analysis. *Urban Studies Research*, *2014*.

Bettencourt, L. M., Lobo, J., Helbing, D., Kühnert, C., & West, G. B. (2007). Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the national academy of sciences*, *104*(17), 7301–7306.

Bettencourt, L. M., Lobo, J., Strumsky, D., & West, G. B. (2010). Urban scaling and its deviations: Revealing the structure of wealth, innovation and crime across cities. *PloS one*, *5*(11).

Bettencourt, L. M., Lobo, J., & West, G. B. (2009). The self similarity of human social organization and dynamics in cities. In D. Lane, D. Pumain, S. E. van der Leeuw, & G. West (Eds.), *Complexity perspectives in innovation and social change* (pp. 221–236). Dordrecht: Springer Netherlands. Retrieved from https://doi.org/10.1007/978-1-4020-9663-1_8 doi: 10.1007/978-1-4020-9663-1_8

Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the national academy of sciences*, *99*(suppl 3), 7280–7287.

Bots, P. W. (2007). Design in socio-technical system development: three angles in a common framework. *Journal of Design Research*, *5*(3), 382–396.

Box, G. E. (1979). All models are wrong, but some are useful. *Robustness in Statistics*, *202*, 549.

Bretagnolle, A., Pumain, D., & Vacchiani-Marcuzzo, C. (2009). The organization of urban systems. In *Complexity perspectives in innovation and social change* (pp. 197–220). Springer.

Carling, J., & Collins, F. (2018). *Aspiration, desire and drivers of migration*. Taylor & Francis.

CBS. (2019). *Minder verhuizingen in 2018*. Centraal Bureau van de Statistiek. Retrieved from https://www.cbs.nl/nl-nl/nieuws/2019/09/minder-verhuizingen-in-2018

Center, U. D. (2020). *Prettig wonen in den haag*. Gemeente Den Haag. Retrieved from https://dashboards.cbs.nl/v2/prettig_wonen_in_den_haag/

Chan, S. (2001). Complex adaptive systems. In *Esd. 83 research seminar in engineering systems* (Vol. 31, pp. 1–19).

Charles, A., Guna, D., & Galal, H. (2017). *Preparing cities to manage migration.* Online.

Chen, L. (2012). Agent-based modeling in urban and architectural research: A brief literature review. *Frontiers of Architectural Research*, *1*(2), 166–177.

Cohen, B. (2006). Urbanization in developing countries: Current trends, future projections, and key challenges for sustainability. *Technology in society*, *28*(1-2), 63–80.

Courgeau, D. (1985). Interaction between spatial mobility, family and career life-cycle: A french survey. *European sociological review*, *1*(2), 139–162.

Dabbaghian, V., Jackson, P., Spicer, V., & Wuschke, K. (2010). A cellular automata model on residential migration in response to neighborhood social dynamics. *Mathematical and Computer Modelling*, *52*(9-10), 1752–1762.

Dam, K. H., Nikolic, I., & Lukszo, Z. (Eds.). (2013). *Agent-based modelling of socio-technical systems.* Springer Netherlands. Retrieved from http://link.springer.com/10.1007/978-94-007-4933-7 doi: 10.1007/978-94-007-4933-7

Dawid, H., & Neugart, M. (2011). Agent-based models for economic policy design. *Eastern Economic Journal*, *37*(1), 44–50.

De Jong, G. F., & Gardner, R. W. (2013). *Migration decision making: multidisciplinary approaches to microlevel studies in developed and developing countries.* Elsevier.

Den Haag, G. (2016). Agenda ruimte voor de stad. *Retrieved on June*, *9*, 2018.

Desmond, M. (2016). *Evicted: Poverty and profit in the american city.* Broadway books.

Doocy, S., Lyles, E., Delbiso, T. D., Robinson, C. W., & Team, I. S. (2015). Internal displacement and the syrian crisis: an analysis of trends from 2011–2014. *Conflict and health*, *9*(1), 33.

Downey, A. (2018). *Think complexity: complexity science and computational modeling.* " O'Reilly Media, Inc.".

Dun, O. V., & Gemenne, F. (2008). Defining 'environmental migration'. *Forced Migration Review.*

Duvell, F. (2019). The 'great migration' of summer 2015: analysing the assemblage of key drivers in turkey. *Journal of Ethnic and Migration Studies*, *45*(12), 2227-2240. doi: 10.1080/1369183X.2018.1468385

Economic, U. N., & Council, S. (2019, may). Special edition: progress towards the sustainable development goals. *United Nations.*

Edmonds, B., Le Page, C., Bithell, M., Chattoe-Brown, E., Grimm, V., Meyer, R., ... Squazzoni, F. (2019). Different modelling purposes. *Journal of Artificial Societies and Social Simulation*, *22*(3), 6. Retrieved from http://jasss.soc.surrey.ac.uk/22/3/6.html doi: 10.18564/jasss.3993

Edmonds, B., & Moss, S. (2005). From kiss to kids – an 'anti-simplistic' modelling approach. In P. Davidsson, B. Logan, & K. Takadama (Eds.), *Multi-agent and multi-agent-based simulation* (pp. 130–144). Berlin, Heidelberg: Springer Berlin Heidelberg.

Edwards, S. (2008). Computational tools in predicting and assessing forced migration. *Journal of Refugee Studies*, *21*(3), 347–359.

en Cultureel Planbureau SCP., S., Hart, J., Knol, F., Maas-de Waal, C., & Roes, T. (2002). *Zekere banden: sociale cohesie, leefbaarheid en veiligheid.* SCP.

Feitosa, F. F., Le, Q. B., & Vlek, P. L. (2011). Multi-agent simulator for urban segregation (masus): A tool to explore alternatives for promoting inclusive cities. *Computers, Environment and Urban Systems*, *35*(2), 104–115.

Filatova, T., Verburg, P. H., Parker, D. C., & Stannard, C. A. (2013). Spatial agent-based models for socio-ecological systems: challenges and prospects. *Environmental modelling & software*, *45*, 1–7.

Forrester, J. W. (1970). Urban dynamics. *IMR; Industrial Management Review (Pre-1986)*, *11*(3), 67.

Frith, M., Simon, M., Davies, T., Braithwaite, A., & Johnson, S. (2019). Spatial interaction and security: a review and case study of the syrian refugee crisis. *Interdisciplinary Science Reviews*, *44*(3-4), 328-341. doi: 10.1080/03080188.2019.1670439

Gane, N. (2005). Max weber as social theorist: 'class, status, party'. *European journal of social theory*, *8*(2), 211–226.

Gaube, V., & Remesch, A. (2013). Impact of urban planning on household's residential decisions: An agent-based simulation model for vienna. *Environmental Modelling & Software*, *45*, 92–103.

Ge, Y., Meng, R., Cao, Z., Qiu, X., & Huang, K. (2014). Virtual city: An individual-based digital environment for human mobility and interactive behavior. *Simulation*, *90*(8), 917–935.

Gigerenzer, G., & Selten, R. (2002). *Bounded rationality: The adaptive toolbox.* MIT press.

Groenewegen, P. P., Kerssens, J. J., Sixma, H. J., van der Eijk, I., & Boerma, W. G. (2005). What is important in evaluating health care quality? an international comparison of user views. *BMC health services research*, *5*(1), 16.

Haan, A., & de Heer, P. (2015). *Solving complex problems: Professional group decision-making support in highly complex situations.* Eleven International Publishing.

Harvey, D. (2008). The right to the city. *The City Reader*, *6*(1), 23–40.
Haug, D. S. (2008). Migration networks and migration decision-making. *Journal of Ethnic and Migration Studies*, *34*(4), 585-605. doi: 10.1080/13691830801961605
Head, B. W. (2008). Wicked problems in public policy. *Public policy*, *3*(2), 101.
Heppenstall, A. J., Crooks, A. T., See, L. M., & Batty, M. (2011). *Agent-based models of geographical systems.* Springer Science & Business Media.
Hoekstra, M. S. (2018). Governing difference in the city: Urban imaginaries and the policy practice of migrant incorporation. *Territory, Politics, Governance*, *6*(3), 362–380.
Hosseinali, F., Alesheikh, A. A., & Nourian, F. (2013). Agent-based modeling of urban land-use development, case study: Simulating future scenarios of qazvin city. *Cities*, *31*, 105–113.
Huijnk, W. (2016). Arbeidsmarktpositie en inkomen. *Integratie in zicht?*, 84.
Huynh, B. Q., & Basu, S. (2019). Forecasting internally displaced population migration patterns in syria and yemen. *Disaster medicine and public health preparedness*, 1–6.
Janssen, M. A., & Ostrom, E. (2006). Empirically based, agent-based models. *Ecology and society*, *11*(2).
Jantz, C. A., Goetz, S. J., Donato, D., & Claggett, P. (2010). Designing and implementing a regional urban modeling system using the sleuth cellular urban model. *Computers, Environment and Urban Systems*, *34*(1), 1–16.
Jin, W., Xu, L., & Yang, Z. (2009). Modeling a policy making framework for urban sustainability: Incorporating system dynamics into the ecological footprint. *Ecological Economics*, *68*(12), 2938–2949.
Kashnitsky, I., & Gunko, M. (2016, Nov). Spatial variation of in-migration to moscow: Testing the effect of housing market. *Cities*, *59*, 30–39. doi: 10.1016/j.cities.2016.05.025
Kirbyshire, A., Wilkinson, E., Le Masson, V., & Batra, P. (2017). *Mass displacement and the challenge for urban resilience.* Overseas Development Institute London.
Klabunde, A. (2014). Computational economic modeling of migration. *Available at SSRN 2470525*.
Klabunde, A., & Willekens, F. (2016). Decision-making in agent-based models of migration: state of the art and challenges. *European Journal of Population*, *32*(1), 73–97.
Kloosterman, R., & Lambregts, B. (2001). A brief history of the duality of the hague. *Built Environment (1978-)*, 176–191.
Kloosterman, R. C., & Lambregts, B. (2001). Clustering of economic activities in polycentric urban regions: the case of the randstad. *Urban studies*, *38*(4), 717–732.
Koot, P., van Elk, R., & Jongen, E. (2019). Inkomensongelijkheid naar migratieachtergrond in kaart. *CPB Achtergronddocument*.
Kritsioudi, M. (2015). Living with the others: Spatial transformations towards liveability of cities of social diversity: the case of the schilderswijk, the hague. *TU Delft*.
Lane, D., Pumain, D., van der Leeuw, S. E., & West, G. (2009). *Complexity perspectives in innovation and social change* (Vol. 7). Springer Science & Business Media.
Lansing, J. S. (2003, October). Complex adaptive systems. *Annual Review of Anthropology*, *32*(1), 183–204. Retrieved from https://doi.org/10.1146/annurev.anthro.32.061002.093440 doi: 10.1146/annurev.anthro.32.061002.093440
Leidelmeijer, K., & Van Kamp, I. (2003). Kwaliteit van de leefomgeving en leefbaarheid. *Naar een begrippenkader en conceptuele inkadering, RIVM, Bilthoven.*
Lempert, R. (2002). Agent-based modeling as organizational and public policy simulators. *Proceedings of the National Academy of Sciences*, *99*(suppl 3), 7195–7196.
Lempert, R. J. (2003). *Shaping the next one hundred years: new methods for quantitative, long-term policy analysis.* Rand Corporation.
Lewis, V. A., Emerson, M. O., & Klineberg, S. L. (2011). Who we'll live with: Neighborhood racial composition preferences of whites, blacks and latinos. *Social Forces*, *89*(4), 1385–1407.
Li, Z., Sim, C. H., & Low, M. Y. H. (2006). A survey of emergent behavior and its impacts in agent-based systems. In *2006 4th ieee international conference on industrial informatics* (pp. 1295–1300).
Liu, J., Dietz, T., Carpenter, S. R., Alberti, M., Folke, C., Moran, E., ... Taylor, W. W. (2007, September). Complexity of coupled human and natural systems. *Science*, *317*(5844), 1513–1516. Retrieved from https://doi.org/10.1126/science.1144004 doi: 10.1126/science.1144004
Määttänen, N., & Terviö, M. (2014). Income distribution and housing prices: An assignment model approach. *Journal of Economic Theory*, *151*, 381–410.
MacDonald, J., & Sampson, R. J. (2012). The world in a city: Immigration and america's changing social fabric. *The ANNALS of the American Academy of Political and Social Science*, *641*(1), 6–15.
Maggi, E., & Vallino, E. (2016). Understanding urban mobility and the impact of public policies: The role of the agent-based models. *Research in Transportation Economics*, *55*, 50–59.

Maslow, A. H. (1943). A theory of human motivation. *Psychological review*, *50*(4), 370.
McFarlane, C., & Rutherford, J. (2008). Political infrastructures: Governing and experiencing the fabric of the city. *International journal of urban and regional research*, *32*(2), 363–374.
McKinney, W. (2011). pandas: a foundational python library for data analysis and statistics. *Python for High Performance and Scientific Computing*, *14*(9).
McLeman, R. A., & Hunter, L. M. (2010). Migration in the context of vulnerability and adaptation to climate change: insights from analogues. *Wiley Interdisciplinary Reviews: Climate Change*, *1*(3), 450–461.
Meerow, S., Newell, J. P., & Stults, M. (2016). Defining urban resilience: A review. *Landscape and urban planning*, *147*, 38–49.
Menezes, R., Evsukoff, A., & González, M. C. (Eds.). (2013). *Complex networks*. Springer Berlin Heidelberg. Retrieved from https://doi.org/10.1007/978-3-642-30287-9 doi: 10.1007/978-3-642-30287-9
Mohlmann, J. (2019). *Hoe vaak verhuizen mensen? de cijfers.* Huislijn.nl. Retrieved from https://blog.huislijn.nl/2019/10/hoe-vaak-verhuizen-mensen-de-cijfers/
Nara, A., & Torrens, P. M. (2005). Inner-city gentrification simulation using hybrid models of cellular automata and multi-agent systems. In *Proceedings of the geocomputation 2005 conference. university of michigan, michigan.*
Nations, U. (2020a, Feb). *Goal 11 .:. sustainable development knowledge platform.* Retrieved from https://sustainabledevelopment.un.org/sdg11
Nations, U. (2020b, Feb). *Sdgs .:. sustainable development knowledge platform.* Retrieved from https://sustainabledevelopment.un.org/sdgs
Netto, V. (2017). *The social fabric of cities [new book].*
Newman, M. (2018). *Networks.* Oxford university press.
Nicoletti, L. (2020). *Access to urban infrastructure: Who benefits?* Retrieved from http://leonardonicoletti.com/#/liveability-and-social-inequality-in-canadian-cities/
Nikolic, I. (2019). *Verification of an abm.* https://brightspace.tudelft.nl/d2l/le/content/195109/viewContent/1369712/View. Delft University of Technology.
Nikolic, I., Lukszo, Z., Chappin, E., Warnier, M., Kwakkel, J., Bots, P., & Brazier, F. (2019). *Guide for good modelling practice for policy support.* Delft University of Technology. Retrieved from http://resolver.tudelft.nl/uuid:cbe7a9cb-6585-4dd5-a34b-0d3507d4f188 doi: 10.4233/UUID:CBE7A9CB-6585-4DD5-A34B-0D3507D4F188
Nwankwo, S., Hamelin, N., & Khaled, M. (2014). Consumer values, motivation and purchase intention for luxury goods. *Journal of retailing and consumer services*, *21*(5), 735–744.
of Technology, D. U. (2020). *Hpc cluster tu delft.* Retrieved from https://login.hpc.tudelft.nl/
of The Hague, M. (2020). *Dataplatform den haag.* Retrieved from https://denhaag.dataplatform.nl/
Ontwikkeling, D. S. (2012). *Citizen prognosis 2012 - 2020.*
Ormerod, P., & Rosewell, B. (2006). Validation and verification of agent-based models in the social sciences. In *International workshop on epistemological aspects of computer simulation in the social sciences* (pp. 130–140).
Orozco, L. G. N., Deritei, D., Vancso, A., & Vasarhelyi, O. (2019). Quantifying life quality as walkability on urban networks: The case of budapest. In *International conference on complex networks and their applications* (pp. 905–918).
Perez, L., Dragicevic, S., & Gaudreau, J. (2019). A geospatial agent-based model of the spatial urban dynamics of immigrant population: A study of the island of montreal, canada. *PloS one*, *14*(7).
Pitner, R. O., Yu, M., & Brown, E. (2012). Making neighborhoods safer: Examining predictors of residents' concerns about neighborhood safety. *Journal of Environmental Psychology*, *32*(1), 43–49.
Poston, B. (2009). Maslow's hierarchy of needs. *Surgical technologist*, *41*(8), 347–353.
Pullum, L. L., & Cui, X. (2012). Techniques and issues in agent-based modeling validation. *Technique Report*, 5–6.
Railsback, S. F., Ayllón, D., Berger, U., Grimm, V., Lytinen, S., Sheppard, C., & Thiele, J. (2017). Improving execution speed of models implemented in netlogo. *Journal of Artificial Societies and Social Simulation*, *20*(1), 3. doi: 10.18564/jasss.3282
Ramachandra, T., Aithal, B. H., & Sanna, D. D. (2012). Insights to urban dynamics through landscape spatial pattern analysis. *International Journal of Applied Earth Observation and Geoinformation*, *18*, 329–343.
Ritchie, H., & Roser, M. (2020). Urbanization. *Our World in Data.* (https://ourworldindata.org/urbanization)
Saltelli, A., & Annoni, P. (2010). How to avoid a perfunctory sensitivity analysis. *Environmental Modelling & Software*, *25*(12), 1508–1517.
Sayama, H. (2015). *Introduction to the modeling and analysis of complex systems.* Open SUNY Textbooks.
Schelling, T. (1978). Micromotives and macrobehavior. *New York.*

Scheutz, M., & Mayer, T. (2016). Combining agent-based modeling with big data methods to support architectural and urban design. In *Understanding complex urban systems* (pp. 15–31). Springer.

Squazzoni, F., Jager, W., & Edmonds, B. (2014). Social simulation in the social sciences: A brief overview. *Social Science Computer Review*, *32*(3), 279–294.

Statline, C. (2020a). Centraal bureau voor de statistiek. *Statline Open Data.* Retrieved from http://statline.cbs.nl/StatWeb/publication

Statline, C. (2020b). *Kerncijfers wijken en buurten 2015-2019.* Den Haag/Heerlen: Centraal Bureau voor de Statistiek.

Sturgis, P., Brunton-Smith, I., Kuha, J., & Jackson, J. (2014). Ethnic diversity, segregation and the social cohesion of neighbourhoods in london. *Ethnic and Racial Studies*, *37*(8), 1286-1309. doi: 10.1080/01419870.2013.831932

Suleimenova, D., Bell, D., & Groen, D. (2017). A generalized simulation development approach for predicting refugee destinations. *Scientific reports*, *7*(1), 13377.

Teddlie, C., & Tashakkori, A. (2003). *Handbook of mixed methods in social & behavioral research.* Sage.

Ten Broeke, G., Van Voorn, G., & Ligtenberg, A. (2016). Which sensitivity analysis method should i use for my agent-based model? *Journal of Artificial Societies and Social Simulation*, *19*(1).

Thiele, J. C., & Grimm, V. (2010). Netlogo meets r: Linking agent-based models with a toolbox for their analysis. *Environmental Modelling & Software*, *25*(8), 972–974.

Tomasiello, D. B., Giannotti, M., & Feitosa, F. F. (2020). Access: An agent-based model to explore job accessibility inequalities. *Computers, Environment and Urban Systems*, *81*, 101462.

Toole, T. M. (2005). A project management causal loop diagram. In *Arcom conference, london, uk, sep* (pp. 5–7).

Torrens, P. M. (2003). Automata-based models of urban systems. *Advanced spatial analysis*, 61–79.

Turbek, S. P., Chock, T. M., Donahue, K., Havrilla, C. A., Oliverio, A. M., Polutchko, S. K., . . . Vimercati, L. (2016). Scientific writing made easy: A step-by-step guide to undergraduate writing in the biological sciences. *The Bulletin of the Ecological Society of America*, *97*(4), 417–426.

United-Nations, U. (2019). *World urbanization prospects: the 2018 revision.* United Nations and Department of Economic and Social Affairs and Population Division.

van Amsterdam, H., Beets, G., Bontekoning, R., van Dam, F., de Groot, C., Hofman, L., . . . van Middelkoop, M. (2015). *De stad: magneet, roltrap en spons. bevolkingsontwikkelingen in stad en stadsgewest.* Planbureau voor de Leefomgeving (PBL).

van Deen, J., Voskamp, W., & Bezuijen, A. (2018). Geokunststoffen en het milieu. *Deltares.* Retrieved from http://ngo.nl/wp-content/uploads/2019/11/Jurjen-van-Deen-geokunststoffen-en-het-milieu.pdf

van de Statistiek, C. B. (2020a). *Grootste huurstijging in zes jaar.* CBS. Retrieved from https://www.cbs.nl/nl-nl/nieuws/2020/37/grootste-huurstijging-in-zes-jaar

van de Statistiek, C. B. (2020b). *Woningmarkt analyse in coronatijd.* CBS. Retrieved from https://www.cbs.nl/nl-nl/visualisaties/welvaart-in-coronatijd/woningmarkt/

van Tongeren, S. A. (2014). Creating a conceptual framework for a deeper understanding of evolving processes in socio-technical systems. *world*.

Verkade, T., & Brommelstroet, M. (2020). *Het recht van de snelste.* De Correspondent.

Verma, T. (2020). *A guide to writing scientific text.* Author. Retrieved from https://research.trivikverma.com/writing/

Vlug, J. (2020). *Impact of migration on cities, an agent-based model on the city of den haag.* https://github.com/Jochem285/OpenDataDenHaag. GitHub.

Wenge, R., Zhang, X., Dave, C., Chao, L., & Hao, S. (2014). Smart city architecture: A technology guide for implementation and design challenges. *China Communications*, *11*(3), 56–69.

Wilensky, U. (1999). *Netlogo.* CCL Northwestern. Retrieved from http://ccl.northwestern.edu/netlogo/

Wilensky, U., & Payette, N. (1998). Netlogo traffic 2 lanes model. *Center for Connected Learning and Computer-Based Modeling.* Retrieved from http://ccl.northwestern.edu/netlogo/models/Traffic2Lanes

Wilensky, U., & Rand, W. (2007). Making models match: Replicating an agent-based model. *Journal of Artificial Societies and Social Simulation*, *10*(4), 2.

Wilensky, U., & Rand, W. (2015). *An introduction to agent-based modeling: modeling natural, social, and engineered complex systems with netlogo.* Mit Press.

Winchie, D. B., & Carment, D. W. (1989). Migration and motivation: The migrant's perspective. *International Migration Review*, *23*(1), 96–104.

Windrum, P., Fagiolo, G., & Moneta, A. (2007). Empirical validation of agent-based models: Alternatives and prospects. *Journal of Artificial Societies and Social Simulation*, *10*(2), 8.

Yang, Y., Roux, A. V. D., Auchincloss, A. H., Rodriguez, D. A., & Brown, D. G. (2011). A spatial agent-based model for the simulation of adults' daily walking within a city. *American journal of preventive medicine*, *40*(3), 353–361.

Yazgan, P., Utku, D. E., & Sirkeci, I. (2015). Syrian crisis and migration. *Migration Letters*, *12*(3), 181–192.

Yuan, X., Ji, X., Chen, H., Chen, B., & Chen, G. (2008). Urban dynamics and multiple-objective programming: a case study of beijing. *Communications in Nonlinear Science and Numerical Simulation*, *13*(9), 1998–2017.

Zhang, M. (2016). Large-scale agent-based social simulation: A study on epidemic prediction and control. *TU Delft*.

Zöllig, C., & Axhausen, K. W. (2011). A conceptual, agent-based model of land development for urbansim. *Arbeitsberichte Verkehrs-und Raumplanung*, *717*.

Zuk, M., Bierbaum, A. H., Chapple, K., Gorska, K., Loukaitou-Sideris, A., Ong, P., & Thomas, T. (2015). Gentrification, displacement and the role of public investment: a literature review. In *Federal reserve bank of san francisco* (Vol. 79).

Appendix A

Appendix I: Time Planning

In Figure A.1, the time planning for the research is shown. The formal meetings are marked in red. A week prior formal meetings, the progress or report that is expected with the meeting will be submitted. The exact dates for greenlight meeting and thesis defence are "to be determined" and will be shared in the mid-term presentation.

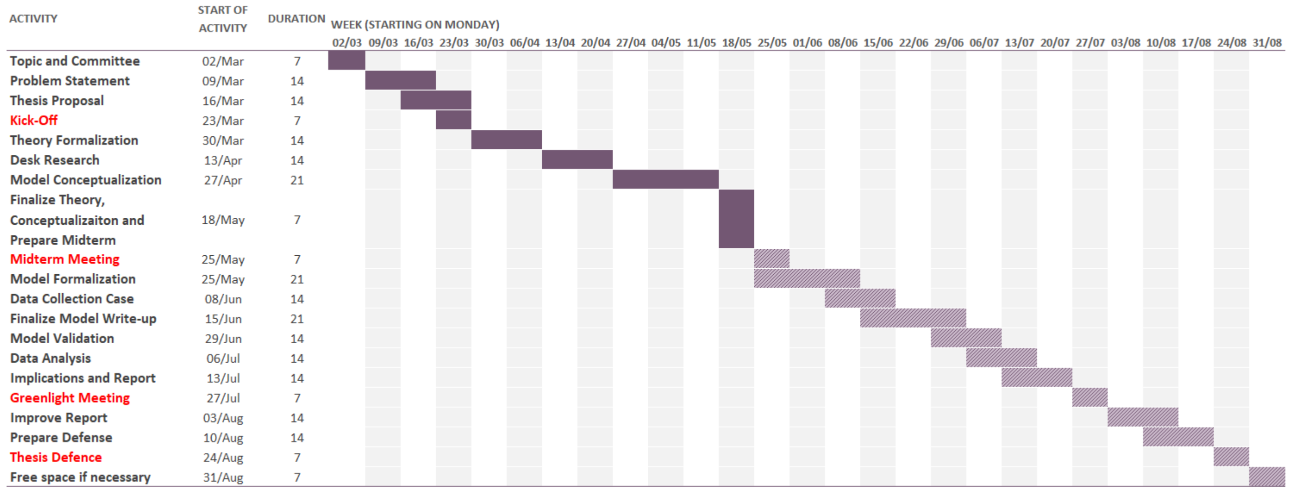


Figure A.1: Time planning thesis proposal. *Formal meetings are highlighted in red.*

Appendix B

Research Approach

B.1 Desk Research / Theory Formalization

First, a ground basis of theory is needed to better understand what the problem is and also to understand what the conceptual model needs to encompass. By reviewing literature and performing a desk research, a better understanding of the concepts of city science is gained. Furthermore, this phase of the project formulates the basis of concepts that are part of the conceptual model. The elements and interactions that define the (changes to the) urban fabric are looked at in literature. By looking at models in current literature, and what has been (and hasn't been) done before, a good basis of current knowledge and lack thereof can be formulated. At the end of this phase, all the theory that is needed to understand both the problem and all the concepts and interactions that will be described in the conceptual model should be gathered.

B.2 Conceptual Model

During this phase, the knowledge gathered in the literature study and desk research is used to formulate a conceptual model. This model is used as a framework which defines the basic concepts of the urban fabric. This conceptual model is formulated in such a way, that it is applicable to any city and should only consist of the components that are present in each type of city that is observed.

B.3 Data gathering and preparation

After defining the conceptual model, case study data is collected to use as an input when creating the agent-based model later on. During this phase, data from open data sources is used, which has been described in Chapter 2. Since the data has already been collected by the efforts of scientific research from institutes such as CBS and the municipality of The Hague, no further collection of data from the real-world is necessary (Statline, 2020a; of The Hague, 2020; Statline, 2020b).

B.4 Agent-Based Model

After defining the (cor)relation between citizen composition and supply of amenities, an Agent-Based model (ABM) will be made. Using the concepts from the conceptual model as a framework and applying the effects observed in historical data, the elements and interactions are simulated using Agent-Based modeling. By using data from the city of The Hague as an input, a model can be made. The creation of said model requires different steps, which are structured using the Guide for Good Modeling Practices (Nikolic et al., 2019). In regards to the creation of the ABM, the different steps involved are respectively conceptualisation, formulation, implementation, verification, and validation (Nikolic et al., 2019).

B.5 Data Analysis

After the model has been made and tested, the results from the model can be interpreted. In this phase, the created model is used to observe model behavior using data analysis (Thiele & Grimm, 2010). By applying statistical analyses, the observed behavior from the experimentation is explored. Furthermore, in this phase the model will simulate an extreme influx of migrants and the changes to the fabric of the city in the simulation are observed. Doing so, gives insight in the possible future and the way the fabric of a city can change. To conclude, results regarding the quantitative analyses produced in this phase are reported.

B.6 Policy Analysis

In the last phase of the project, the results from the data analysis are considered regarding the implications for policy recommendations. Furthermore, conclusions will be drawn regarding the usefulness of the produced conceptual framework. By looking at the results from both the conceptual as well as the simulation model, recommendations will be made for future research and for implications on the city of The Hague(R. Lempert, 2002; Janssen & Ostrom, 2006; Dawid & Neugart, 2011).

B.7 Time Planning

To further elaborate on the planning of the thesis project and obtain a better temporal sense of the project and its phases, an indicative time planning has been made which can be found in Figure A.1 (Appendix I). In this planning, the three major events are the Kick-off meeting, Midterm meeting and Green light meeting. These meetings respectively align the topic and research question, making sure the research is on progress and finally check if the research that was done is sufficient for a Master's Thesis.

B.8 Research Flow

To better understand the flow and the sub goals of this research, an overview is shown in Figure B.1. In this research flow diagram, the aforementioned phases of research are described as well as the relation between the different parts of the research.

B.9 Reporting

The results, findings and conclusions of the research will be reported in a Master Thesis Document. The reporting style of the document will follow a scientific structure to be as informative as possible (Verma, 2020; Turbek et al., 2016).

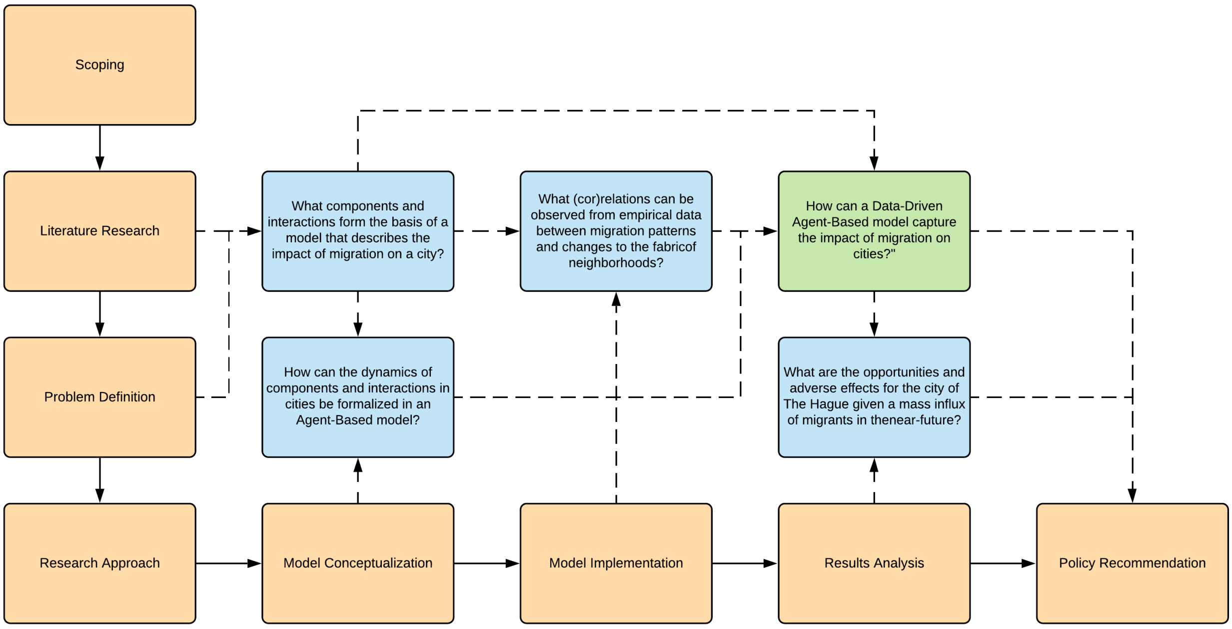


Figure B.1: Research Flow Diagram, depicting the course of research. Each step of the research (highlighted in orange) leads to answers to the sub research questions (in blue) which then leads to an answer to the main research question (in green) and policy recommendations.

Appendix C

Background Literature

This chapter highlights literature, models and research which is out of scope for the development of an Agent-Based Model which looks at the influx of migrants on the meso-level and observes the resulting changes to the city. However, the literature in this chapter does describe many dynamics of cities that are important to better understand the way a city develops over time, and what role the citizens in the city play in this development.

C.1 Transport & Mobility

A Review of literature applying modeling techniques for Transport & Mobility using ABM was made, showing many models are looking at the interactions of agents on the micro-scale as well as some on the macro-scale (Bazzan & Klügl, 2014). This macro-scale is more apparent in research in order to assist policy making regarding mobility (Maggi & Vallino, 2016). Furthermore, simulations were made to calculate human mobility behavior (Barbosa et al., 2018). Some models even go as far as making a complete virtual city to simulate mobility behavior and interaction of citizens (Ge, Meng, Cao, Qiu, & Huang, 2014). When looking at more granular models, a micro-scale is applied to model walking patterns and behaviors of citizens (Yang, Roux, Auchincloss, Rodriguez, & Brown, 2011). ABM is also combined with Discrete-Event Simulation (DEVS) to simulate mobility behaviors and the influence of interactions of agents in an urban area on the spread of an infectious disease (Zhang, 2016).

C.2 Urban Planning & Housing

An overview of literature on ABMs in urban research is presented by Chen, some of which focus on spatial planning and housing on a macro-scale (Chen, 2012). A more refined meso-scale of urban planning is seen in a research of the housing market and the influence of real-estate behavior on urban planning in Vienna, which is using ABM and housing data from big housing brokers (Gaube & Remesch, 2013). Similairly, Hosseinali, Alesheikh, and Nourian made an ABM on land-use in Qazvin, Iran (Hosseinali et al., 2013). Heppenstall, Crooks, See, and Batty made an overview of ABMs used in geographical systems, and uses *the New Science of Cities* as theoretical framework (Heppenstall et al., 2011). By combining ABM with Big Data (and data regression, Machine Learning), Scheutz and Mayer attempts to better understand architectural and urban design choice behavior in cities (Scheutz & Mayer, 2016). By looking at satelite imagery, landscape spatial pattern analysis was used to better understand the macro-scale behavior of spatial changes (Ramachandra, Aithal, & Sanna, 2012). Zöllig and Axhausen studied the interactions in land development for urbanism using ABM (Zöllig & Axhausen, 2011). A much broader approach on urban development was used by Yuan, Ji, Chen, Chen, and Chen to model macro-scale changes to population, resources, energy, economy, environment and ecosystem in Beijing, China using System Dynamics (SD) and multi-objective programming (ISDMOP) (Yuan et al., 2008). A literature review of ABM models in social-ecological systems (such as cities) displayed that ABMs can also greatly benefit the research of ecology (Filatova, Verburg, Parker, & Stannard, 2013). Lastly, an overview of papers on modeling social sciences using "social simulation" was made, which provides helpful insights in the behavior of citizens (Squazzoni, Jager, & Edmonds, 2014).

Appendix D

Neighborhoods of The Hague

This appendix highlights the different neighborhoods of The Hague, their names, size, location and how they are used in the Agent-Based Model. This is done to prevent confusion, since the naming and numbering scheme for the neighborhood does not necessarily follow a "logical" order.

The overview of neighborhoods and respective properties of neighborhoods is based on the data of CBS research from 2019. This is chosen because it is the most recent dataset that is complete and shows the most representative overview of the city.

First, an overview is presented by plotting the map of The Hague. Using Python, and some geoplot packages, the city is plotted with their respective neighborhood "codes" inside each of the neighborhoods. Furthermore, the population size of each neighborhood is represented using a YlGn colormap, the hue is explained using the legend at the right side of the plot. Figure D.1 shows this plot.

Next, more detailed information of each neighborhood is presented using a table, in which the full name, neighborhood code, population size and "wijkcode" are presented. The "wijkcode" represents codes for aggregate clusters of neighborhoods. Most neighborhoods are better known for their "Wijk" name locally, however, for completeness of the data, the table presents the name based on the most granular level, the neighborhood or "buurt" level. This is shown in Table D.1.

Table D.1: List of all neighborhoods of The Hague.

BUURTCODE	WIJKCODE	buurtname	total_citizens
1	7	Oud Scheveningen	2915
2	7	Vissershaven	3870
3	7	Scheveningen Badplaats	5495
4	7	Visserijbuurt	3885
5	6	Van Stolkpark en Scheveningse Bosjes	765
6	18	Waldeck-Zuid	2080
7	9	Statenkwartier	9705
8	9	Geuzenkwartier	4215
9	13	Vogelwijk	5335
10	21	Rond de Energiecentrale	6250
11	28	Kortenbos	7300
12	28	Voorhout	2430
13	28	Uilebomen	3215
14	28	Zuidwal	6910
15	29	Schildersbuurt-West	14220
16	29	Schildersbuurt-Noord	9980
17	29	Schildersbuurt-Oost	7185
18	27	Huygenspark	7870

Continued on next page

Table D.1: List of all neighborhoods of The Hague.

BUURTCODE	WIJKCODE	buurtname	total_citizens
19	38	Laakhaven-Oost	4085
20	36	Moerwijk-Oost	3245
21	37	Groente- en Fruitmarkt	5290
22	38	Laakhaven-West	5960
23	38	Spoorwijk	4350
24	38	Laakkwartier-West	8135
25	38	Laakkwartier-Oost	11330
26	38	Noordpolderbuurt	8390
30	31	Rustenburg	6095
31	31	Oostbroek-Noord	4525
32	30	Transvaalkwartier-Noord	3940
33	30	Transvaalkwartier-Midden	5015
34	30	Transvaalkwartier-Zuid	7220
35	31	Oostbroek-Zuid	8075
36	35	Zuiderpark	155
37	36	Moerwijk-West	6245
38	36	Moerwijk-Noord	7035
39	36	Moerwijk-Zuid	4765
40	18	Nieuw Waldeck	7285
41	10	Zorgvliet	530
42	11	Stadhoudersplantsoen	2025
45	22	Zeeheldenkwartier	12085
46	5	Archipelbuurt	6160
47	23	Willemspark	1590
48	4	Nassaubuurt	1610
49	24	Haagse Bos	445
50	12	Bloemenbuurt-West	2305
51	12	Bloemenbuurt-Oost	6495
52	12	Bomenbuurt	5940
53	19	Vruchtenbuurt	7215
54	20	Heesterbuurt	6820
55	20	Valkenboskwartier	11425
60	39	Binckhorst	2275
61	25	Landen	5205
62	27	Rivierenbuurt-Zuid	850
63	27	Rivierenbuurt-Noord	3650
64	26	Bezuidenhout-West	3610
65	26	Bezuidenhout-Midden	4500
66	26	Bezuidenhout-Oost	9045
67	25	Kampen	2450
68	25	Marlot	795
69	25	Burgen en Horsten	6315
70	1	Oostduinen	0
71	2	Belgisch Park	8245
72	7	Rijslag	1605
73	3	Westbroekpark	930
74	3	Duttendel	1075
75	4	Uilennest	2105
76	4	Duinzigt	2525
77	4	Waalsdorp	3905
78	4	Arendsdorp	1365
80	34	Morgenstond-Zuid	6205

Continued on next page

Table D.1: List of all neighborhoods of The Hague.

BUURTCODE	WIJKCODE	buurtname	total_citizens
81	14	Bosjes van Pex	370
82	18	Rosenburg	2915
83	19	Eykenduinen	2615
84	32	Leyenburg	15010
85	17	Kerketuinen en Zichtenburg	55
86	17	Houtwijk	12890
87	33	Venen, Oorden en Raden	8730
88	34	Morgenstond-West	7505
89	34	Morgenstond-Oost	5630
90	15	Ockenburgh	780
91	15	Kijkduin	1565
92	14	Bohemen en Meer en Bos	4425
93	18	Componistenbuurt	1825
94	18	Waldeck-Noord	3150
95	17	Kom Loosduinen	5025
96	33	Zijden, Steden en Zichten	7955
97	16	Kraayenstein en Vroondaal	6255
98	33	Dreven en Gaarden	10820
99	33	De Uithof	1525
100	8	Duindorp	5875
101	40	Erasmus Veld	2165
102	40	Hoge Veld	8890
104	40	Lage Veld	4270
105	40	Zonne Veld	3405
106	41	Vlietzoom-West	165
107	41	Vliegeniersbuurt	0
108	42	Bosweide	2025
110	41	De Reef	95
111	42	De Venen	1790
112	42	Morgenweide	7140
113	42	Singels	5595
114	42	Waterbuurt	4810
115	42	De Bras	5850
117	43	De Rivieren	40
118	44	De Lanen	5440
119	44	De Velden	4200
120	44	De Vissen	8620
121	44	Rietbuurt	2795

It might be noted that 1) some neighborhood numbers are "missing" and 2) some neighborhoods have a population of 0 (or a very low number). The first observation can be explained by the way the municipality has chosen to number the different neighborhoods, and has a historic reason which is not relevant for this research. The second observation can be explained by the fact that there are several "neighborhoods" which are in fact industrial sites, big parks, or sporting facilities. For example, the "Vliegeniersbuurt" consist solely of industrial and commercial buildings, and has no residents.

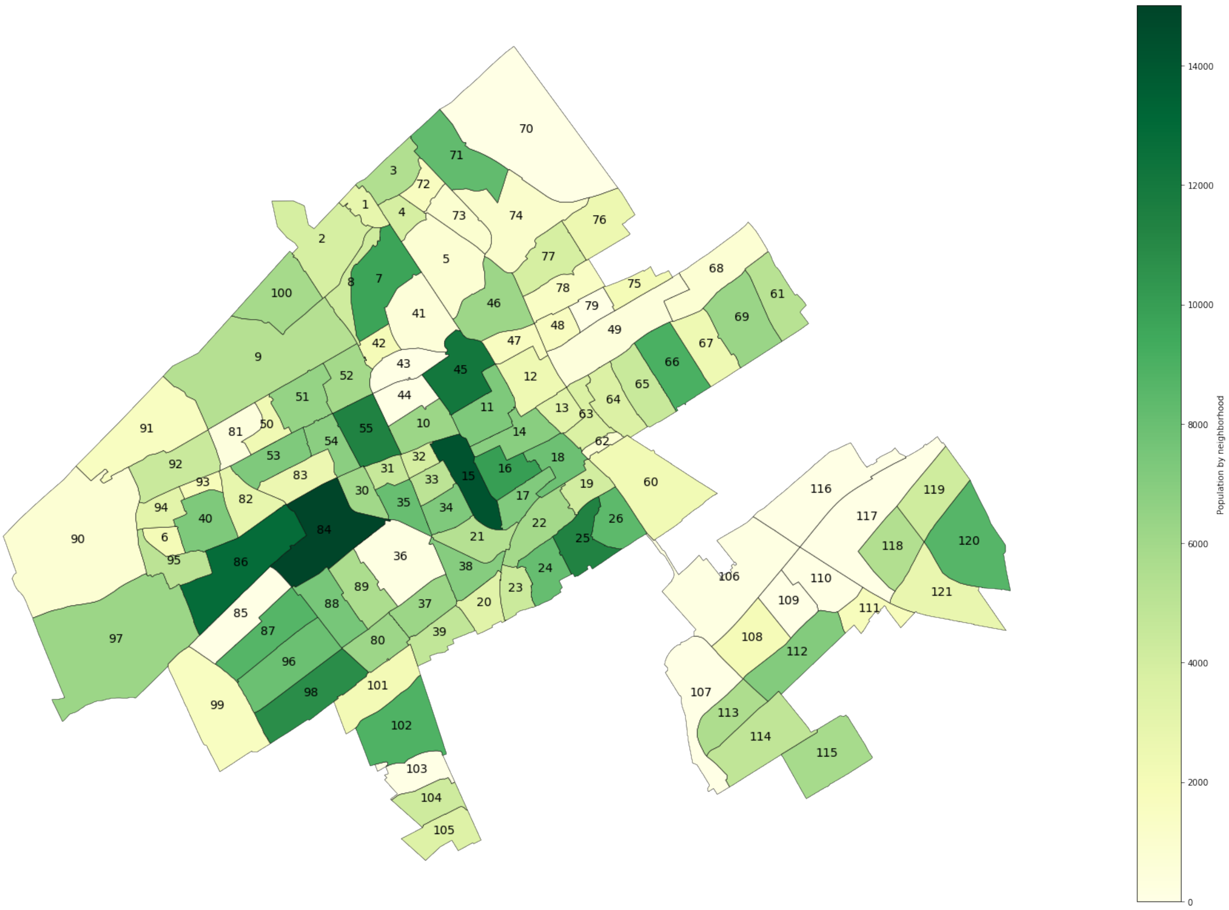


Figure D.1: Map of the city of The Hague, using the unique neighborhood identifier codes as labels. The color represents the population count of each neighborhood.

Appendix E

UML

This UML diagram is a work in progress, since the model narrative has not been fully completed yet. It should give an overview of the system that is going to be modeled.

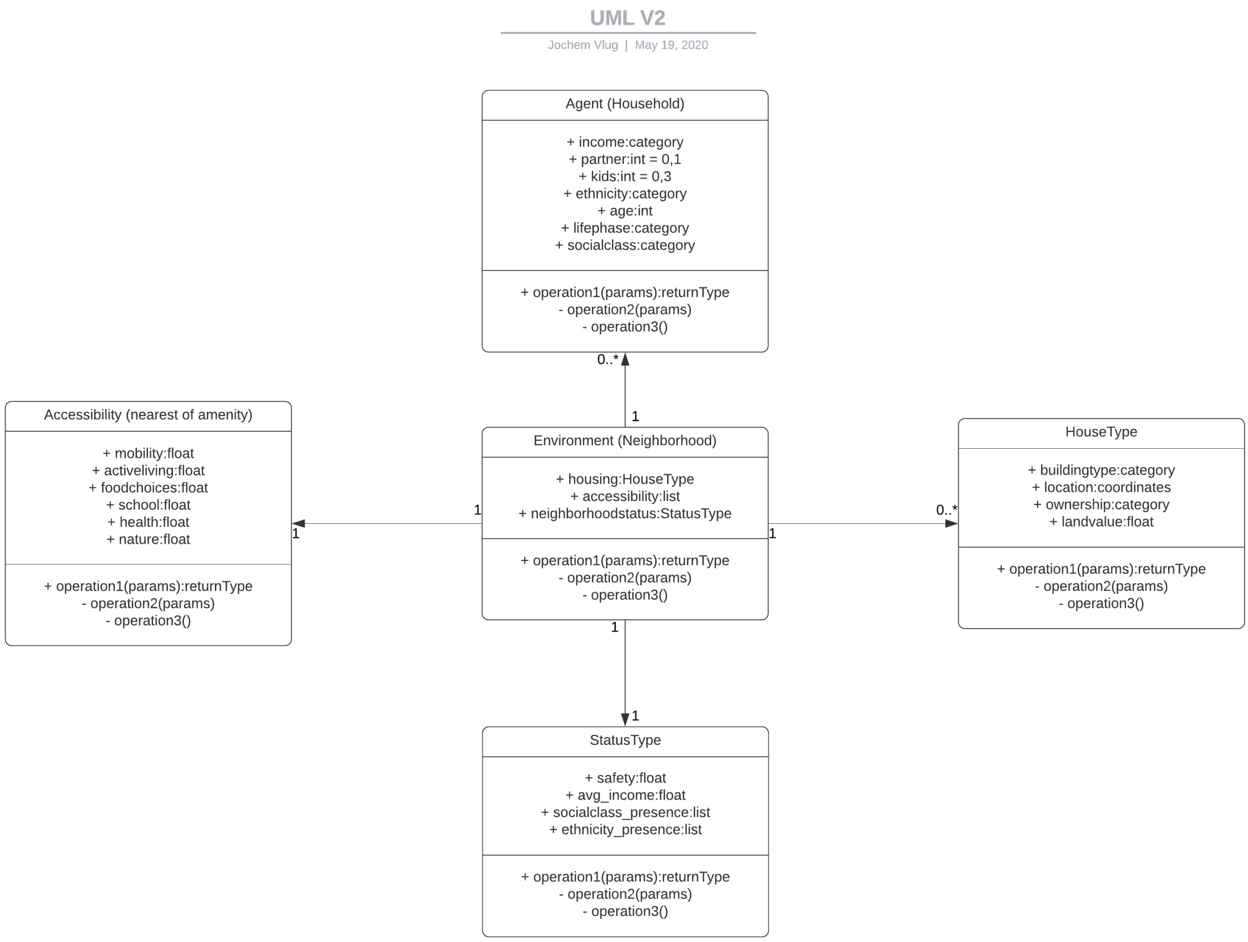


Figure E.1: UML diagram

Appendix F

Social Group Composition

This appendix highlights the way the social groups are defined and how the data leads to categorization of citizens to the different social groups.

F.1 Social Groups

There are many theories on social class, categorizing the different classes in quantities between 2 (Karl Marx) and more than 10 in some economical theories. Influential work on the definition of social class comes from economists and sociologists such as Max Weber (Gane, 2005).
Based on work from economic theory and for the sake of simplicity of categorizing from data, it is assumed there are four different social groups based on their needs: those with basic needs, stability needs, amenity needs and extended needs (Argyle, 1994). By categorizing the needs of people, the different groups can account for the decision-making behavior of agents in the model. Using this grouping, the importance of certain aspects becomes increasingly more important whether someone can "afford" certain facilities or amenities (Nwankwo, Hamelin, & Khaled, 2014).
Using data from a real-world opinion research in the city of The Hague, a similar relation between the importance of certain needs and the affordability of amenities was found. Using this information, the categorization of social groups can be made (Center, 2020).

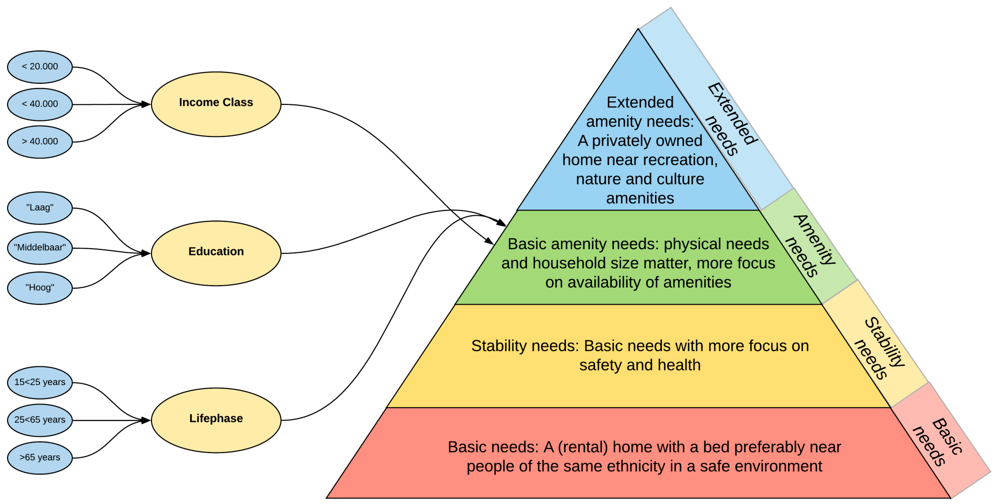


Figure F.1: Composition of Social Groups

F.2 Factors for Categorization

Looking at the available data, three factors are selected as predictors of one's social grouping: Income, Education and Age. By looking at these categories it is possible to categorize citizens into the different groups. An overview of the factors included is shown in Figure F.1.

F.3 Calculation of Categorization

A decision tree is made to calculate the affordability of needs for all citizens within the data and which social groups they can be categorized in. This decision tree includes the three factors (age, education, income) in such a way that all citizens can easily be categorized. The process of categorization happens during the simulation of the model, in such a way that it is possible to define the social group of new agents entering the model, such as migrants.

It should be noted that this does not necessarily reflect the complete complexity of the real-world, but is a fair assumption for the sake of simulating behavior within the model. For simplicity's sake, this grouping can solve the trade-off between over fitting the complexity of decision-making and under representation of social mechanics in the psychological rationale of citizens. Figure F.2 shows the decision tree for categorizing agents in social groups with differences in their affordability of needs.

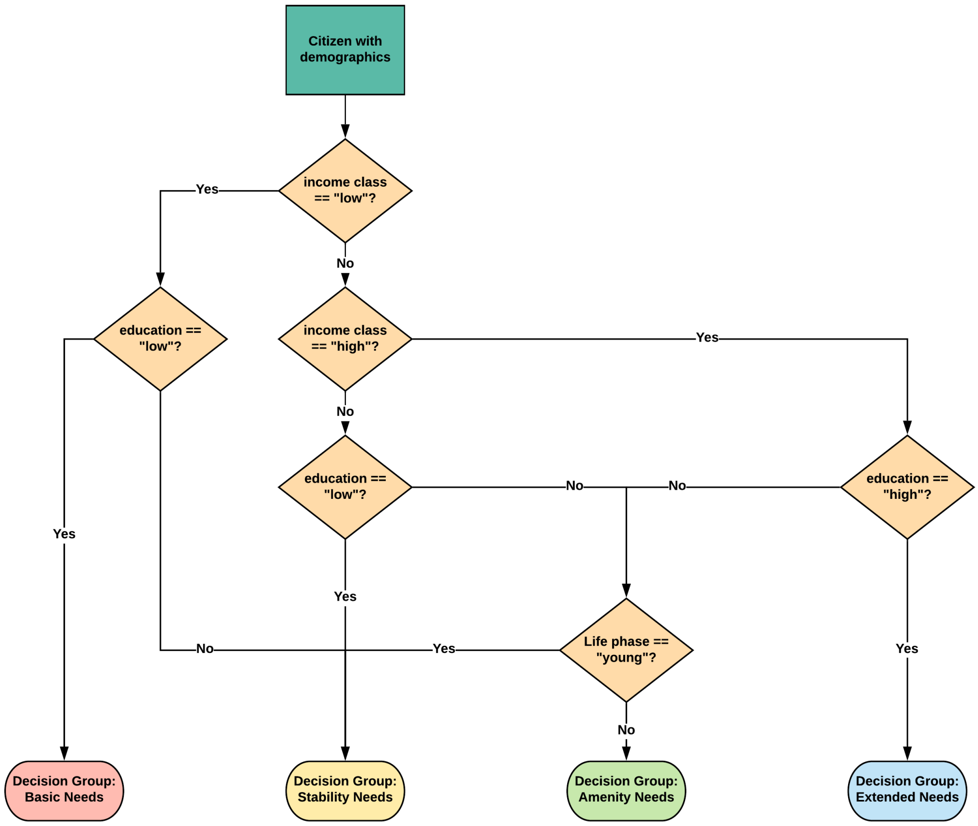


Figure F.2: Decision Tree of Social Groups

A minor side note: In the model, the same logic has been applied but uses a different terminology. To make it easier to quickly differentiate the different social groups, the groups are named lower, working, middle and upper class in the model which corresponds to social class theory of Argyle. This terminology does not reflect the actual categorization into different classes and is meant to display the concepts explained in this appendix chapter.

Appendix G

Assumptions

This appendix sums a list of all assumptions made in the model formulation and conceptualization steps. These assumptions are made because it is impossible to completely model real-life. It is important to record the assumptions, to be aware of the aggregation of simulation and be able to explain certain results or behavior in the model (Nikolic et al., 2019).

1. The time step of the model is 3 months per tick. And the model runs for 30 ticks, which is a time span of 10 years.
2. The model starts in time where most data has already been collected for the city of The Hague, which is the year 2019.
3. The value of data for properties that are static (such as expected influx of migrants) is calculated using historical data.
   (a) This calculation relies on linear regression (extrapolation of data).
4. Citizen agents are modeled as an aggregate for the complete household.
   (a) Because of social complexity, it is assumed households do not split up within the time frame of the simulation.
   (b) The size of the household is not significant for the modeling purpose, since only the state of housing is observed.
   (c) Average household size of neighborhoods are therefore static and gathered from data.
   (d) It is assumed exactly 1 household agent can occupy each housing availability.
5. All citizens in the model are assumed to have an age of 0 to 100 years.
6. The age of agents does not change during the model, as the impact of this change is negligible because of the short simulation time of the run.
7. Because of the short time scale of the model, it is assumed citizens do not change in education, lifephase, income or social class.
8. Based on the social group a citizen is in, social cohesion with other citizens is calculated using either the prevalence of similar ethnicity or social group (see also Chapter F).
   (a) This implies that in the model, social cohesion does not reflect an individual's ability to connect to others, it reflects the presence of similar people in their vicinity.
9. The amount of crimes occurring is static and equal every year. Because of a lack of data and insight in the dynamics, this is out of scope of the research.
10. Citizens are categorized in life phases based on their age, 0-25 (young), 25-65 (middle) and 65+ (old).
11. The lifephase of a household is based on the age of the agent it represents, therefore an age below 18 assumes a young household where at least 1 citizen is an adult.
12. Agents will only decide to move if there is a significant change in their life that causes them to reconsider.

13. Citizen agents cannot die or be born in the model. To cope with the natural change of citizens over time, the availability of housing changes in accordance with the average natural change of a neighborhood.
    (a) This natural change is gathered from data, and linear regression is used to extrapolate natural change in the future.
14. The distribution of citizen's income follows a Normal distribution. The spread of income is simulated using a standard deviation of 0.05 * the average of incomes of the neighborhood (see Chapter L).
15. Citizens can not move out of the city in the model, only move to another neighborhood within The Hague.
    (a) Based on statistics from CBS, it is assumed people move 7 times in their lifetime (Mohlmann, 2019; CBS, 2019). Since the moving happens more in younger phases of life, and most citizens are of the age group with where most moving occurs, it is assumed people move approximately once every 8 years.
    (b) Because of this assumption, the calculation for deciding to attempt to move is based on a random function with a chance of 1/32. Because each ticks models 3 months (1/4th of a year) and the average movement happens once every 8 years.
16. For simplification and a lack of granular data, it is assumed that the percentage of free houses in a neighborhood is evenly distributed between houses up for sale and for rent.
17. The municipality of The Hague is able to improve the city by means of policies.
18. There is no money involved in the performance of policies, and no budget is limiting the amount the municipality can do.
19. The feasibility otherwise for each project is assumed to be out of scope as well. For example, it is assumed there is always enough space to build new homes when this policy is active.
    (a) This is because the goal of the model is show effectiveness of policies, not their cost efficiency.
    (b) The amount of change a policy can bring (such as the amount of new houses that can be built in a year) is based on empirical data and is estimated.
    (c) The influence of the numbers for these amounts is varied to check the sensitivity of outcomes in the sensitivity analysis.
20. When migrants enter the model, they spawn in as homeless and directly start looking for suitable homes.
21. There is no restriction for migrants to start looking for homes.
22. Migrants start with an income based on an assumed average of 19.000 Euros.
23. There is no care for migrants that cannot find/afford a home, thus all migrants that fail to do so are homeless.
24. Similarly, if citizens get pushed out of a certain neighborhood and are unable to afford other housing options, they too become homeless.
25. At the end of each year, the average pricing of houses adjusts its value to the average income of its residents. This happens each year in model simulation.
26. At the end of each year, all citizens of a neighborhood check if they are still able to afford to live in a certain neighborhood. If their budget is lower than the (new) average housing price, they immediately move out.

Appendix H

Model Narrative

This chapter outlines the model narrative, which is a list of all procedures in the simulation model outlined in such a fashion and in such detail, that it can be read as a "script" or story of what is happening in the model. Part of the conceptualisation of the model is creating the model, and by creating the model narrative, the modeller gets insight into all the components that are necessary to model the system.

H.1 The actors

First, the model narrative defines the actors that are part of the system and relevant to include in the model to observe their behaviour. Chapter 3.1 outlines the included actors within the system, and Table H.1 sums up all included actors and a short description of how they are included in the model.

Actor	Role	Short Description
Citizens of The Hague	Agent in the model	Citizens of The Hague are modeled as agents, where each agent represents a household within the city.
Migrants	Agent in the model	Migrants arriving in the city of The Hague are modeled in the same fashion as current citizens, but have an extra property that keeps track of the fact that they have arrived as migrants.
Neighborhoods	Agent in the model	The averages of city properties are simulated on the neighborhood aggregation level. For each neighborhood, the amount of available amenities and averages of citizen statistics are tracked.
Municipality of The Hague	Can alter model behavior	By constructing policy levers, which change properties of the model, the influence the municipality can have on the city is modeled. For instance, a policy lever can be adopted that alters the availability of new houses, to model a policy of facilitating housing construction.
Uncertainty / Parameters	Can alter model behavior	To see the influence of certain (uncertain) properties of the model, parameters can be altered to check the change to the behavior of the model. For example, the expected influx of migrants can be altered and the resulting outcomes can be observed.

Table H.1: Actors in the Agent-Based model, as defined in the model narrative.

H.2 Model Procedure

Figure H.1 outlines the model procedure, which show the input of the model, the initialization, the model simulation loop and the output of the model. After the output is collected using the *Behaviorspace* tool in Netlogo, the data is analyzed using Python with packages such as *OS, Pandas, Numpy, Geopandas* and visualized using

MatplotLib, *Seaborne* and *Plotly*.

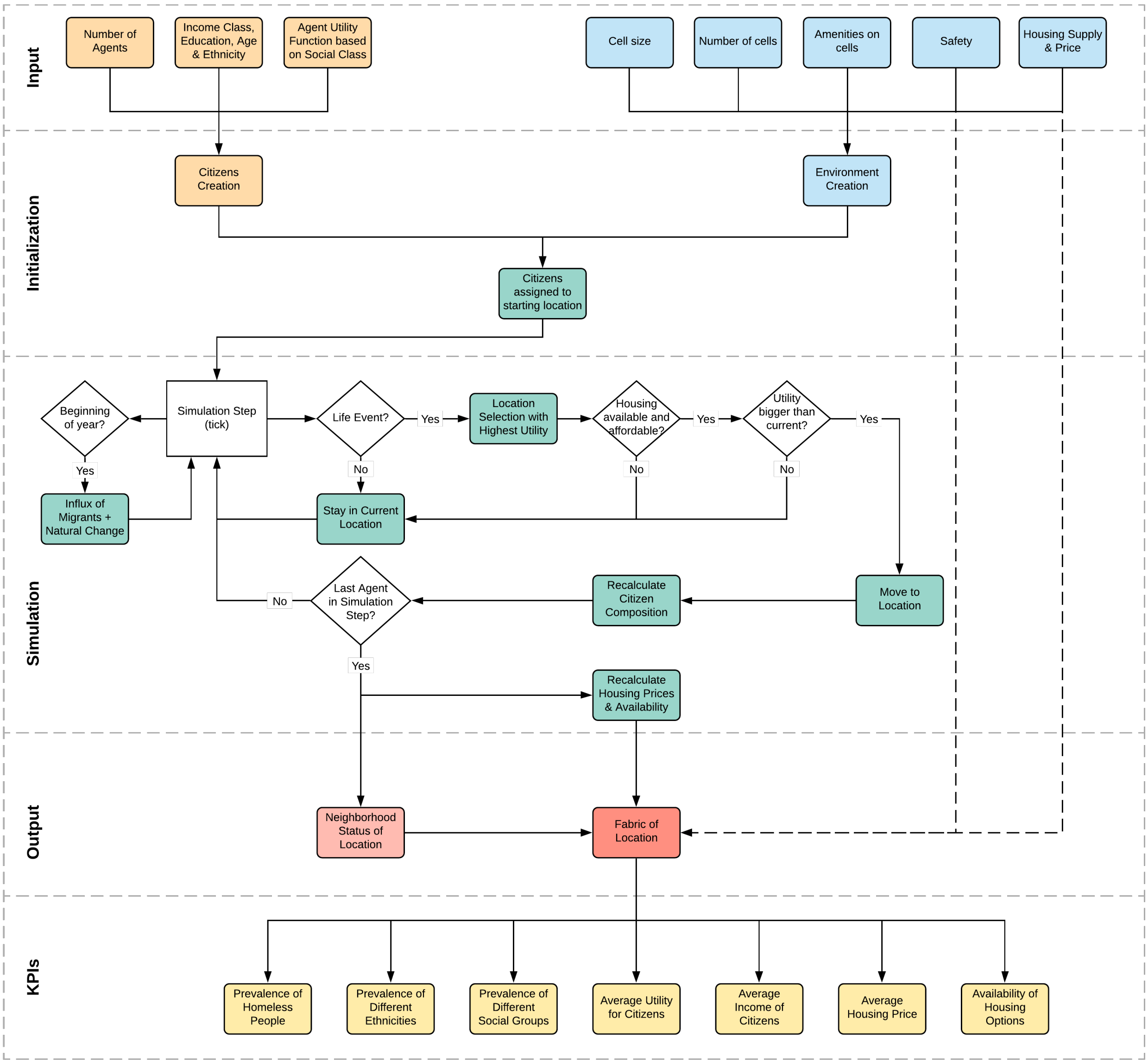


Figure H.1: Procedure of the model.

H.2.1 Input

By using data from CBS and Open Data Den Haag (of The Hague, 2020), the inputs for the citizen agents are formulated. Furthermore, the input for the environment can also be generated. The process of gathering data and preparing it for model use is described in the Data Chapter (Chapter 2).

H.2.2 Initialization

After the data has been prepared and entered as inputs into the model, the agents and their initial location are defined. First, the map and layout of the observed areas are loaded into the model. This is done using GIS or shape files. Then, the neighborhoods are initialized as agents of the same class or "breed", with fixed locations based on data. The neighborhoods retrieve their respective properties such as average housing price, population size, availability of housing and many more from the data that has been initialized in the Input phase.

Next, the citizens of each respective neighborhood are imported into the model. This is done by checking the properties of each neighborhood and using the properties to define the amount of citizens that are spawned into each neighborhood. By looking at the aggregation size parameter in conjunction with the population data, the amount of citizens is determined. The properties of citizens is then determined in either a randomly distributed fashion or based on correlation logic. The first manner checks the properties of the neighborhood and uses them as distributions for a random draw. For example, if a certain neighborhood has a population size of 100 households and 5% of its inhabitants are of Turkish ethnicity, the random distribution assumes that 100 citizens will be created, each have a 5% chance of being Turkish. This is then done for all the relevant demographics of citizens.

The second manner in which citizens can be created at the start of the model is by assuming correlations between certain properties. If this approach is used, the initialization of citizens correlates to properties. For example, if an agent is created with a high income, the model assumes there is a higher chance of this citizen owning a home instead of renting. Similarly, the model assumes citizens of lower income and education to have a small chance of owning a home, and rather will be renting for a place to stay.

After the initialization of the city map, neighborhoods and its citizens is completed, the model is then ready to perform the last step of the initialization phase, which is to add an initial influx of migrants. The model assumes that at the start of each year, a new influx of migrants enter the city. For consistency's sake, this is also the case in year 0 (at the end of initialization).

To finalize the initialization, all citizens (including migrants) recalculate the social group they should be part of based on their age, income and education. After this, the model is ready to start the simulation phase.

H.2.3 Simulation

The main logic loop is presented in the simulation phase. The most important aspects of the simulation loop are discussed in this section. However, a complete and detailed overview of all the procedures and mechanics within the simulation can be found in the model narrative Appendix (Chapter H).

As Figure 3.1.3 shows, at the start of the simulation loop (which is every simulation *tick*) the citizen agents are asked if they are currently deciding about moving to another location. As discussed in Section 3.1.2, it is assumed that agents only move when a life event occurs in their life. Using a logit implementation (Klabunde & Willekens, 2016), it is calculated for each citizen if a life event is currently happening, and the agent decides to move.

If this is the case, the citizen agent reviews all locations which have housing availability and checks if they are within its price range. Using this short list of locations, the agent then checks if the locations provide higher utility than its current housing location. If there are no locations that can provide higher utility then their current location, the loop ends and the citizen agent does not move.

If there are locations that can provide a higher utility than the current location, the locations that meet this requirement are compared. The location that can provide the highest utility for this citizen agent will become the chosen location for the agent to move towards.

Now the citizen agent moves from one location to another. As a result, the status of both the location the agent leaves as well as the location the agent moves in to are recalculated since there is a change in citizen demographics for those locations. If the change in citizen composition leads to a change in average social group or income of the location, the average housing price is recalculated as a result of housing market dynamics.

After completing the simulation loop for one agent, the same is done for all other citizen agents within the model. Because the chance of a life event occurring are relatively low, this means that during the simulation of a *tick* only a small part of agents move.

After all citizens have decided to move or not move, the simulation *tick* is concluded by advancing time with 1 *tick*, or 3 months of simulation time. During this phase, the percentage of social groups and ethnicity are recalculated for each location based on the changes from the simulation. This is done after the movements to minimize computations.

Once every year (or 4 simulation *ticks*), a new influx of migrants is simulated entering the city. Based on migration data, the amount of migrants is determined (Statline, 2020a).After all agents have finished looking for suitable locations to move to, the simulation recalculates housing pricing and income of locations based on the changes of citizen composition. It is assumed that a rise in average income in a location results in a rise in average housing price as a result of housing market dynamics (Gaube & Remesch, 2013).

H.2.4 Output

The output of the model is generated at the end of a simulation loop, at which the citizen composition for each location is recalculated. Furthermore, once every 4 *ticks*, the housing price and average of citizen statistics is recalculated.

Using the citizen composition and demographic statistics such as income and social group are used to then define the location's neighborhood status. The fabric of the current location can be calculated using the housing price and availability, the safety of the neighborhood and availability of amenities. This output can be later analyse to see if there are correlations to be observed between changes to the citizen composition and city fabric. This is done in the data analysis phase of research, highlighted in Chapter 4.

H.3 Decision-making Logic

In the flowchart, the different factors included in the decision-making of citizen agents are shown. The factors that are included in the decision-making process are internal and external to the properties of the agent. In other words, to align with the framework as defined in the conceptual framework (Section 3.1.2), the agent assesses properties from himself (citizens submodel) as well as its observed location's environment (environment submodel).

Based on a citizen's affordability of needs (social group), certain factors are either included or excluded in the decision-making process. This is highlighted in more detail in Section F. Based on their affordability of needs, the agent might just only check neighborhoods for people of the same ethnicity but if the affordability of needs allows for it, might also include a wide range of needs such as amenities, safety and social grouping needs.

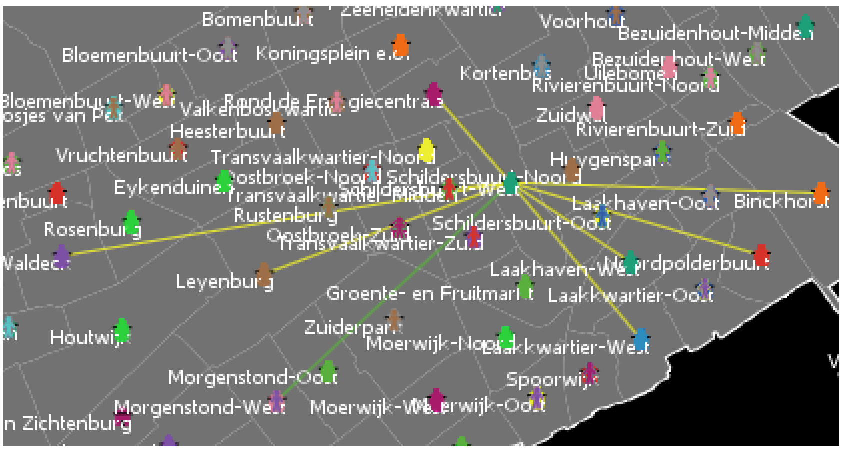


Figure H.2: A citizen agent looking at different neighborhoods and comparing his utility score for each of the neighborhoods. Green lines indicate a social rent housing option, yellow indicates private rent and red lines indicate houses for sale.

By checking neighborhoods for the availability of these factors, the utility of each neighborhood is calculated. This is done by multiplying the different factors (which are normalized) with a "significance factor", based on the relevance of each factor calculated based on survey work (Center, 2020). Then, each factor (with its multiplication) is then summed to make up for a total utility score. This is done for every potential location. Afterwards, all possible locations are compared and the location with the highest utility is chosen as the location to move to. Figure H.2 shows an agents comparing different neighborhoods on his utility scores, by making a temporary link with the neighborhoods it is comparing. In the case the highest possible utility is in the location the citizen already lives in, it is assumed the agent does not need to move at this moment and no movement occurs. An overview of the decision-making is shown in Figure H.3.

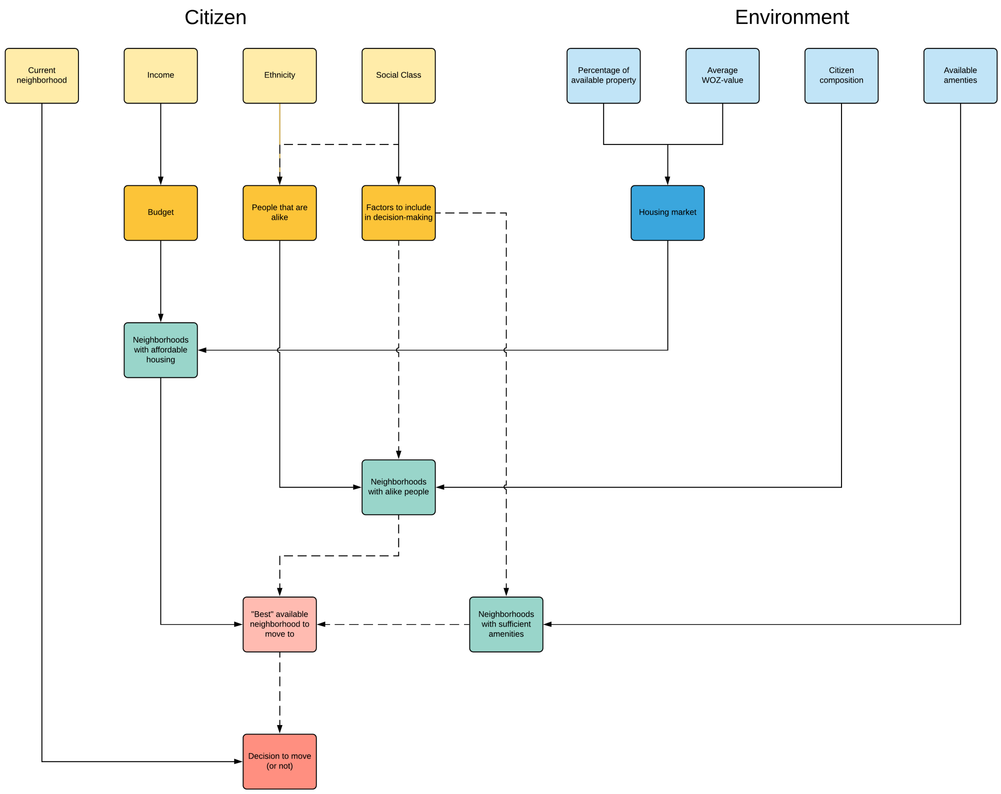


Figure H.3: Decision Logic for citizens deciding to consider moving.

H.4 Exogenous Factors

There are factors within the model which cannot be influenced by the decision-making of the municipality, but are important for the outcomes of the model. Most of these exogenous factors have some uncertainty in the value of parameters (in the future). Therefore, to cope with this uncertainty, the model has been tested for multiple values of these parameters, the results of which can be found in Chapter L. Table H.2 gives an overview of the tested parameter space, which is simulated during the running of the model.

Parameter	Variable	Metric	Values	Default Valu
P1	Standard deviation of income distribution	Multiple of average	0.00, 0.05, 0.10, 0.15	0.05
P2	Multiplier of expected influx of migrants	Multiplier expected influx	0.5, 1.0, 1.5, 2.0	1
P3	Average spendable income of migrant	Thousands of Euros / year	15, 17, 19, 21	19
P4	Randomly-assigned attributes	Boolean	yes, no	no
P5	Housing market and income inflation	Boolean	yes, no	no
P6	Size of aggregation of citizen agents	# Households / agent	1, 5, 10, 100	10

Table H.2: Parameter space of experimental parameters for testing sensitivity of agent properties.

Appendix I

Model Verification

To make sure the model that has been made "does the right thing", verification is needed (Nikolic et al., 2019). A framework outlines 4 steps involved in the verification process of Agent-Based Models (Nikolic, 2019). Using this framework, the steps involved are as follows:

1. Recording and tracking agent behavior
2. Single-agent testing
3. Interaction testing in a minimal model
4. Multi-agent testing

In essence, the verification tests and checks if the model behaves as intended in its most basic, aggregated version. This is in line with the design philosophy of "keeping it simple, stupid" (Edmonds et al., 2019; Edmonds & Moss, 2005). Furthermore, literature points out that verification of ABMs with a focus on social aspects should focus on the most simple realistic behavior that can be explained or simulated in a model (Ormerod & Rosewell, 2006). By testing the behavior of parts model in isolation, the observed interactions can be verified.

I.0.1 Recording and tracking agent behavior

By using output variables as a monitor, the values and changes to values in the model can be observed. This is necessary to see how the model behaves over time, and if the model mechanisms work as intended.

An example of such a monitor is to count the amount of *links* currently present in the model at all times. The *links* are used to calculate the utility of a citizen agent with regards to one location. After calculating the utility for all possible locations which the agent might move to and deciding which link has the highest value, the links should be deleted again (to save memory). To check if the model behaves as intended, a counter was added that monitors the amount of links currently present in the model, as shown in Figure I.1.

Many more monitors are present in the model, such as date representations, amount of agents in the model (to check if the correct amount of agent have been created), total movement counter, graphs that plot an overview of distributions of demographics and more.

Another form of tracking that was used during the creation of the model is the use of *error catching*. Error catching or handling is a technique where an input is monitored for unexpected outputs, and warns if the model encounters such an event by logging the output to the *log*. An example is the *logging* of agents that were not able to find any locations at all, during their search for a new home. By defining the different stages of the location search process, it becomes easier to determine why the agent was not able to find a suitable location. If there are no locations present at all, the monitor would log an error describing something in the simulation must have gone wrong, whereas an agent unable to find a suitable home because his budget is too low is considered an outcome that should be considered possible within the simulation.

Other forms of *error logging* include checking if an agent correctly leaves his former home when moving, making sure an agent only moves to a new neighborhood if there are homes available for purchase/rent, checking if agents are correctly created and have all their properties set correctly, checking if there are no agents without

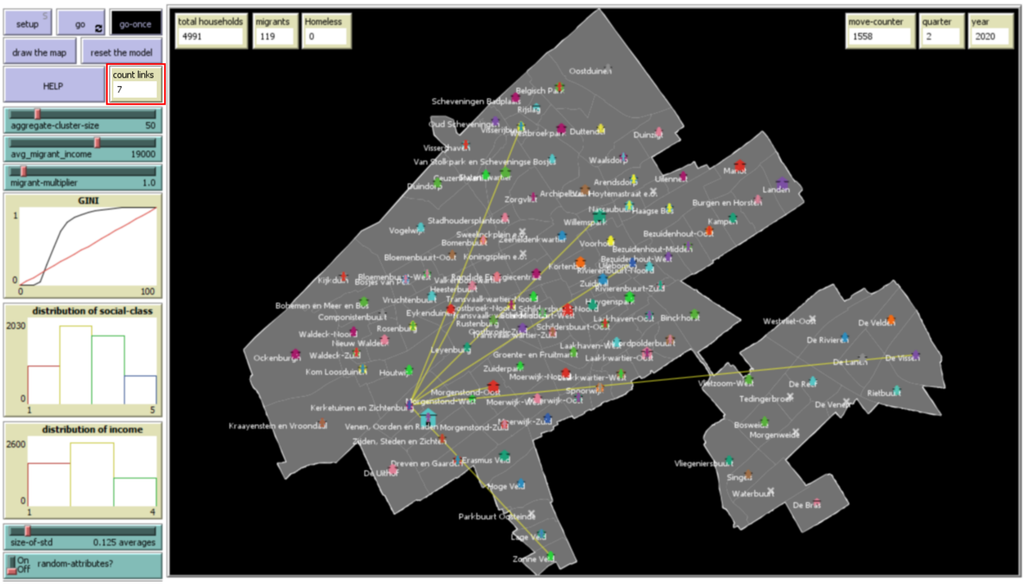


Figure I.1: A highlight of the link counter monitor. This box displays the current amount of *links* present in the model.

a budget or income, making sure neighborhoods with 0 population (non-residential areas) do not inherit properties such as average income and average housing price.

During the building of the model, an even more extensive form of logging was present to make sure behavior was implemented correctly. Because of the complexity of the decision-making process of agents looking for suitable locations to move to, the first iteration of the model would *log* every single interaction within this process to make sure the calculation of utility worked as intended, the calculation of the best option worked and lastly to make sure all possible options were considered (and no options that were not a possibility!).

With the help of tracking and *logging* the agent behavior, many bugs and unintended behavior was found and corrected. For example, in one iteration of the model a slight oversight in coding resulted in the agents not ever considering their budget when searching for suitable locations in decision-making. This was quickly discovered when the *link counter* monitor always showed the same amount of links, showing something was not working as intended.

I.0.2 Single-agent testing

To make sure the model works as intended, it can be helpful to reduce the model to a single agent. This means there are no interactions to be observed, and makes it easier to observe the behavior in its most basic form. Testing the model using a single agent is done using a two-fold test, first the model behavior is observed under normal conditions to see if the model with only one agent works as expected. For the first test to have added value, it is necessary to define a prediction of behavior before running the single-agent test.

After conducting and reporting the single-agent test under normal circumstances, the model is *stress-tested* to see what extreme values of model properties can *break* the model, which is to say observing the point at which the model stops working as intended or new unintended behavior starts to appear.

What follows in this section is a list of predictions of model behaviors followed by a either a confirmation that the prediction is correct, or an explanation of behavior that was unexpected.

Because the neighborhood agents are bound to imported GIS data, they are considered not part of the single agent. In other words, single-agent testing in this case means single citizen-agent testing. Furthermore, because the citizens are imported from data as well, the only way to force a single agent is to increase the *aggregate-cluster-size* value to extreme amount (higher then 6500 works to make sure only one citizen agent is created).

Setting up the model multiple times should change the properties of the agent, as it is an aggregate agent representing many agents at once. **Confirmed.**
Because of the insane aggregate-cluster-size, only the biggest neighborhood is able to accommodate a citizen (and thus the citizen will never move and always be spawned in the same location). **Confirmed.**
Although it can only be a tiny number, migrants can spawn at the beginning of a year. **Confirmed.**
Location checking still works, but no suitable location are found since the utility is already at its best for the initial agent. **Confirmed.**
Changing the properties of the only citizen agent in single-agent testing results in massive changes to average income, housing price and demographics. **Confirmed.**
When *aggregate-cluster-size* is higher than the highest household count of neighborhoods, no initial citizens spawn in setup. **Confirmed.**
When *aggregate-cluster-size* is higher than the highest amount of expected migrants to enter the model each year, no agents spawn at all. **Confirmed.**
Agents can still perform the check for locations even if no alternatives are available (in other words, the model does not crash). **Confirmed.**

I.0.3 Interaction testing in a minimal model

Interaction testing is done by making a minimal model with the least amount of agents present to see the bare minimum of interactions. Doing so makes it possible to confirm that the interactions work as intended as well as seeing the interactions in their most basic form.

Because of the way the model has been built, with a heavy reliance on the use of data, the previous step already partly outlines the minimal model testing. One exception to the minimal model testing is to include such an amount of citizen agents that movement in its most basic form is possible, which is at citizen count 2. This is achieved by setting the *aggregate-cluster-size* to a value of 6000.

Agents in minimal model testing are able to switch housing locations. **Confirmed.**
Movement of agents results in changes to demographics and housing availability. **Confirmed.**
Movement of agents ultimately results in changes to housing prices. **Confirmed.**
Agents cannot move into a neighborhood that is "fully occupied" by another agent. **Confirmed.**

I.0.4 Multi-agent testing

The last type of model testing that is performed during the verification is the multi-agent testing. In this phase, more agents are added to the model to make sure the model runs as intended. Then, using predictions, edge cases are sought after to see if the model breaks under certain (extreme) conditions. This last phase is guided by 4 types of testing: 1) Theoretical prediction tests and sanity checks 2) Break-the-agent tests 3) Variability tests 4) Timeline sanity tests (Nikolic, 2019).

Theoretical prediction tests and sanity checks

Increasing the amount of agents in the model by decreasing the *aggregate-cluster-size* does not change the percentages of demographics, housing supply or averages of income or housing prices. **Confirmed.**
Increasing the multiplier for migrant inflow results in a housing shortage, causing more agents to end up homeless. **Confirmed.**
Migrants with a higher income have no trouble finding a home, since they have a budget to do so. **Confirmed.**
Homeless agents always have a low income. **Confirmed.**
The standard deviation values lead to normal distributions of income that are similar to those observed in the real-word. **Confirmed.**
Agents cannot live in areas without homes. **Confirmed.**

Break-the-agent tests

Increasing the amount of agents to real-world size (1:1 ratio) causes the model to become extremely slow since there are over 260.000 agents present. **Confirmed.**
Setting the *aggregate-cluster-size* to 1 should not break the model, but setting it lower (0 or negative) should break the simulation. **Confirmed.**
Changing income of agents in one location drastically results in model-breaking inequality. **Partially Confirmed.** *Model still works but this location becomes inhabitable for anyone that does not live there yet.*

Variability tests

The goal of the variability tests is to explore the variability of the output in different regions of the parameter space. Because the best results for exploring the output space is done by testing a big range of values, this process is automated using the *Behaviorspace* tool in Netlogo. This tool enables the possibility of simulating the model using variations of values in an automated fashion. Furthermore, the results of these simulations are automatically gathered and summed in a single spreadsheet containing results of each individual run as well as summaries of all simulation runs.

A *Behaviorspace* experiment was run using the following settings:

```
["size-of-std" 0.05]
["avg_migrant_income" 19000]
["random-attributes?" false]
["housing-market-inflation?" false]
["aggregate-cluster-size" 10]
["migrant-multiplier" 1]
```

In short, all relevant levers were set to default measures except for *aggregate-cluster-size*, which is set to 20 to speed up the simulation time. It is expected that this has no influence on the model behavior as proven in the verification tests prior to this test. The repetition count is set to 160, as running the model for more than 100 cycles can prove the stability of the model (Nikolic, 2019). Furthermore, since the PC running the model has 16 cores (which can run in parallel), a multiple of 16 (ergo 160) was chosen.

Figure I.2 shows the results of 160 runs when observing the average housing price for each neighborhood. The model proves to be very stable, as most neighborhoods show little spread in their end results. Furthermore, looking at the outcomes in the figure, no unintended behavior can be observed in the outcome space. Therefore, the verification of the model can conclude no unintended behavior is observed.

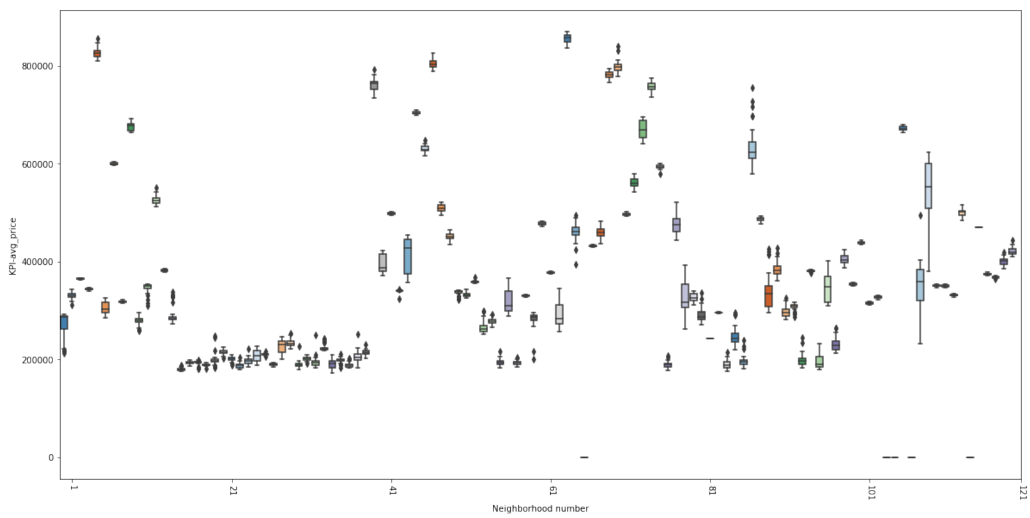


Figure I.2: Box plot graph of the normal verification run, showing average housing prices of neighborhoods from 160 "baseline" simulation runs.

As a last check, the model is run on a 1:1 aggregation size, meaning every agent in the model represents a household in the real-world city of The Hague. This means the model had over 260.000 agents, which significantly increases memory usage, calculations and run time. Because of this, this last verification step is only simulated 16 times (since the model can run 16 simulations in parallel on a 16-core CPU, an AMD Ryzen 3800X in this instance). The results of this run are shown in a map overview in Figure I.4.

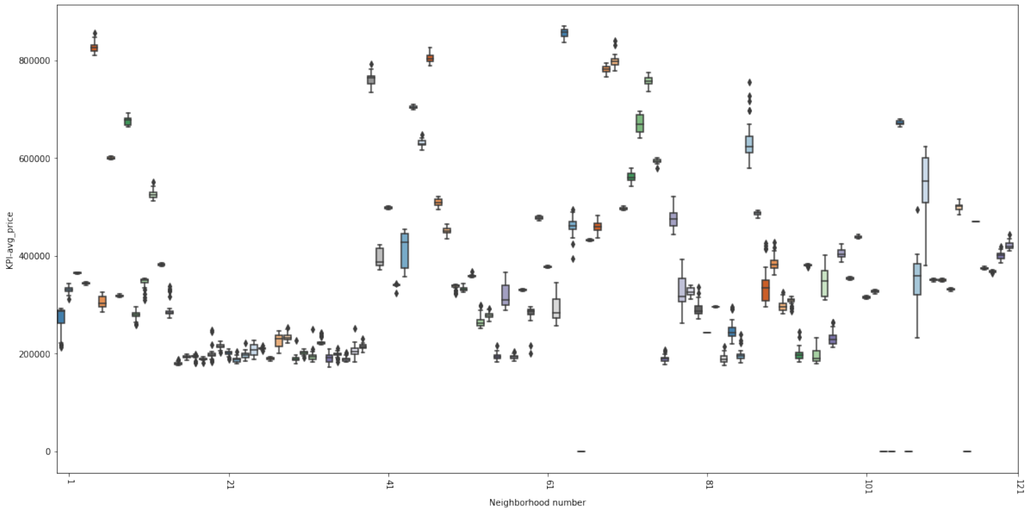


Figure I.3: Box plot graph of the third iteration of the normal verification run, showing average housing prices of neighborhoods from 16 "normal" simulation runs with an aggregation of 1:1.

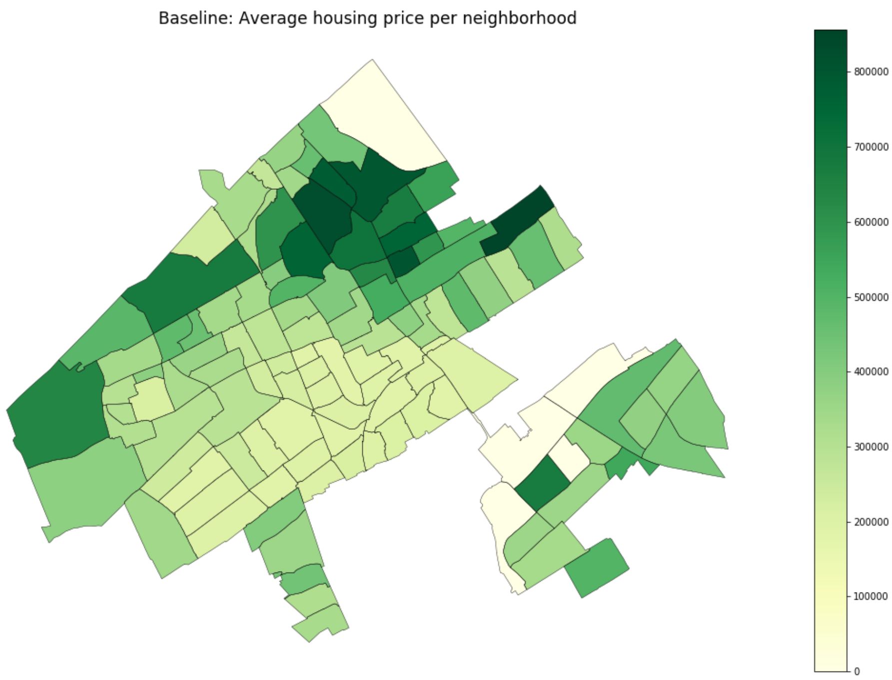


Figure I.4: Map plot overview of the normal verification run, showing average housing prices of neighborhoods from 16 "normal" simulation runs with an aggregation of 1:1.

Since this process is closely aligned to the exploration of outcomes of the model, some more details of the process are discussed in the results section of the research (Chapter 4). Any unexpected results encountered during the exploration of the results is reflected in that chapter as well, followed by an exploration of limitations and shortcomings of the model, discussed in Chapter 6.

Timeline sanity tests

Timeline sanity checks are performed to answer the question if the outputs can be explained by reasoning through the model logic. In other words, by looking at the results of the model run under normal circumstances, does the output make sense. Furthermore, the results of the model are checked to see if the outcomes can be explained by the behavior that was implemented in the model, or that some of the outcomes have unexpected components in them. The latter can either be caused by flaws or bugs in the model or be a product of emergent behavior (Li, Sim, & Low, 2006). Emergence is the phenomenon where new behavior is the result of interactions between agents and other agents or (parts of) the environment of the model (Dam et al., 2013).

During the timeline sanity tests, the matching of neighborhood properties has also been checked. It is important to check if all neighborhoods inherit the right properties, and the right data is produced on the right time. By doing the timeline sanity tests, the observed behavior verified the intended mechanisms of the model.

Appendix J

Model Simulation Appendix

This chapter highlights the results from the full model simulation runs which were performed using a distributed network of servers to run the model. This was done using the HPC server system (in specifics, the TPM cluster; TBM1-4) from the Delft University of Technology (of Technology, 2020).

First, the settings used to run the simulation are discussed, followed by an overview of the data that is the result of the simulations. A list of graphs and descriptive statistics is then introduced as a ways of indicating the data structures and general outcomes of the model.

J.1 Model Simulation Settings

Because of the size of the model, and the amount of agents that are present, the model takes a long time to simulate. On average, a computer can take up to 30-45 minutes to simulate one full simulation cycle of the model, depending on the processing power of the computer. Since we want to test multiple scenarios and policy levers and run the model frequently enough to be able to correct for stochasticity of the model, the model was simulated more than a thousand cycles.

To cope with this large number of runs, the simulation was run using a distributed network of computer servers, called the HPC (of Technology, 2020). This cluster of computers is specifically built to be able to compute models and calculations for the use of researchers and students of the University. With the help of Dr.ir. I. Nikolic, the model was imported to the HPC server and run twice. The first run ended in a problematic failure, and thus, after debugging the issues and fixing the model, the model was run again on a second occasion. The runs took over 20 hours to simulate on 4 separate server units consisting of 32-core systems, equipped with ample RAM, VRAM and storage space. For details on the specifications of the server, see the specs on the HPC website (of Technology, 2020).

To use the HPC servers most efficiently, the simulation runs were divided into 4 separate *BehaviorSpace* experiments in Netlogo. The advantage of this, is that the model could be run in parallel on 4 systems at the same time, the only downside is that the interference between certain policy levers is not tested since the runs were performed on separate machines. To run the model on the servers, a Linux command line *one-liner* was written to run the *BehaviorSpace* experiments using the *Headless* version of Netlogo (Wilensky, 1999).

The repetition count for each of the experiments was set to 32. This number is high enough to be able to account for the stochastic properties of the model, and get a good sense of the behavior of the model. Furthermore, the number 32 was chosen because all of the servers would run the model in parallel mode on all 32 cores within each system. This makes it possible to run 32 model at the same time, hence it is wise to run the model in a multiple of said number. In total, this equated to 1056 runs. Below is a detailed list of settings of the experiments for each of the 4 servers.

J.1.1 Unexpected Outcomes

After simulating the model using the HPC servers, analysis of the data started. This chapter outlines the analysis of the data to get a sense of its dimension and the way variables are stored. However, during the analysis of this data, a major flaw was found in the outcomes of data. The distribution of incomes was far greater than expected, ranging from 0 to extremely high numbers. After thorough investigation, the culprit was found to be a wrong value for the standard deviation of income distribution of all citizen incomes in the model. Unfortunately, because this value was used all-throughout the model, all outcomes from the HPC simulation were rendered invaled and therefore useless.

J.1.2 New Simulation Runs

Because the runs were considered invalid, new simulation runs have to be performed. Some other changes were made to the simulation process as well, such as a new approach for testing the experiments. Furthermore, some more minor issues which were found in the model were addressed, and the new version of the model was used for simulation. Because of a limiting time constraints and no ability to run the new version of the model on the HPC on short notice, the model was simulated on the Author's own computer, running on an *AMD Ryzen 7 3800X* 8-core, 16-threads CPU running at 4,3GHz. Although faster than the CPUs of the HPC, the limiting factor for running on this computer is the core count. Only 16 runs could be simulated in parallel, limiting output. Because of these limitation, the model was chosen to be run in an aggregation size of 10, which means every agent in the model represents 10 households. The impact of this change has been described in detail in the sensitivity analysis (Chapter L).

J.2 Simulation Run Parameters

Running the simulation requires strategic thinking on what to simulate, and when. There are many parameters in the model that can be adjusted and policy levers to be turned of or on, which leads to many possibilities of running the simulation. To do this efficiently with the time that is available, while at the same time preventing interference in the outcome space between parameters, a single-parameter simulation approach was chosen. In short, this means all tested parameters were simulated in isolation. In total, this leads to several dozens of isolated simulation runs, which afterwards can be combined for data analysis. Table J.1 sums the different simulation runs and their parameter values, and Table J.2 shows are runs with regards to the testing of policy levers.

Parameter	Variable	Metric	Values	Default Valu
P1	Standard deviation of income distribution	Multiple of average	0.00, 0.05, 0.10, 0.15	0.05
P2	Multiplier of expected influx of migrants	Multiplier expected influx	0.5, 1.0, 1.5, 2.0	1
P3	Average spendable income of migrant	Thousands of Euros / year	15, 17, 19, 21	19
P4	Randomly-assigned attributes	Boolean	yes, no	no
P5	Housing market and income inflation	Boolean	yes, no	no
P6	Size of aggregation of citizen agents	# Households / agent	1, 5, 10, 100	10

Table J.1: Parameter space of experimental parameters for testing sensitivity of agent properties.

J.3 Simulation Descriptive Statistics

An overview is made of the statistics for each of the servers that were used to run the model simulations. The descriptive statistics and overview plots of each server are shown.

J.3.1 Full Data Descriptive Statistics

Lever	Name	Description
L1	transform-houses?	By the municipality buying properties, the 5 neighborhoods with the least amount of rental properties transform 50% of houses for sale into rental properties (both private and social sector).
L2	improve-health?	By building more healthcare facilities, the 5 neighborhoods with the worst healthcare get a 10% increase in the availability of healthcare facilities.
L3	increase-social-housing?	At the beginning of the model simulation, 10% of all rental properties in the private sector are bought and transformed to social rent housing options.
L4	build-more-houses?	By funding projects and granting construction permissions, the 5 neighborhoods with the least amount of housing availability will increase the availability in all sectors by 2,5%.
L5	mixed-use-zoning?	By allowing mixed-use of commercial or industrial zones, new housing is created in non-residential areas. This equates to 400 new housing options per year in each mixed-use zone.
L6	improve-safety?	By increasing the budget of the police and social workers, the amount of crimes in the worst 5 neighborhoods drops with 0-15%. Because of uncertainty, this drop is not a fixed value but randomly drawn.

Table J.2: Table showing all policy levers in the model simulation. Values of parameters in the simulation are fixed but tested for different values in the sensitivity analysis.

	mean	std	min	50%	max
run number	16.500000	9.233124	1.00	16.500	32.000
transform-houses?	1.000000	0.000000	1.00	1.000	1.000
step	20.500000	11.543436	1.00	20.500	40.000
KPI-homeless	8.075000	118.435808	0.00	2.000	4148.000
KPI-homeless-migrants	4.198437	98.867838	0.00	0.000	3460.000
Neighborhood number	61.991228	35.563876	1.00	64.500	121.000
KPI-available_buy_houses	230.459313	454.430774	0.00	45.000	2575.000
KPI-available_part_rent_houses	130.590865	319.884342	0.00	4.000	2427.000
KPI-available_social_rent_houses	122.602728	232.337900	0.00	10.000	1476.000
KPI-avg_income	37383.963980	19307.884452	0.00	34841.000	93736.000
KPI-avg_price	341700.933470	180886.202757	0.00	317231.500	875656.000
KPI-avg_utility	0.238033	0.668811	-7.21	0.269	1.804
KPI-citizen-count	262.507237	306.029873	0.00	164.000	2356.000
KPI-p-sc-lower	0.145481	0.263783	0.00	0.000	1.000
KPI-p-sc-working	0.299365	0.246015	0.00	0.274	1.000
KPI-p-sc-middle	0.273102	0.259343	0.00	0.214	1.000
KPI-p-sc-upper	0.215459	0.314672	0.00	0.000	1.000
KPI-p-dutch	0.433902	0.281009	0.00	0.527	1.000
KPI-p-other-western	0.106797	0.118305	0.00	0.083	1.000
KPI-p-antilles	0.025772	0.086141	0.00	0.010	0.990
KPI-p-morocco	0.038657	0.116839	0.00	0.012	1.000
KPI-p-suriname	0.032320	0.031075	0.00	0.028	0.429
KPI-p-turkey	0.064241	0.082659	0.00	0.036	0.912
KPI-p-indonesian	0.053062	0.110295	0.00	0.014	0.916
KPI-p-eastern-eu	0.064910	0.149816	0.00	0.027	0.980
KPI-p-other-nonwestern	0.113724	0.164224	0.00	0.082	1.000

Appendix K

Key Performance Indicators

This chapter describes the Key Performance Indicators (KPIs) of the model, which are used to quantify the performance of the simulation. A complete list of KPIs is shown in Table K.1. The details of each KPI and their respective measuring procedure are explained below.

KPI name	Variable	Measurement level
KPI-homeless	Total homeless people	Global
KPI-homeless-migrants	Total homeless migrants	Global
KPI-available_buy_houses	Available housing to buy	Neighborhood
KPI-available_part_rent_houses	Available private sector rent housing	Neighborhood
KPI-available_social_rent_houses	Available social rent housing	Neighborhood
KPI-avg_income	Average yearly spendable income of households	Neighborhood
KPI-avg_price	Average value of housing	Neighborhood
KPI-avg_utility	Average utility of citizens (happiness of living)	Neighborhood
KPI-citizen-count	Amount of households living in neighborhood	Neighborhood
KPI-p-sc-lower	Percentage of citizens in "lower social group"	Neighborhood
KPI-p-sc-working	Percentage of citizens in "working social group"	Neighborhood
KPI-p-sc-middle	Percentage of citizens in "middle social group"	Neighborhood
KPI-p-sc-upper	Percentage of citizens in "upper social group"	Neighborhood
KPI-p-dutch	Percentage of citizens of Dutch ethnicity	Neighborhood
KPI-p-other-western	Percentage of citizens of other Western ethnicities	Neighborhood
KPI-p-antilles	Percentage of citizens of Antillean ethnicity	Neighborhood
KPI-p-morocco	Percentage of citizens of Moroccan ethnicity	Neighborhood
KPI-p-suriname	Percentage of citizens of Surinamese ethnicity	Neighborhood
KPI-p-turkey	Percentage of citizens of Turkish ethnicity	Neighborhood
KPI-p-indonesian	Percentage of citizens of Indonesian ethnicity	Neighborhood
KPI-p-eastern-eu	Percentage of citizens of Eastern European ethnicity	Neighborhood
KPI-p-other-nonwestern	Percentage of citizens of other Non-Western ethnicities	Neighborhood

Table K.1: Table outlining all Key Performance Indicators in the Agent-Based Model.

K.1 Homelessness

Both *KPI-homeless* as well as *KPI-homeless-migrants* are calculated at the end of each *tick* using the *update-KPI* procedure which is run by the *observer*. It simply tallies the agents that are homeless, and the agents which are homeless and have the *migrant? = True* property respectively. The result is a number stored as an integer for every step of the model.

K.2 Measuring on Neighborhood Level

Apart from homelessness, all other KPIs are measured on the neighborhood level. This means that each neighborhood reports their neighborhood code along with the value of the KPI for its neighborhood and adds it to a list of number as an output. The resulting dataframe might contain an output cell formatted as following:

```
[[1 123][15 412][34 365]...]
```

Where the first number denotes the neighborhood's unique identifier and the second value show the value for the KPI that has been reported for this time step.

K.3 Available Housing Options

There are three KPIs that check the availability of housing options: *KPI-available_buy_houses*, *KPI-available_part_rent_hou* and *KPI-available_social_rent_houses*. Each of these KPIs measures the amount of houses that are currently not occupied by citizens for each of the 114 neighborhoods. The resulting output is a list of counts for each respective neighborhood. The distinction between private sector rent and social rent is made to show a difference in rent options, which are assumed to correlate to housing price. Social rent is assumed to be cheaper than private sector rent options.

K.4 Average Income & Housing Value

The average income of all citizens in a neighborhood on a certain time step is measured using *KPI-avg_income*, which is rounded down to an integer. Furthermore, the average price of housing options (all properties are included for the neighborhood) is listed in *KPI-avg_price*, which is rounded down as well to an integer.

K.5 Average Utility

The average utility, or in other words, the average "fit" of citizens in a neighborhood is calculated using the *KPI-avg_utility*. This KPI gives a good indication of how happy citizens are on average to live in a certain neighborhood. Because the utility is calculated individually for all the citizens in a neighborhood, and the calculation values can differ based on the individual's properties, it might not always reflect how well a neighborhood is performing. It does however show a good indication if the citizens living there feel happy to live there.

K.6 Citizen Count

The tally for total citizens living in each neighborhood is shown using *KPI-citizen-count*. It is interesting to keep track of the amount of citizens living in an area to get a better idea of which areas are attracting or deterring people from living there. However, more detailed information on the citizens living in a neighborhood can be found by looking at the ethnicity and social group prevalence within neighborhoods.

K.7 Social Group Prevalence

There are 4 KPIs which look at the respective social groups present in neighborhoods. These correspond with the theory as explain in Chapter F. The KPI reports the percentage of each of the four groups respectively for each neighborhood in The Hague. This can be used to estimate the type of residents living in a certain area, as well as observe changes to the people living in a certain neighborhood.

K.8 Ethnicity Prevalence

Lastly, to measure the prevalence of ethnicities within neighborhoods, nine KPIs show the percentage prevalence of each of the ethnicities observed in neighborhoods. The reason for categorizing into these 9 ethnicity groups is based on the available data of The Hague citizens from real-world data sources. Using the ethnicity KPIs can help understand moving behavior such as the clustering of citizens of the same (or similar) backgrounds.

Appendix L

Sensitivity Analysis

The sensitivity analysis is needed to check the influence of both internal (endogenous) and external (exogenous) influences on the behavior of the model. By defining factors that can be considered exogenous and then parameterizing these factors into variables, it becomes possible to test the influence of each of these factors by means of a sensitivity analysis. Furthermore, the exploration of interesting behavior in the model that is the results of endogenous factors can be explored by checking a broader range of values for variables that are in the model.

The values chosen for the sensitivity analysis are based on the default value which is used in the simulation of the ABM when observing the interactions between the model and the policy levers. Using the default value, values are chosen which look at both sides of the spectrum of values, for instance, a the sensitivity of a default value of 1 will be alternated to 0.5, 1, 1.5 and 2 to see how the model behaves. To prevent interference of variables in the output space, one parameter is tested at the time, thus creating 6 different simulation runs (Saltelli & Annoni, 2010).

All sensitivity runs are tested using an aggregation of households with a factor of 10. This means that one agent in the model represents an average of 10 households with similar or the exact same properties in The Hague. This is done by using the data sources available on the current status of the real-world, and aggregating the data. For instance, when the data show 1% of people in a certain neighborhood with 5000 citizens has a Turkish ethnicity, this equates to 50 citizens in the real-world, but because of the aggregation of the model, the model simulates 5 agents of Turkish ethnicity. To make sure the impact of this aggregation is also tested, the last sensitivity analysis checks if there are any significant differences in outcome when aggregation sizes are varied.

L.1 Exogenous Sensitivity

The parameters that are regarded within the model are defined during the conceptualization of the model (Section 3.1.2). Table L.1 overviews a detailed list of exogenous parameters that are explored and the values which are tested. All parameters are tested in isolation, to prevent interference between parameters (Ten Broeke, Van Voorn, & Ligtenberg, 2016). Furthermore, all metrics are observed for the averages at the last time step in the simulation, to see how the parameter has influenced the outcomes. A more in-depth look at the changes over time is discussed in Section L.1.7

Parameter	Variable	Metric	Values	Norm
P1	Standard deviation of income distribution	Multiple of average	0.00, 0.05, 0.10, 0.15	0.05
P2	Multiplier of expected influx of migrants	Multiplier of expected influx	0.5, 1.0, 1.5, 2.0	1
P3	Average spendable income of migrant	Thousands of Euros / year	15, 17, 19, 21	19
P4	Randomly-assigned attributes	Boolean	yes, no	no
P5	Housing market and income inflation	Boolean	yes, no	no
P6	Aggregation size of citizen agent	Amount of households represented	1, 5, 10, 100	10

Table L.1: Table showing the parameterization of experimental parameters for the sensitivity analysis of exogenous variables.

As Table L.1 shows, there are 6 different parameters that are varied in value to check the sensitivity to the observed behavior in the model. The first three variables are varied in value using 4 different value steps, to get an indication of the impact of these exogenous factors. The next two variables are mechanisms that can either be turned off or on in the model, and are thus checked to see how they influence the model behavior. Lastly, the analysis is checked by making sure the aggregation size does not influence the model.

L.1.1 Variation in Standard Deviation of Income Distribution

The model assumes a normal distribution (Figure L.1) for income within neighborhoods, where the mean of the distribution is defined using data from real-world observations (see also Chapter G). However, the "spread" or distribution of income is does not only determine on the mean value μ, but also relies on the standard deviation or σ. By looking at income distribution data from the Netherlands, three plausible values for σ are chosen, and as a control a σ of 0 is added. Since there are 114 neighborhoods each with their own respective income average, the values for σ are multipliers for the average. For example, a neighborhood with an average income of 30.000 Euros where the Standard Deviation Multiplier is set to 0.1 (the default for the model), leads to an effective σ value of 3000 Euros.

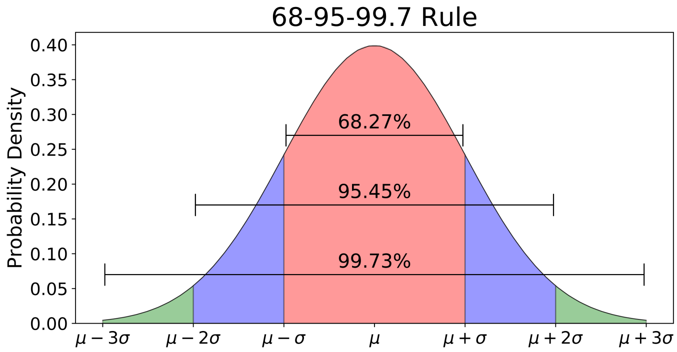
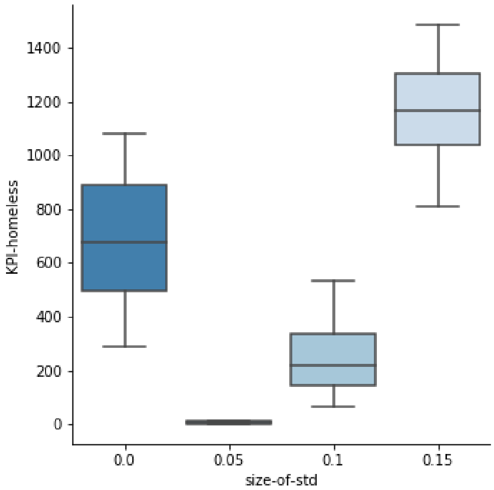


Figure L.1: Mathematical representation of a normal distribution.

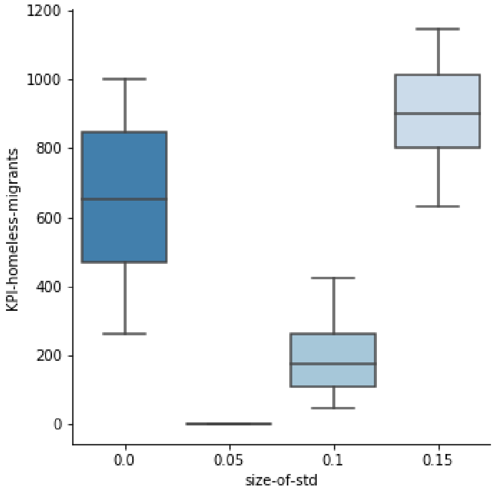
In short, increasing the value for the standard deviation multiplier simulates bigger inequality between citizens in a neighborhood. The most significant outcomes of changing the value of the multiplier are shown in Figure L.2.

As Figure L.2a and L.2b show, almost all homelessness in the model can be explained by migrants not being able to find a suitable living space. Furthermore, there seems to be an interesting correlation between the distribution of average income and homelessness. When all citizens within the same neighborhood have equal income (*size-of-std = 0*), migrants struggle to find a home. This can be explained by the fact that the income of migrants is also equal for all, and since it is lower than most citizens of The Hague, they cannot compete on housing options. However, when a slight variance in income distribution is present, the migrants seem to have little issues finding a home. At this moment, the author does not have a clear explanation for this observed outcome. When the variance becomes bigger, the increase in inequality between incomes leads to more homelessness of migrants.

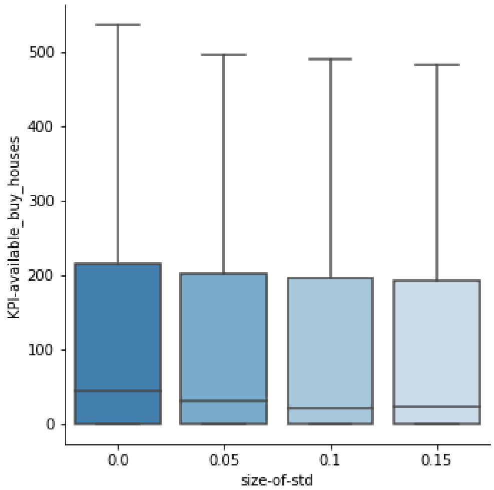
When looking at housing availability and average income (Figure L.2c and L.2d), it becomes clear that bigger inequality in income distribution leads to more rich people being able to afford to buy a house, lowering the availability whilst more poor people cannot find a home at all, which shifts the average income upwards. Similarly (but not shown in these graphs), a minor increase in social rent housing availability can be observed since the poor can no longer afford it and more people can afford more luxurious housing options on the other end of the income spectrum.



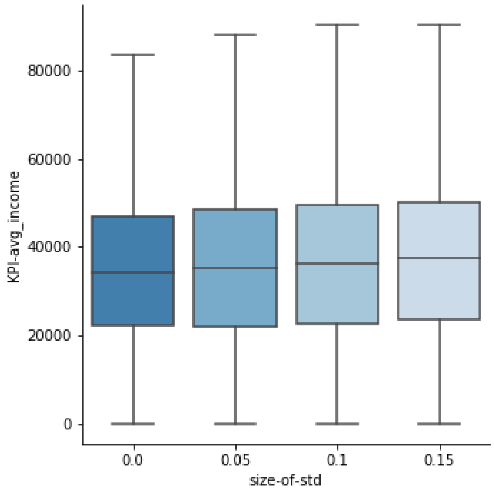
(a) Total amount of homeless people



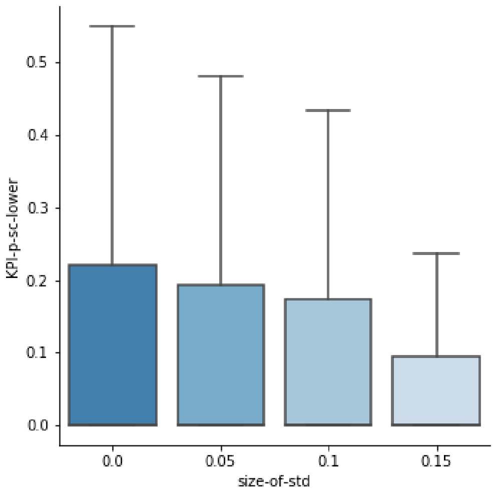
(b) Amount of homeless migrants



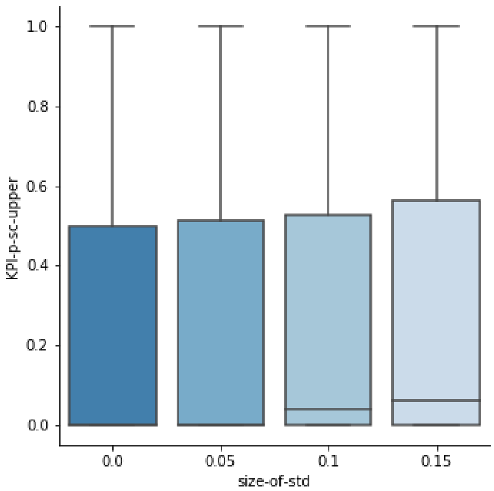
(c) Amount of houses available to buy



(d) Average income of neighborhoods



(e) Percentage of citizens of the lower social group



(f) Percentage of citizens of the upper social group

Figure L.2: Boxplots showing significant differences in outcomes of KPIs when testing the sensitivity of the standard deviation of income.

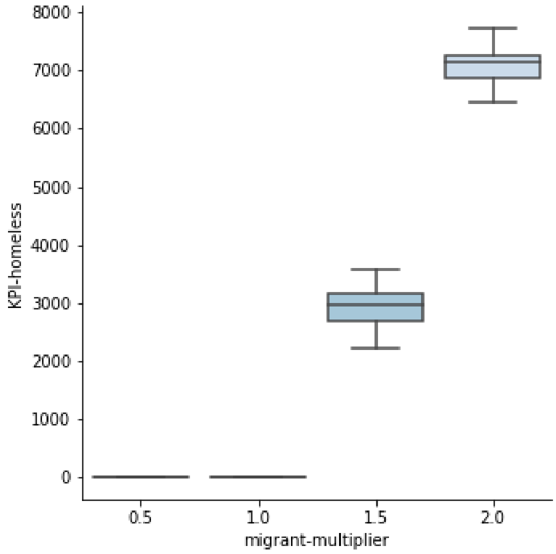
Figure L.2e and L.2f show that an increase in standard deviation of income leads to more people in the upper social groups, and less people in the lower social group. This can be explained that there are more extremes, on both the poor and rich side of the income spectrum. The decrease in lower social group prevalence is because many of these citizens have now become homeless, and thus are not accounted for in neighborhood averages.

A side note that should be mentioned is that during testing and verification of the model, a severe bug was found when miscalculating the values for the standard deviation. When using a value of 0.5, the income distribution is so wide, that incomes can reach negative values and very extreme heights on the other end of the distribution, which in term lead to unintended behavior of the model. More information on this can be found in Section 3.4.

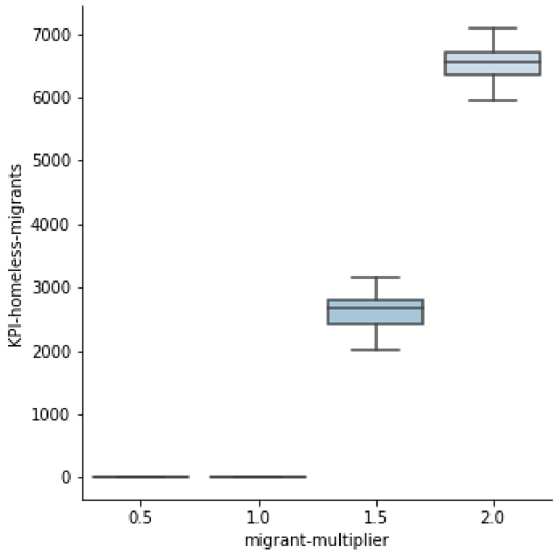
L.1.2 Multiplier of Influx of Migrants

When modeling the changes of the city fabric as a result of the influx of migrants, the influx itself is a very important aspect of the model. However, as the city cannot (directly) influence the influx of migrants (see also Chapter G) and there is much uncertainty in the expected influx, this variable has also been tested for its sensitivity. There is granular data on past migration patterns which is divided into country of origin. Using this data, the expected influx of migrant from non-western countries is predicted using linear regression (see also Chapter 2). To test the sensitivity of the influx, the multiplier value can adjust the expected influx by a multiplicative factor.

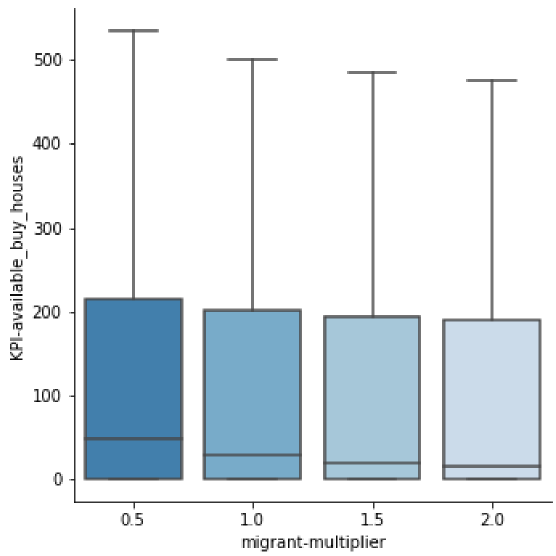
Figure L.3 shows the significant changes to model given differences in the value for the migrant influx multiplier. As expected, an increase in migrants arriving in the city leads to an increase in homeless migrants, since there are not enough houses to accommodate all migrants (Figure L.3a and L.3b). As the amount of migrants entering the city increases, the supply of housing options decreases (as shown in Figure L.3c, but can be observed for all three sectors). When more migrants enter the city, it also influences the average composition of neighborhoods. As there are proportionally more migrants from a lower social group, the prevalence of this group increases when more migrants enter the city (Figure L.3d).



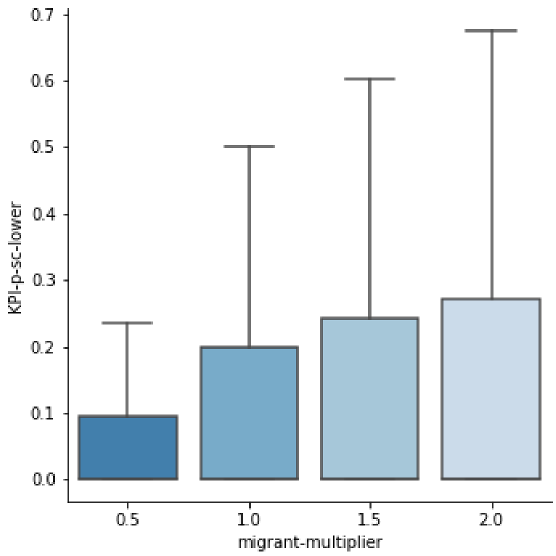
(a) Total amount of homeless people



(b) Amount of homeless migrants



(c) Available houses for sale

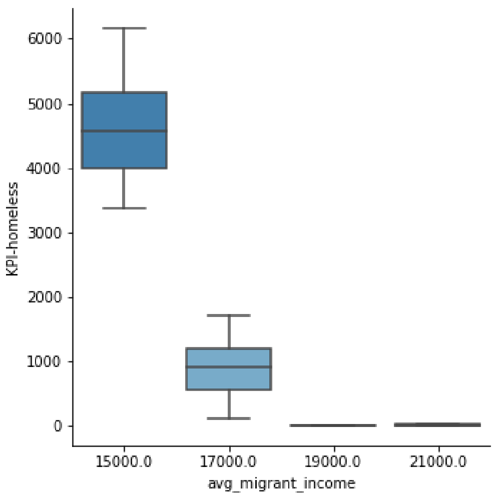


(d) Percentage prevalence of citizens in lower social group

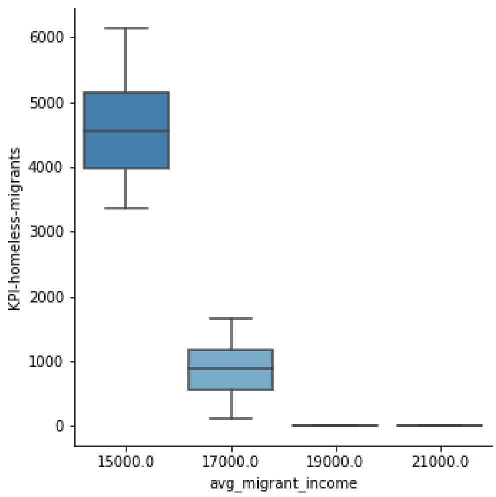
Figure L.3: Boxplots showing significant differences in outcomes of KPIs when testing the sensitivity of the multiplier for expected migrant influx.

L.1.3 Average Spendable Income of Migrants

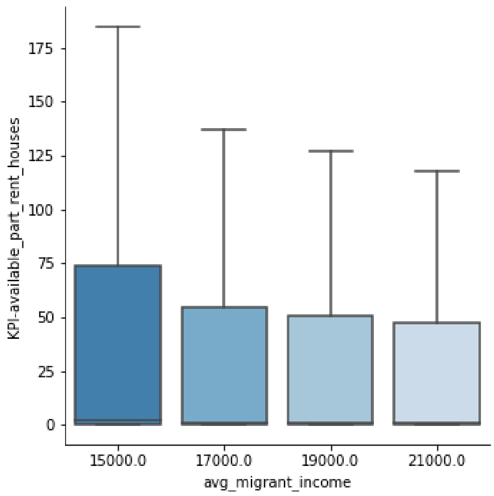
Unfortunately, there is little data available on the distribution of income of migrants arriving in the city. Because of this, assumptions have to be made with regards to the distribution. There are statistics on distribution averages in relation to country of origin which is used as an indicator for the default value for income of migrants. This default is set to 19.000 Euros. To test the sensitivity of income, three more values are tested. The resulting outcomes are shown in Figure L.4.



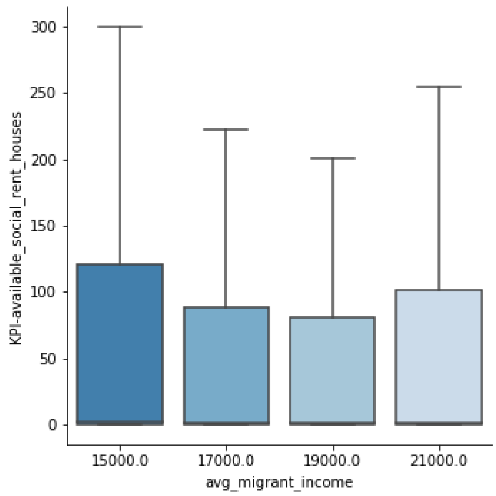
(a) Total amount of homeless people



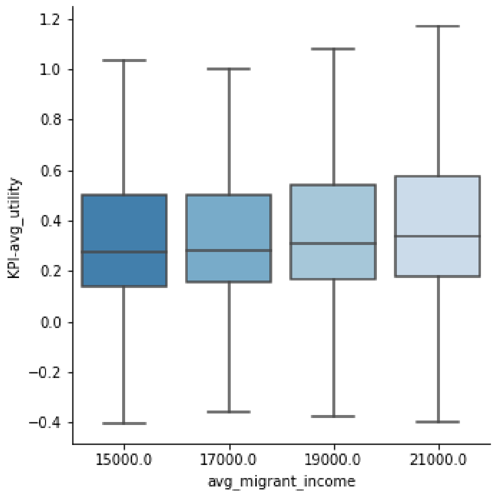
(b) Amount of homeless migrants



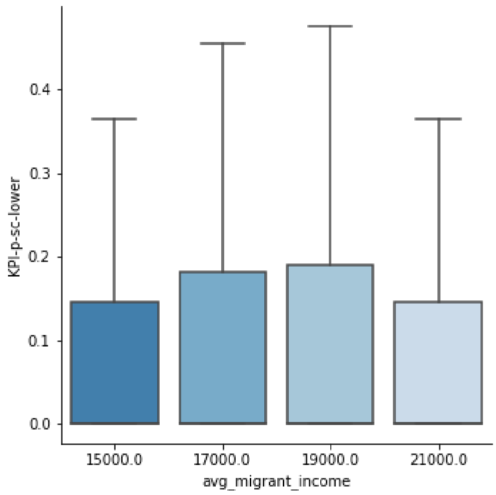
(c) Availability of houses for rent in the private sector



(d) Availability of houses for rent in the social sector



(e) Neighborhood average utility of citizens



(f) Percentage prevalence of citizens from lower social group

Figure L.4: Boxplots showing significant differences in outcomes of KPIs when testing the sensitivity of the multiplier for average income of arriving migrants.

As Figure L.4b shows, there seems to be a break-off point for migrant homelessness. When the average income of migrants is higher than 17.000 Euros per year, migrants are almost always able to find a home and are not forced to live on the streets. Furthermore, an increase in average income directly correlates to a decrease in availability of rental options in the private sector (Figure L.4c), which is due to the affordability of migrants to rent more expensive houses. Similar outcomes can be observed when looking at the social rental housing options (Figure L.4d), but when average income surpasses 19.000 Euros, migrants start looking more into private rent instead of social rent options.

Another interesting observation is the increase of average utility (Figure L.4e). When migrants have more income to spend, they can afford more suitable (or nicer) housing options which increases the overall utility of all citizens.

Lastly, an increase in average income among migrants leads to a decrease in the prevalence of citizens in the lower social group (Figure L.4f). This can be addressed to the calculation of social groups, where higher income groups tend to group in working, middle or upper social groups. Thus, less migrants with low income leads to less people in the lower social group.

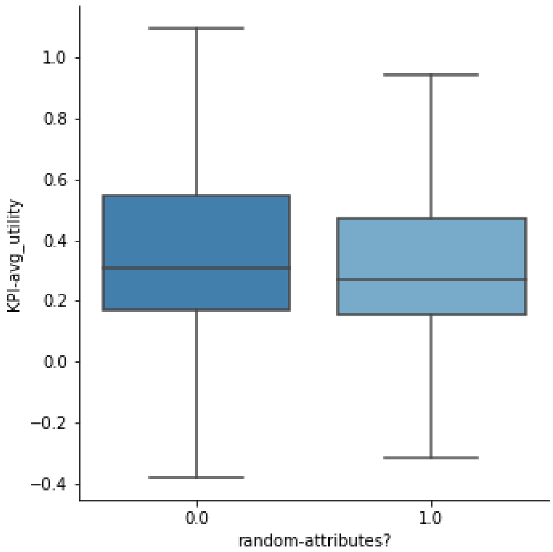
L.1.4 Randomized Attributes of Citizens

In the model, some properties of citizens are defined based on other properties of the agent such as an assumed relation between income and education (see also Chapter H and Chapter F). Because this relies on assumptions of citizen composition (see also Chapter G), the sensitivity of these assumptions is tested. The way the model simulates agent properties (dependently or randomly) can only be "switched on" or "switched off", and thus the experimentation space for this parameter is a Boolean. Since there are a lot of differences observed from changing this parameter, the plots showing the differences in outcomes are drawn in both Figure L.5 and Figure L.6.

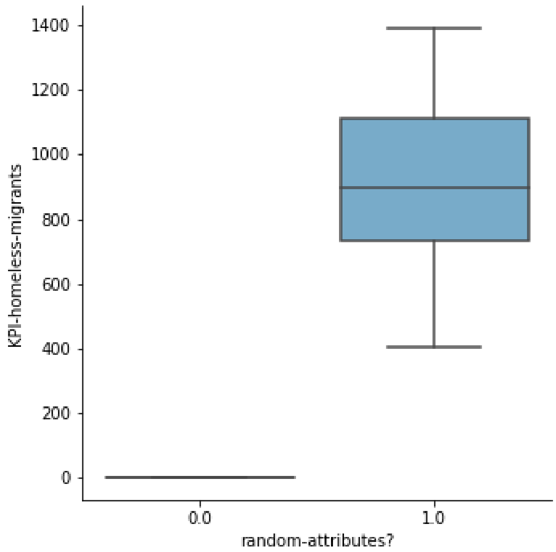
The average utility of citizens seems to be negatively impacted by the random distribution of properties. When properties are not correlated, citizens have a harder time adjusting their needs to their budget. For example, in the real-world, there is a correlation between income and education. Educated people tend to have more needs (which they can afford with a higher income). However, as Figure L.5a shows, if the attributes are assigned randomly, needs are met less often.

Looking at Figure L.5b, a positive trend can be observed when attributes are randomized in the case of homelessness. When agent properties are randomized, more "unfortunate" combinations of properties might occur.

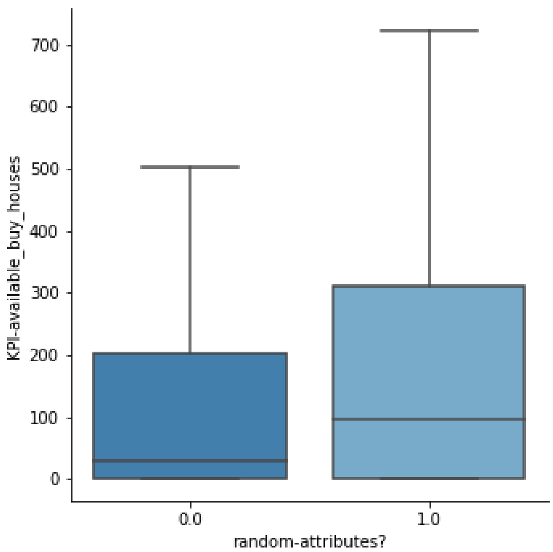
Furthermore, when properties are assigned randomly, more houses that are for sale are left available (Figure L.5c) whilst the amount of available rent houses decreases (Figure L.5d). This can be explained due to the fact that higher educated citizens tend to look for houses to buy, but because of the random assignment of properties do not always have an income to support this wish. At the same time, because of this unfortunate selection of properties, more citizens tend to resort to rental housing options as an alternative.



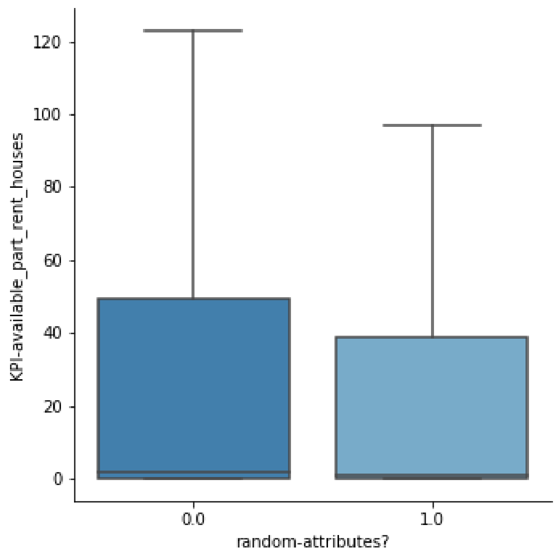
(a) Neighborhood average utility of citizens



(b) Amount of homeless migrants

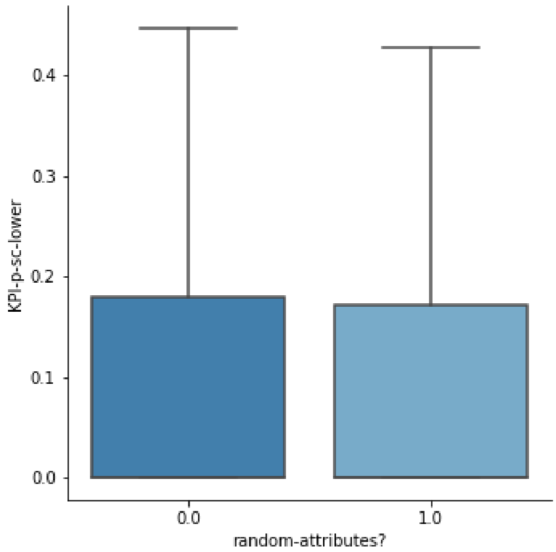


(c) Amount of available houses for sale

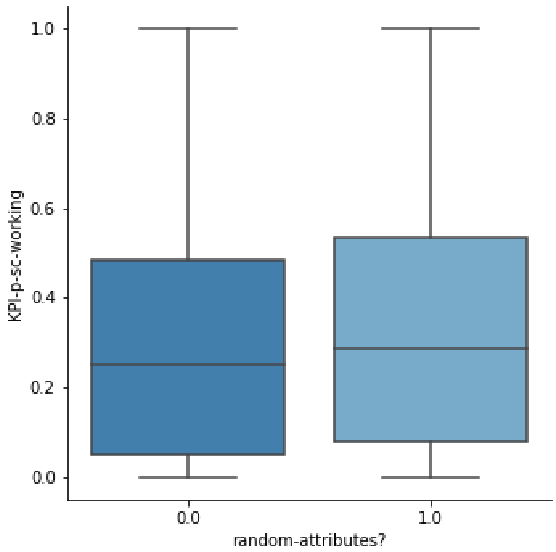


(d) Amount of available private sector housing rent options

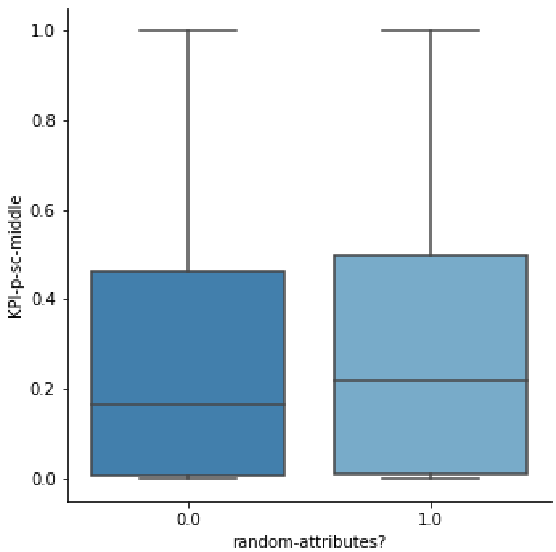
Figure L.5: Part 1 of Boxplots showing significant differences in outcomes of KPIs when testing the sensitivity of the random assignment of agent attributes.



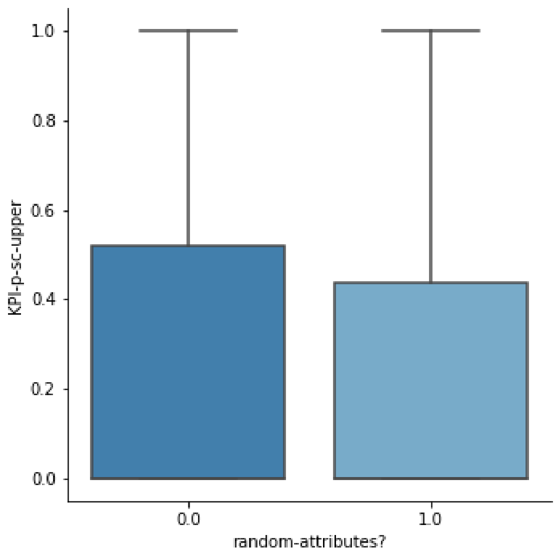
(a) Percentage of people in the lower social group.



(b) Percentage of people in the working social group.



(c) Percentage of people in the middle social group.



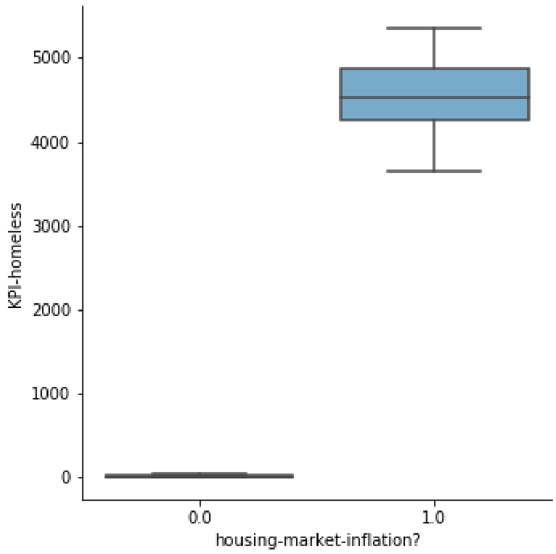
(d) Percentage of people in the upper social group.

Figure L.6: Part 2 of Boxplots showing significant differences in outcomes of KPIs when testing the sensitivity of the random assignment of agent attributes.

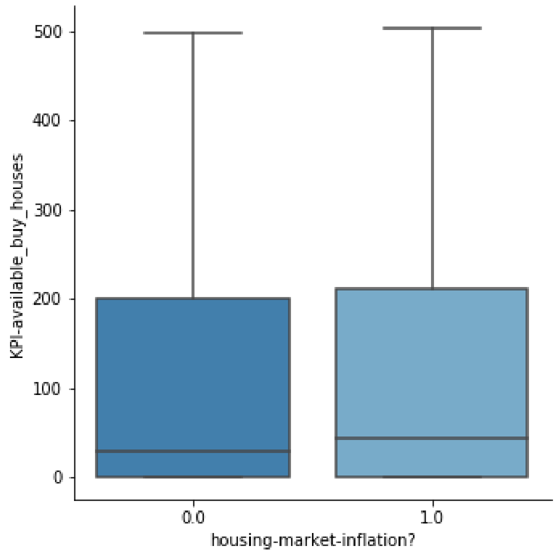
There is also a significant change in citizen composition when looking at the percentage prevalence of each social group when the attributes of citizens are selected randomly (as seen in Figure L.6). This can easily be explained when looking at the way the social group of citizens is calculated (Chapter F). Because education and income no longer relate to each other, a rise in working group and middle social group can be observed, whilst at the same time a decrease in upper social group can be seen, as the combination of both high income and high education is now more rare.

L.1.5 Inflation

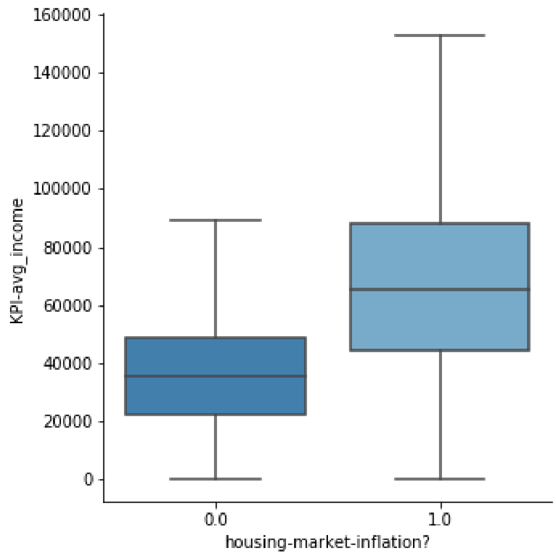
One exogenous factor that is not included in the model by default, but might play a role in real-life and thus in the model, is inflation. Income increases over time as a result of inflation. However, when looking at the trend of average housing market prices, an upward increase in average housing price can be observed as a "market reaction" to a shortage of supply. Since there is a shortage of housing options in the *Randstad*, the supply and demand are not in equilibrium (R. C. Kloosterman & Lambregts, 2001). To add this effect to the simulation, a parameter was added that models an increase in housing price whilst at the same time modeling an increase in income (due to inflation). The increase in housing price is simulated as a steady increase of 7% per year, where inflation increases income by 3-10% (with an average increase of 6%). The resulting changes to the outcomes are shown in Figure L.7.



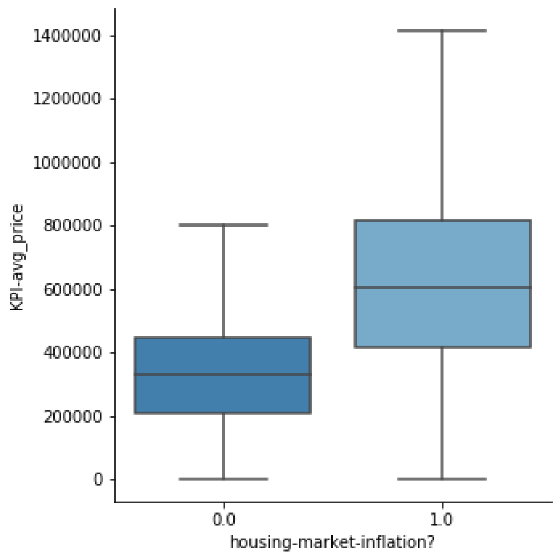
(a) Total amount of homeless people



(b) Availability of houses for sale



(c) Income distribution



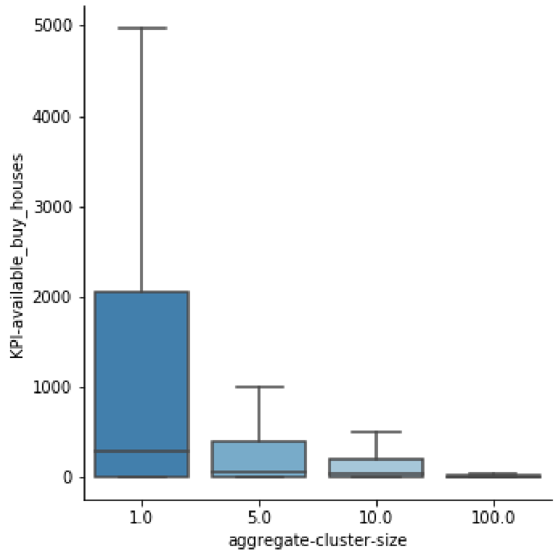
(d) Housing price distribution

Figure L.7: Boxplots showing significant differences in outcomes of KPIs when testing the sensitivity of the inflation mechanism in the model.

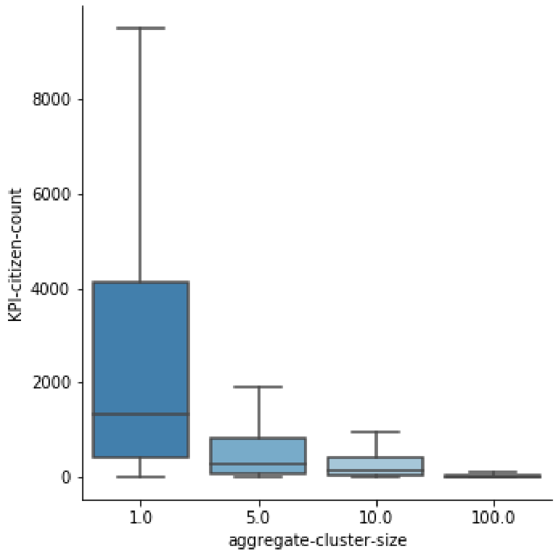
As expected, Figure L.7c shows an increase in average income since the inflation increases income over time. Similarly, the average price of housing options also increases as shown in Figure L.7d. More surprisingly, Figure L.7a shows that the increased housing prices and income also lead to significantly more homelessness. This might be explained by the fact that the range of affordable housing options for migrants is already small (as shown in Section L.1.3) and any change in housing market prices causes them to lose the ability to afford housing. This hypothesis is further confirmed when looking at the availability of houses for sale (Figure L.7b), which rises because less people are able to afford buying a home.

L.1.6 Aggregation Size of Household Agents

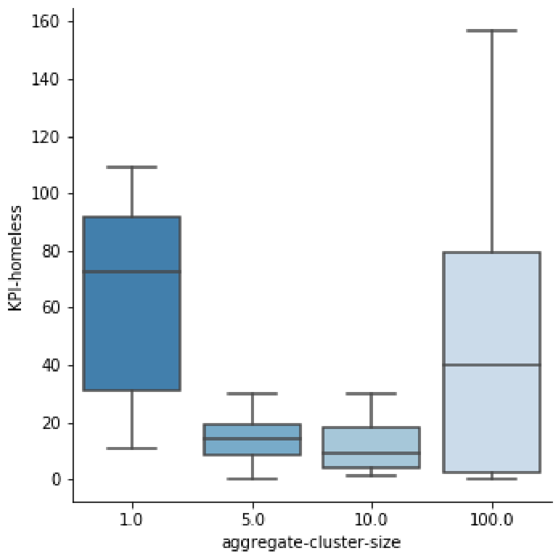
To make sure there is no interference with the outcomes of the model when applying aggregation, the size of aggregation is tested using sensitivity analysis. Here, different aggregation sizes of household aggregation are checked to see if they change the output of the model. The aggregation models the amount of similar households that are represented by one agent. In the most granular form, at an aggregation size of 1, each agent represents a single household, where the least granular value chosen for this analysis is set to 100. The impact of aggregation is shown in the boxplots in Figure L.8.



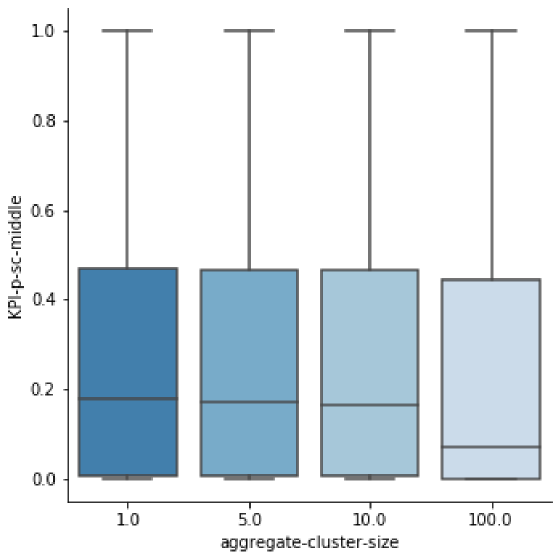
(a) Available houses



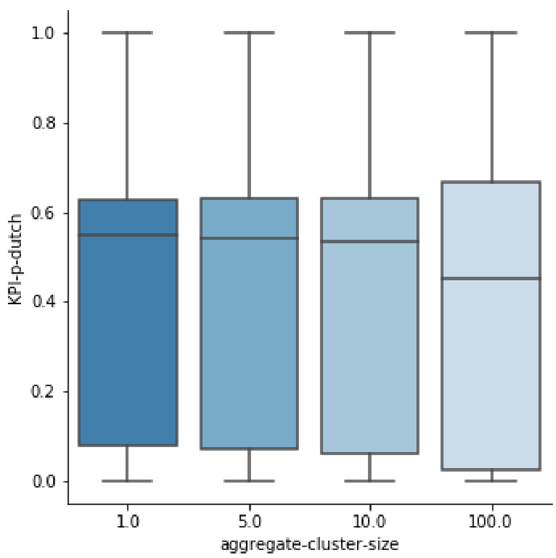
(b) Citizen count



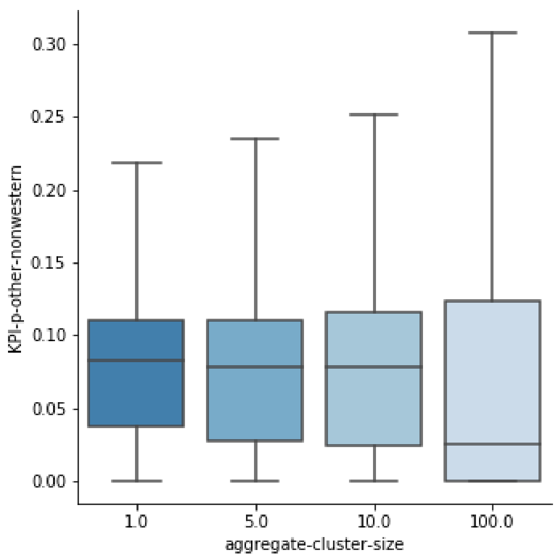
(c) Amount of homeless people



(d) Middle social group



(e) Percentage prevalence of Dutch ethnicity



(f) Percentage prevalence of other Non-Western ethnicities

Figure L.8: Boxplots comparing outputs of KPIs with significant differences when comparing based on the aggregation size of households for citizen agents in the model.

Looking at Figure L.8a and L.8b, there are significant differences visible between the outcome of the KPIs and the parameter values. However, this can be easily explained by the fact that the counting of both housing options and citizen agents in the model is directly tied to the size of aggregation of each agent. In other words, if one agent represents more than one household it will divide the citizen count and availability of houses by the same amount as the aggregation. This relation can be seen in all types of housing availability, but for brevity's sake only the houses for sale graph is shown.

An interesting correlation between homelessness and aggregation size can be observed in Figure L.8c. Confusing at first glance, since there seems to be a non-linear pattern between aggregation size and homelessness, it becomes more clear when taking into account the fact that there are less citizens in the model when aggregation size is low. To account for this difference, the resulting output should be corrected for its multiplier. Figure L.9 shows the homelessness when correcting for the aggregation size, which clearly shows that the model is only negatively impacted when aggregation size is set to 100, which is most probably causing by rounding errors and inefficiency in finding a suitable home.

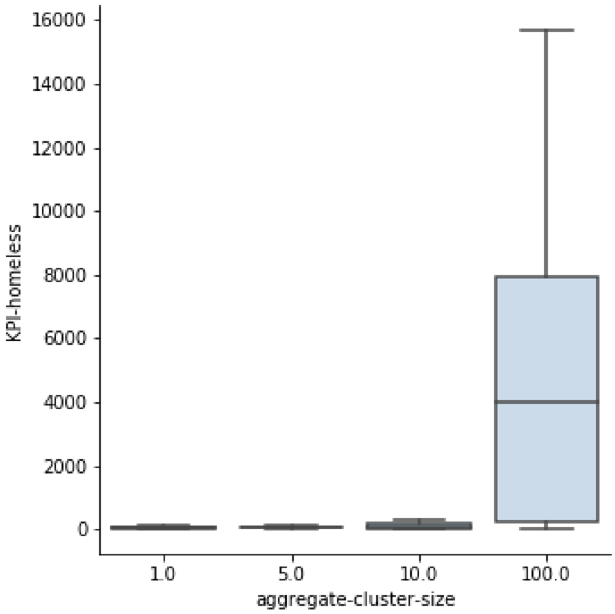


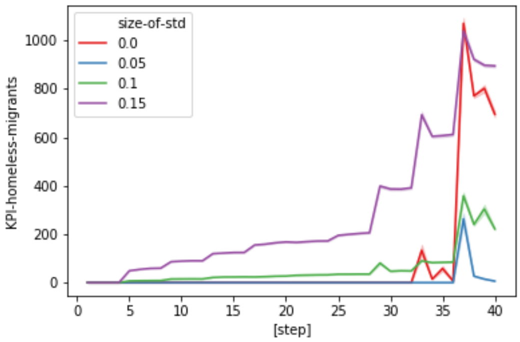
Figure L.9: Homelessness observed with different aggregation sizes, but correcting for the aggregation size (multiplying by aggregation size value).

In Figure L.8d the most significant difference between simulation values can be found for the percentage prevalence of social groups in the model. The other three social groups also show similar differences, but not as significant as the middle social group does. The author has no other explanation for the change in citizen composition than rounding differences, causing the total percentage of all four social groups to be lower than 100%. It should be noted that the difference only start to appear when the aggregation is at the extreme value of 100.

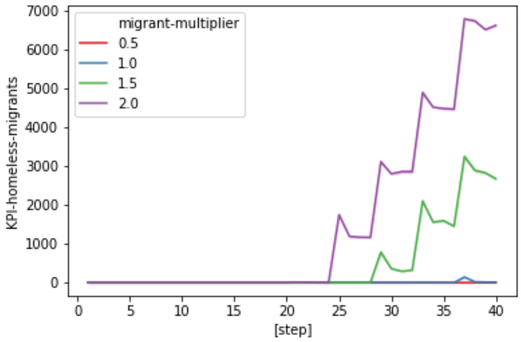
All demographic KPIs which measure ethnicity prevalence show significant differences between aggregation sizes. The Dutch ethnicity is the biggest and the Non-Western ethnicities are most relevant for the calculation of migrant properties and thus these two KPIs and the difference in outcomes are show in Figure L.8e and L.8f respectively. As the outcomes show, there is a steep drop in percentage prevalence for both ethnicities at aggregation size 100 (which can be observed in all of the ethnicity KPIs). This can be explained by looking at the model code. In the model, the percentage prevalence of each ethnicity (from Open Data sources) is divided by the aggregation size and the resulting number is rounded down, which leads to the amount of households generated for each neighborhood with certain ethnicities. Because rounding down occurs, a large aggregation of 100 leads to a rounding down to 0 instead of 1 in many cases (for example, a neighborhood with 9000 citizens and 1% prevalence of Turkish people with an aggregation of 100 would lead to 90 / 100 = 0.9 which is rounded down to 0).

L.1.7 Sensitivity of Parameters over Time

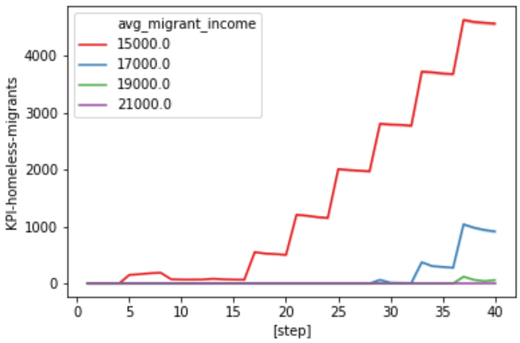
To better understand the impact of differences in parameter values, the changes are observed over time. The easiest way to observe these changes is by plotting the value of the most significant KPI over time, whilst comparing the different parameter settings. Since all parameters show significant impact on the homelessness of migrants, Figure L.10 gives an overview for each of the exogenous parameters over time by looking at the homelessness of migrants.



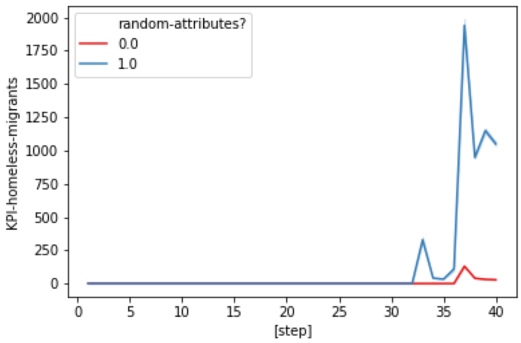
(a) Standard Deviation



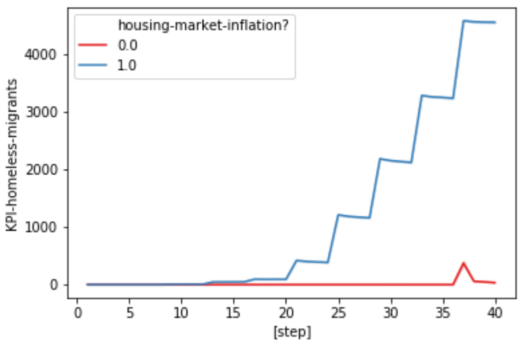
(b) Influx of Migrants



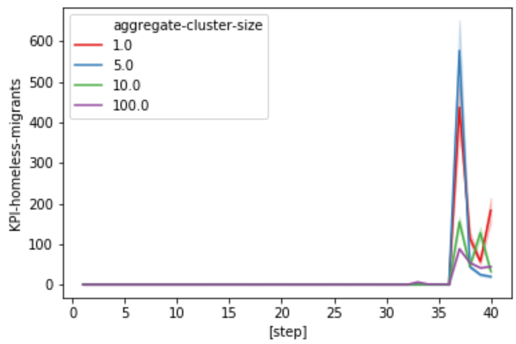
(c) Average Income of Migrants



(d) Random Attributes of Citizens



(e) Inflation of Prices and Income



(f) Aggregation Size of Agents

Figure L.10: Timeline plot showing the development of migrant homelessness over time for each of the exogenous parameters.

When observing the development of homelessness among migrants given the exogenous parameters, most parameters show a cyclical pattern. However, differences in the standard deviation of income shows wildly different behavior at the end of the simulation (Figure L.10a), and assigning agent properties randomly has

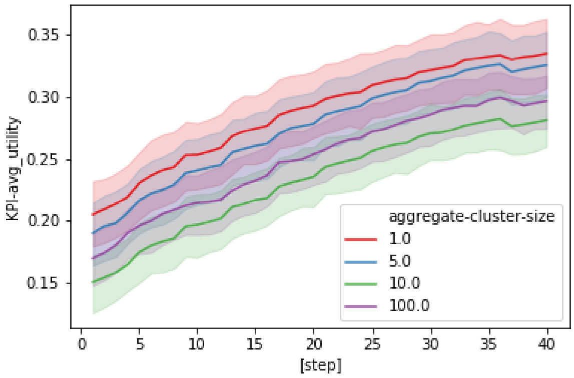


Figure L.11: Development of citizen's utility over time given aggregation of agents

similar unpredictable behavior at the end of simulations (Figure L.10d). All other parameters show "stable" behavior, which is easier to explain given the parameterization.

Figure L.10f shows the relation between the development of homelessness and aggregation of agents. Here we can see that the overall development of homelessness seems to follow the same pattern for each of the aggregation sizes, but the height of homelessness differentiates between aggregation sizes, which can be explained by the fact that a lower aggregation size leads to more agents being present in the model and thus more homelessness occurs proportionally.

As a final check, the development of utility over time for the differences in aggregation size is shown in Figure L.11. Although the author is unable to explain the difference in height between aggregation size 10 and 100, it is clear that the pattern of averages in all runs is the same. This leads to the belief that the *relative* difference between runs is insignificant, and the interpretation of utility between runs is the same, as long as the absolute values of utility are not considered as representative for the quality of a neighborhood. Observing the changes of a neighborhood should thus always look at utility in relation to the utility of other neighborhoods in the same simulation run.

L.2 Endogenous Sensitivity

Apart from testing the influence of parameter values with regards to external factors and exogenous variables, it is also insightful to get a better feeling of the effect of certain policy levers. The levers themselves and their observed changes to the outcomes of the model are described in Chapter 4, but some of the results require a more granular outcome analysis to define how good a policy lever performs, and under what conditions or values.

Because of time constraints, the sensitivity of policy parameters has not been done. Running the simulations for variations of values is expected to take around 36 hours of run time, and interpretation and analysis of data might be equivalent. However, this report is due in 4 hours and so the sensitivity analysis of policies is out of scope of the research (for the time being).