City growth and Zipf’s law

**Delft University of Technology**

**Faculty of Technology, Policy and Management**

**Jaffalaan 5,**

**2628BX, Delft**

**01-Feb-15**

Matthijs Brouns Hélène van Heijningen

[Type the abstract of the document here. The abstract is typically a short summary of the contents of the document. Type the abstract of the document here. The abstract is typically a short summary of the contents of the document.]

# Preface

…

Executive Summary

…

# Table of contents

Table of contents 1

Introduction 2

System analysis 3

Life stage 3

Family 3

Jobs 4

Housing 4

Other concepts 4

Narrative 5

Model logic 6

Household agent 6

Model setup 6

Life stage transitions 6

Moving 7

Verification 10

Single-agent verification 10

Minimal interaction level 11

Multi-agent level 12

Conclusion 15

Validation 16

Methods of validation 16

Model experimentation 17

Conclusions 19

Appendix A: Job graphs 20

Primary and secondary sector jobs 20

Service and non-profit jobs 20

Finance and IT jobs 21

Appendix B: Changing model input values 22

MODEL\_SETUP.nls 23

Households in City 23

MOVE\_CITIES.nls 23

Resistance to move based on Time since moving + Children + Age 23

Moving Distance modifiers for City Attractiveness 23

Attractiveness Jobs Countryside 24

# Introduction

Economic activity is geographically extremely concentrated in cities, and even more in large cities. At present, competition between countries to a large part amounts to competition between their cities. This notion is starting to be embedded in, for instance, Dutch economic policy thinking.

Most countries have a city that is by far the biggest and most dominant city. For instance, London accounts for about 20% of the UK’s GDP and the next biggest city is much smaller. Indeed, it appears that the distribution of city sizes within a country often follows the Zipfian distribution, that is, the size of any city is inversely proportional to its ranking in the list of city ranked by size (Gaujal et al.; 2014). So, competition between countries often boils down to competition between their biggest cities.

How this Zipfian distribution emerges however has not yet been agreed on. There are many ideas on how cities form such a distribution, many of which rely on economic theories. There are of course other ways to approach this phenomenon. The Netherlands Bureau for Economic Policy Analysis (CPB) for example is interested to find out if decisions made at household level can explain the emergence of the Zipf’s law. They wish to understand more on the emergence of the Zipf’s law and how it may be influenced by future policy making or how it will influence future policies. This research has been conducted in order to bring light to these matters. Another interesting phenomenon concerning the Zipf’s law is that for some countries the Zipfian distribution does not hold strictly, i.e. with a power coefficient of one. The Netherlands is one of those countries, with many cities of more equal size than the Zipfian distribution would predict. The CPB wonders if that is a problem or an asset? What can be the reason for this? Their first approach is focused on decisions at household levels, which leads to the following research question posed in this report:

*To what extent do decisions made at household level influence moving behaviour in cities to cause the emergence of the Zipf's law?*

Two studies (Mansury and Gulyas, 2007, Gaujal et al., 2014) showed that a Zipfian distribution of cities can be generated with an agent based model. An agent based model tries to explain complex macro patterns, like city size, on the basis of individual decision making. These individuals take account of their preferences and their perceived environment and they possess some heterogeneity in decision making. In comparison with traditional economic methods, the agent based modeling approach is more simulation oriented and less analytical. Since CPB is interested in a focus on how individual decision making influences city growth, the research question posed in this report will be answered by using an agent based model of decisions made at household level and how these decisions influence city growth. The CPB is also interested to see what this method of research offers and how it might be of help to them in future research.

This report will further explain the narrative on which the model will be based, the assumptions made, and how this narrative is transformed into model logic. This will all be explained in chapter X and X. Next, the verification and validation of the model will be discussed after which the experimental design will be elaborated on in chapter X and the model experimentation will be described in chapter X. Finally, this report will finalize with an answer to the research question in the concluding chapter X.

# System analysis

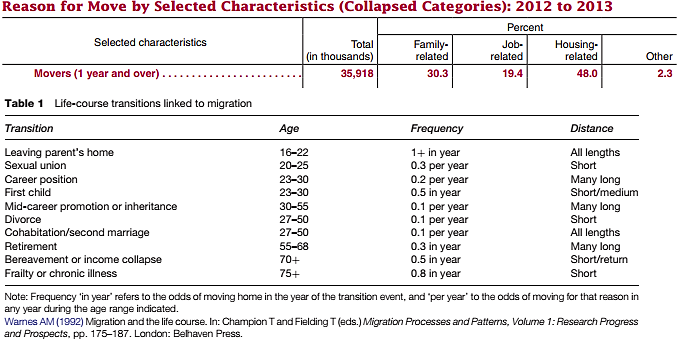
This chapter will discuss the basic concepts that have been identified which will be taken into account when conceptualizing the agent based model. The concepts have been identified using a literature research and by expert consultation.

There are of course many reasons for a family to move to another city. The most important reasons for moving can be categorized by reasons based on someone’s life stage, family, job opportunities or housing (XXX, XXXX). Next to reasons for moving, there are other concepts in play which influence people’s decision-making. Together with CPB we have decided on several concepts that will be added into our analysis.

## Life stage

Research done by XXX shows that the amount of times someone moves as well as the distance between their current location and future home changes quite a bit for different life stages. This is shown in table 1.

Table 1: Life-course transitions linked to migration (Census, XXXX)



Not all life-course transitions are interesting enough to take into account. For example, life-stage transitions where the moving distance is short (within the same city), will not be taken into account because the model will focus on moving behaviour between cities. Furthermore, this level of detail is not needed in order to build a stable number of households during a long running period. Overall, it can be concluded from this graph that young adults and adults up until the age of 55 move relatively often with medium up till long moving distances. After the age of 70 households do not move over long distances anymore. Between the age of 16 and 22 people move very often as well as all types of moving distances.

## Family

Family-related reasons for moving overlap the life stage transitions as family-related reasons are for example leaving parent’s home, sexual union, first child, divorce and second marriage. Other reasons could be that one wants to live closer to his or her family. Since these reasons are overlapping with the life stage transitions, family-related moving will not be taken into account as a separate category in the model.

## Jobs

Research done by XXX (XXXX) shows that people who move for job-related reasons travel a very large distance to their new home. This makes this category very interesting. Furthermore, as many researchers indicate that a Zipf’s law is probably related to economic factors, we find this category very important. There are three types of jobs that are often distinguished in moving models (TIGRIS , CPB), namely primary sector jobs (agriculture, manufacturing), secondary and quaternary sector jobs (service and non-profit jobs) and finally tertiary sector jobs (for example, IT and finance jobs). These types of jobs all have different effects on city growth as for example primary sector jobs are mostly found in small cities where-as tertiary sector jobs are found in large, growing cities. These concepts are very interesting and because they might have a large influence on how cities grow, they will be taken into account in the model.

## Housing

The housing category focusses on people moving to either a smaller or a larger house. This is tied to the housing market. In order to implement this category the housing market would have to be modelled as well. Because of the time available for this project this cannot be modelled and will therefore not be taken into account in the rest of the modeling or analysis.

## Other concepts

There are various other reasons for people to move or for certain cities to be more attractive than others. Together with the CPB, the most important concepts that are within the scope of this project have been identified. The decision to move is then further influenced by the following concepts:

* Having children
* Aging
* Time since moving
* Distance from current location
* City size (facilities)
* Borrowed utility

These concepts will now briefly be described in more detail. Since moving in this project means moving between cities and families with young children often only move within their current city, the concept of having children is implemented so that when a family has young children they will move less often.  
  
The next concept, aging, influences how often someone will move. Since literature indicates that young adults move more often than older adults, the aging of a household can be used to determine how often they move.  
  
The time since moving also influences how often a family moves. Since people will not move right away after they have just moved or after they have been living somewhere for over 30 years, the time since moving can be used to incorporate these effects.  
  
The distance from the current location can be used to mimic the disadvantage of moving to a city which is very nearby as well as moving to a city which is very far away from your current location. Moving to a city nearby (within cycling distance) is not needed for many families. Moving to a city very far away is in reality not advantages either because a household will have connections, friends and family nearby at their current location.

Lastly, the concepts of ‘city size’ and ‘borrowed utility’. Studies of Gulyas (2007) and Gaujal et al. (2014) emphasize the well-known phenomenon in urban economic literature; individuals are attracted by large cities. This is because large cities have many facilities, making it more attractive to live in such a city. The concept of ‘borrowed utility’ is closely linked to this concept as people in one city can use facilities of a large city when it is located nearby, thereby increasing the attractiveness of the smaller city.

# Narrative

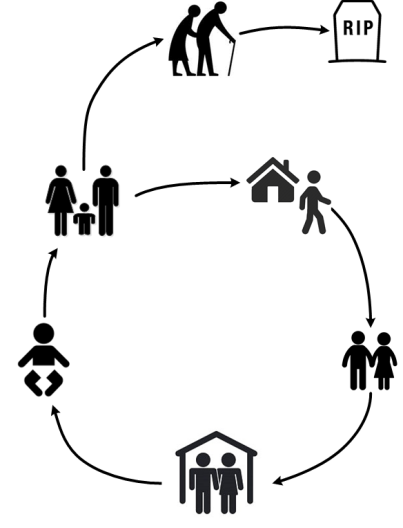
In this chapter the concepts discussed in chapter X will be translated into the model narrative. The agents in the model are households. This narrative will therefore focus on what a household goes through during a typical model run.

Figure 1 visualizes the life stage transitions a household will go through. We will start with a household of two parents with a young boy, shown in the left of the figure. The boy will grow up and decide to leave his parent’s home. The young man will then move to the city which is most attractive for him. He will then in the coming years find a girl who he will be a couple with. These matches always happen in the same city. Two single households are then merged into one. During these years the couple is relatively young and will move quite often due to finding new jobs or being promoted to a higher function. Then the couple gives birth to a child, after the child grows up and leaves the house the parents grow older until they retire at the age of 65. Then they find the countryside more attractive than cities, thereby mimicking the moving behaviour of retired adults in reality. This cycle will go on and on during the model run. During a households lifespan it will move on several occasions. Whether a household moves is dependent on many factors next to the life-stage transitions, as described in the previous section. These concepts and how they influence a household are described shortly in table 2.

Figure 1: Household life cycle

Table 2: Moving concepts and their influence on a households decision making

|  |  |
| --- | --- |
| Having children | Households are less inclined to move to another city when there are children. |
| Aging | A young household moves more often than an older one. |
| Time since moving | This determines when the household will be more inclined to move again. When they have just moved, they will not move immediately again, but when they have not moved for a long time, they are more inclined to stay. |
| Distance from current location | This distance determines whether moving to the city in question becomes less attractive because it is either very close and therefore moving is not needed, or it is too far away from their connections. |
| City size | The larger a city, the more attractive it is because of its increased number of facilities. |
| Borrowed utility | Utility gained from nearby cities will increase the attractiveness of the city in question. |
| Job attractiveness | Cities are more attractive based on the attractiveness of a certain job in that city. |

# Model logic

This chapter will discuss how the narrative has been translated into the actual model code.

## Household agent

Agents in the ABM model represent households. A household has members, which is coded as a list of people. Every household member has the following attributes:

* Age (0 to 100)
* Job preference (1 to 7)
* Sex (Male / Female)

Next to household members, a household also has a location. This location can be in one of the cities or on a patch in the countryside.

# Model setup

The model starts by choosing locations for the X number of cities the model is run. Then the X number of households are divided over these cities and the countryside. The first city is always placed in the middle of the map, after which the other cities are distributed randomly in a circle with a certain width around this first city. This width is minimally 250 patches and maximally 500 patches. The households are distributed over the cities and the countryside. 40% of the households are placed in the cities whereas 60% of the households are placed in the countryside at the start of the model run.

## Life stage transitions

Every member in a household gets older each tick of the model. They progress through different life stages in the model according to the flowchart shown in figure 2.

When a child is born a member is added to the household of 2 or more members. A child is aged 0 and the sex is chosen randomly. At birth they are already given a job preference from 1 to 7:

1. Primary sector jobs (18% of people)
2. Secondary sector jobs (32% of people)
3. Tertiary sector jobs; Financial (8% of people)
4. Quartary sector jobs; Non-Profit (34% of people)
5. Jobless (8% of people)

These numbers are based on data from CBS (2014) so that the model is based on a realistic distribution of job types.

Each tick the child grows a year older. Between the ages of 16 and 23 the child will move out of its parents’ home and move to the city that is most attractive for him/her. This will spawn a new household in the model.

Figure 2: Life stage transitions as implemented in the model

Between the ages of 23 and 30 adults will then find partners that live in the same location. When a couple is matched, the households are merged into one household and they stay in the same location. It is possible for older people to be in search for a partner as well (<50 years old), because some people are not able to find a partner before they turn 30.

The couple then has a change to reproduce each year equal to 1 in 7. In order for a couple to reproduce they both have to be younger than 40. When a child is born, this child is added to the household and that child’s life stage progress starts at the beginning of the flowchart.

Adults retire when they are 65 years old. They then become more interested in moving to the countryside, thereby mimicking the moving behaviour of the elderly who move out of the cities. Adults die between the age of 60 and 100. This is coded using a normal distribution with u = 80 and s = 6. When the last member of a household dies, the household is removed from the system.

## Moving

Each household has a resistance to move and a willingness to move which is actually the attractiveness of a city for each household. When the attractiveness of one or a number of cities is higher than the household’s resistance to move, the household will move randomly to one of the cities which attractiveness exceeds their resistance to move. This can be translated into one rule the model follows:

“A household moves to a random city for which its ‘Attractiveness’ is greater than the households’ ‘Resistance to move’.

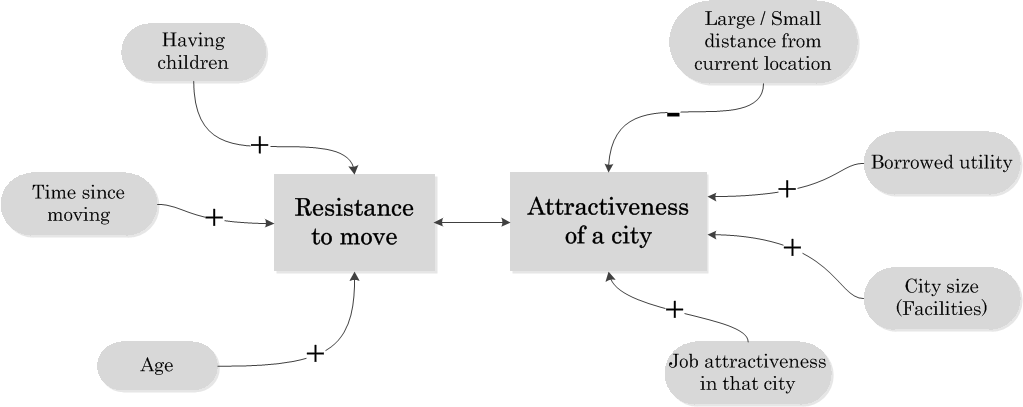
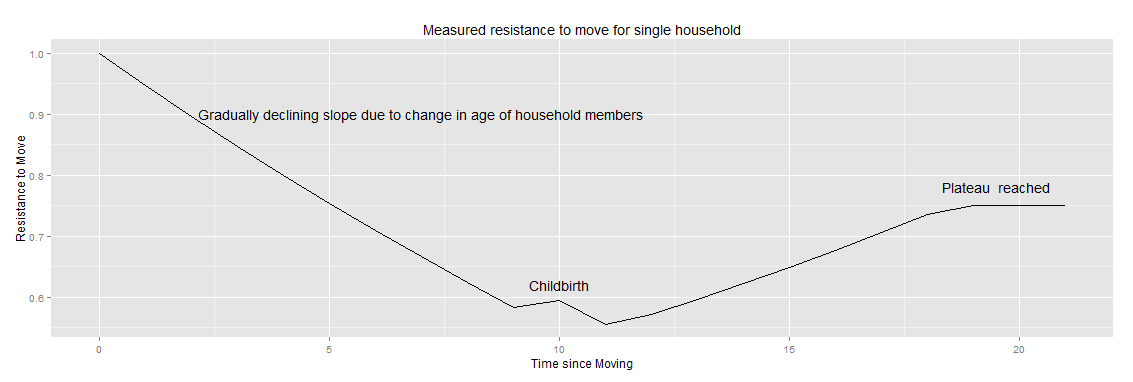
Figure X shows the factors that influence the resistance to move and the attractiveness of a city.

Figure 3: Model logic with regards to the moving decision making process

### Resistance to move

As you can see in figure 2, the resistance to move is influenced by three factors:

* Number of children in a household
* Time since moving
* Age of adult household members

The effects of these factors are summarized in a graph which is shown in figure 3. The x-axis show the number of years since the last time the household moved to another location. This resistance to move (y-axis) can then become higher when a household has children and the lowest point of the graph is moved to the right when a household is relatively old, so the elderly do not move as often as young people do. The resistance to move is a value between 0 and 1.

### Attractiveness of a city

Figure 4: Measured resistance to move for a single household

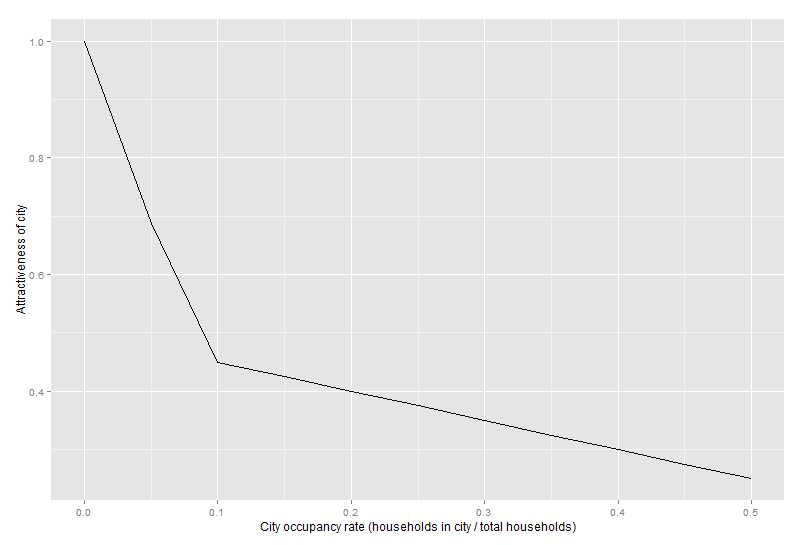
The attractiveness of a city to a certain household is dependent on the job preferences in that household and the distance from their current location to a city. The attractiveness of a city also takes on values between 0 and 1. We assume that people prefer to stay at their current location when a more attractive city is very close by their current location. We also assume that people do not move to cities that are very far from their current location. This translates into lower city attractiveness for cities that are close by or very far away from a household’s current location.

The attractiveness is also influenced by the household members’ job preference. Each city has a unique attractiveness score for each job based on the amount of people in the city with that job. The way in which this attractiveness is determined differs per job type:

Figure 5: Primary job attractiveness graph in one city as a function of the city occupation rate

*City occupation rate*

*Attractiveness*

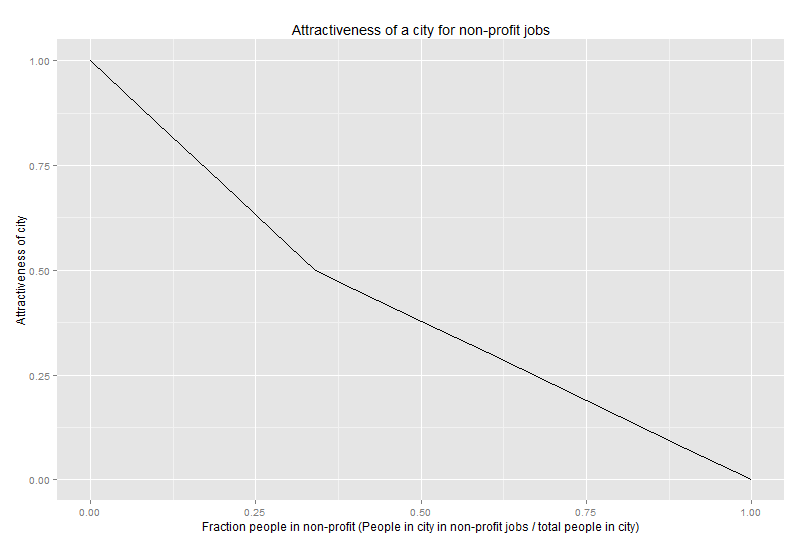


Cities are more attractive to manufacturing and agriculture jobs if they are relatively small cities. The size of cities in this case is used as a proxy for land prices, which greatly influence the profitability for these types of jobs.

Figure 6: Job attractiveness graph for one city as a function of the fraction of people in 2nd & 4th sector jobs

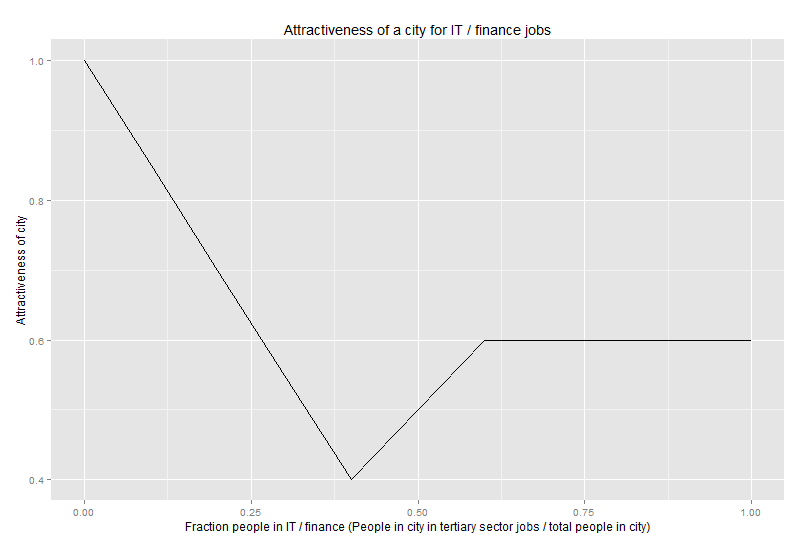
*Fraction people in 2nd & 4th sector jobs*

*Attractiveness*



Jobs in the service and non-profit domain are not influenced by city size in general. Instead, jobs in these domains are expected to spread evenly over all cities. This means that cities with less service jobs than average is more attractive for service jobs.

Figure 7: Job attractiveness graph for one city as a function of the fraction of people in 3rd sector jobs



Service-like effect

Specialization effect

*Fraction people in 3rd sector jobs*

*Attractiveness*

There are also two groups of jobs which show a networking effect. Cities are more attractive for IT and finance jobs if more people are working in that domain in the city.

Finally, each city is evenly attractive for jobless people. In Appendix A more detailed explanations of the job attractiveness graphs are presented for the different job types.

### Attractiveness of the countryside

The attractiveness of the countryside is a constant value for the different job types, so that for IT and finance job types the countryside is not very attractive, but for primary and secondary jobs the countryside is very attractive. This results in the following values from job type 1 to 7: 0.55; 0.55; 0.50; 0.45; 0.45; 0.5; 0.5.

### Distance from current location

…

### City size

…

### Borrowed utility

…

# Verification

Before employing the model to examine the effect of household decision making on the emergence of a Zipf's law, it is necessary to ensure the model is verified and valid. Model verification is used to determine whether the built model performs in a manner in which it was intended to perform. The model was verified on three distinct levels: Single-agent level, in which the behaviour of a single household is analysed; minimal interaction level, in which the interaction between a small set of agents is observerd; and multi-agent verification which analyses the emergent behaviour of multiple agents.

## Single-agent verification

There is only a single type of agent in this urban migration model, namely households. Households live in a certain location and consist of household members.

As described in section x, households follow a certain lifestage progression chain. This lifestage progression has been verified by attempting to falsify the outcomes. To do this, we have first determined the possible composition s of households. Then, a large number (100,000) of households were generated and the model was run for 100 years. In addition, a test was written in NetLogo to verify whether any of the households were in a state which should not be possible.

The possible household compositions are:

* Single member household with 16 < age < 30 OR age > 53
* Two member household with age > 23
* Households with children with two ages > 23 and children with 0 < age < 23

No households were found outside of the expected composition ranges and therefore it is assumed the household progression is implemented correctly.

The second main behaviour of the households is determining the attractiveness of cities and migrating. Since determining the attractiveness of cities largely depends on the households living in them, this aspect cannot be tested at the single-agent level, but will be tested on the minimal interaction level.

Aspects which can be verified on a single-agent level are the calculation of the resistance to move, of which a plot is shown in figure 4 and the actual choice of moving to a different city with preset city attractiveness. Both were found to show no errors.

### Agent Robustness

The implementation of agents in this model can be considered to be quite robust as no normal model runs can cause the agent to break. However, some notions need to be considered when generating custom household agents or when expanding on the model. First, the model assumes that the list of household members is always sorted descending by age. This assumption is made since it drastically improves the performance of several checks in the model. Furthermore, households with children that have fewer than two parents can cause the model to behave unexpected in several fields such as the calculation of city attractiveness and the overall lifestage progression.

## Minimal interaction level

The main behaviour on a minimal interaction level is determining the attractiveness of cities based on the people living in them. Several scenarios were set up and in each of these scenario's the attractiveness for a representative set of households was determined.

In the first scenario the attractiveness of cities for manufacturing and agriculture jobs was tested. Three cities were set up containing respectively 5% of total inhabitants, 20% of total inhabitants and 50% of total inhabitants. The attractiveness of two households consisting of solely agricultural members and solely manufacturing members were then checked against the assumptions and found to be correct. The results are shown in table x.

Table 3: Results of three scenarios on the attractiveness of cities for primary sector jobs

|  |  |  |
| --- | --- | --- |
| Inhabitants | Jobs | Attractiveness |
| 5% | Agriculture | 0.68 |
| 5% | Manufacturing | 0.68 |
| 20% | Agriculture | 0.4 |
| 20% | Manufacturing | 0.4 |
| 50% | Agriculture | 0.25 |
| 50% | Manufacturing | 0.25 |

The second and third scenarios was set up to verify the attractiveness calculation for service and non-profit jobs. Three cities for each job were set up containing respectively a low number of people in that job (1%), a medium number of people (30%), and a high number of people (80%). The results are shown in table x.

Table 4: Results of three scenarios on the attractiveness of cities for secondary and quaternary sector jobs

|  |  |  |
| --- | --- | --- |
| People in Job | Job | Attractiveness |
| 1% | Service | 0.84 |
| 1% | Non-Profit | 0.85 |
| 30% | Service | 0.53 |
| 30% | Non-Profit | 0.56 |
| 80% | Service | 0.14 |
| 80% | Non-Profit | 0.15 |

The final two scenarios were set up to test the attractiveness for IT and Finance jobs. Four cities for each job were set up containing respectively a low number of people in that job (1%), an average number of people (4%), a moderately high number of people (6%), and a high number of people (8%). The results are shown in table x.

Table 5: Results of four scenarios on the attractiveness of cities for tertiary sector jobs

|  |  |  |
| --- | --- | --- |
| People in Job | Job | Attractiveness |
| 1% | IT | 0.87 |
| 1% | Finance | 0.87 |
| 4% | IT | 0.4 |
| 4% | Finance | 0.4 |
| 6% | IT | 0.6 |
| 6% | Finance | 0.6 |
| 8% | IT | 0.6 |
| 8% | Finance | 0.6 |

## Multi-agent level

On the multi-agent level the overall emergent behaviour of the model is tested under various circumstances. First, an extreme value analysis is performed on the two global model parameters which set the total number of households and the total number of cities in the model.

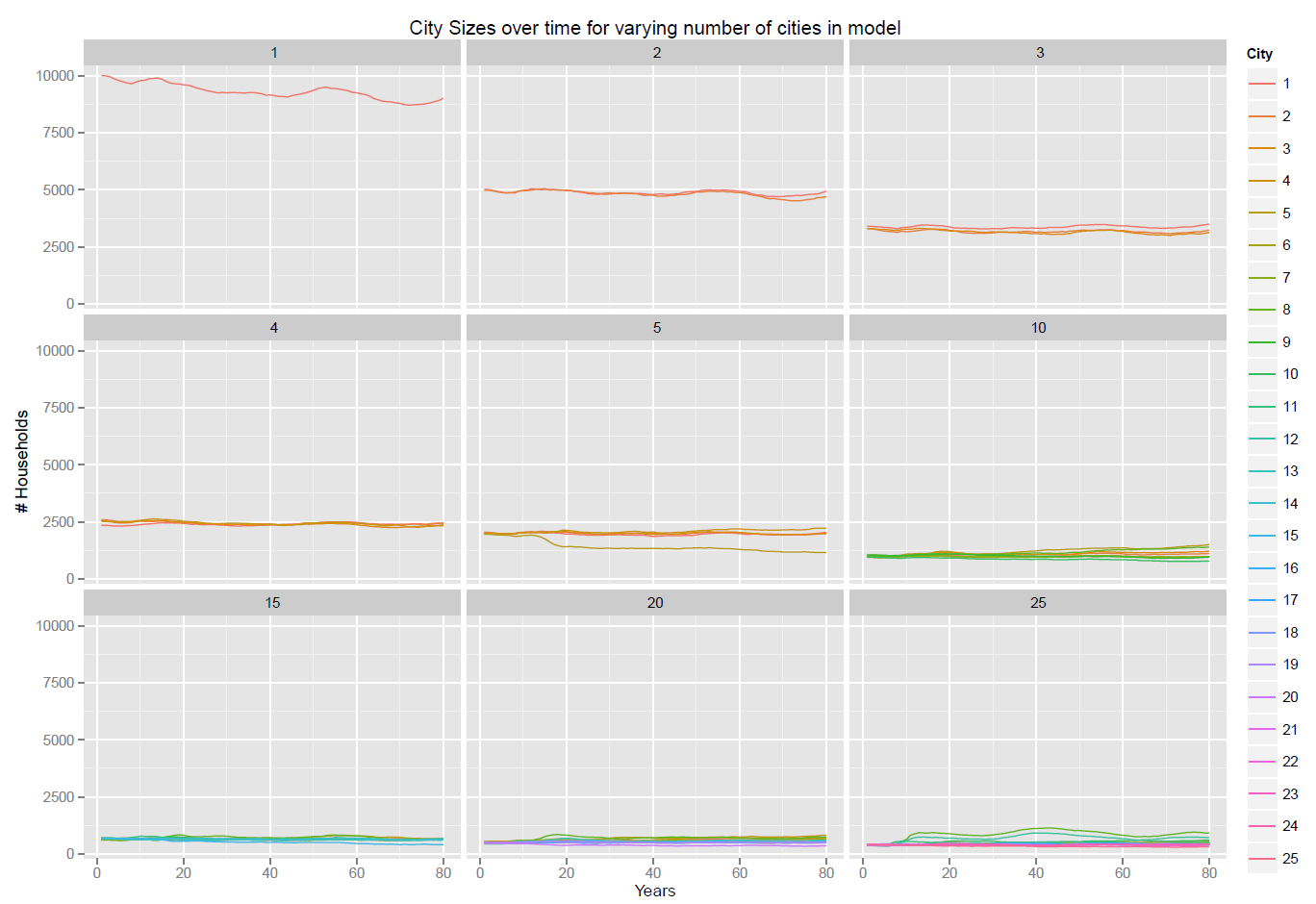
Varying the the total cities in the model from 1 to 25 gives results as shown in figure x. As can be expected, the model does not behave with a very low number of cities. This is mostly because the option space of agents is then too small for interesting behaviour to emerge. From 10 cities and up the model shows expected behaviour. When further increasing the number of cities (>25) some care should be taken that the total number of households will also be increased. If this is not done, the average households per city will be too low which causes unexpected behaviour as described in the following section.

Figure 8: City Sizes over time for a varying number of cities in model

A similar test was performed by varying the the total number of households in the model from 1,000 to 100,000 (figure x). This analysis shows that the model is less suitable for use with a very low number of households (< 10,000). This has several causes: First of all, with such a low number of households it is more difficult for agents in the model to find partners, leading to a decline in total households over time. Furthermore, the city attractiveness effects are smaller which leads to a lower than intended rate of migration. When running the model with a very large number of agents, it can happen that there are large oscillations of city size in the start of the run. This behaviour is caused by two aspects. Firstly, there are several large oscillations in the population system underlying the model in the transitory state (see figure x). Another reason is that normally the city attractiveness only gets calculated on a yearly basis to improve performance. In order to increase the fidelity of large-scale models the "updateCityAttractivenessFreq" slider can be used to set the model to update this attractiveness on a more frequent basis. When this slider is set to a suitably small value, the model results will closer reflect the behaviour as it is with lower amounts of households.

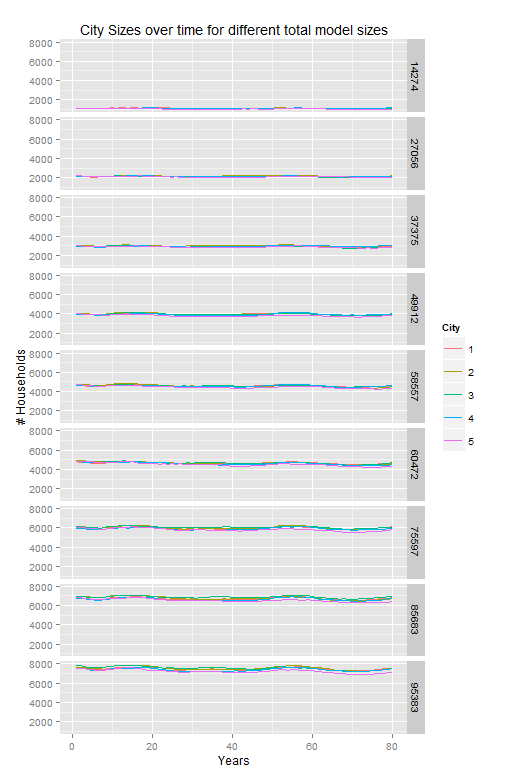
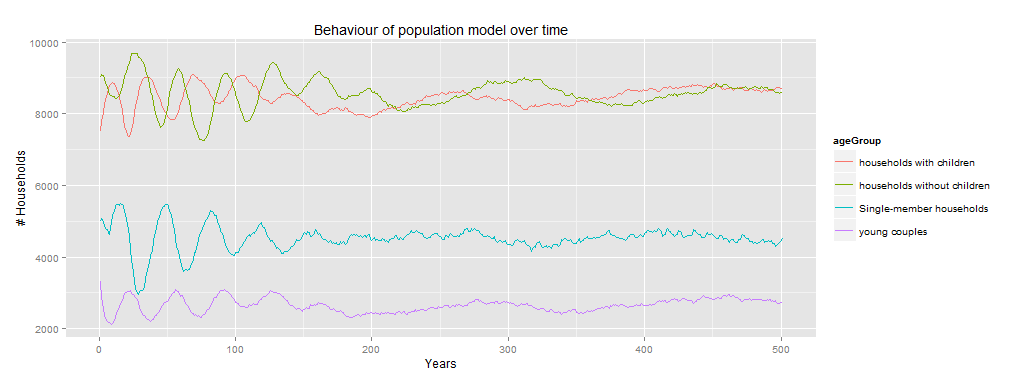


Figure 9: City Sizes over time for different total household sizes

Figure 10: Behaviour of population model over time

Apart from extensively testing these two model input parameters a full parameter sweep with 300 model runs on all  other possible input parameters has been performed. In none of these model runs did the model give any errors and the results are all within the expected bounds (see figure x).

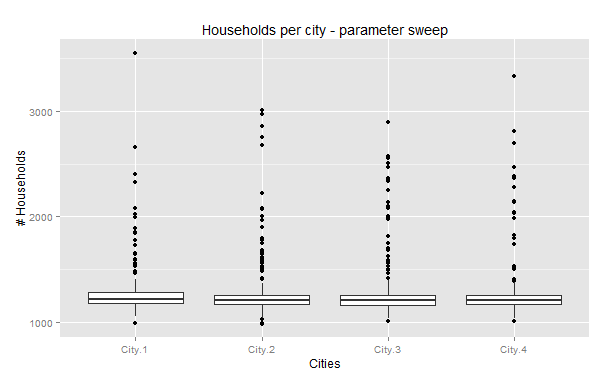


Figure 11: A parameter sweep for the number of households per city

## Conclusion

This model has been verified by 6 tests on a single-agent, minimal interaction and multi-agent level. All of these tests show that the model is implemented in line with the intended implementation as described earlier. However, some care must be exercised in setting up the model for experimentation to make sure the input parameters are within certain bounds to ensure the model behaves as expected.

# Validation

In this chapter the validity of the model is discussed. In order to determine whether the model is valid, it must be checked whether the model outcomes correspond with reality, i.e. does a Zipf’s law emerge, and if the model makes correct predictions. The goal of this research was to understand how individual decisions at household level influence moving behaviour in cities to cause the emergence of the Zipf's law.

In order to answer the questions posed above, two types of validity tests have been conducted, namely literature validation and face validation through expert consultation.

## Literature validation

The model parameters in the population model are based on literature research and this part of the model is therefore considered valid. Since the part of the model focused on the decision making on whether to move or not is also based on concepts found in literature and concepts identified by experts, the factors included in this part of the model are also deemed valid. What exactly has been validated through expert consultation will be discussed in the next paragraph.

## Face validation through expert consultation

The model consists of two parts, the life stage progression of households and the moving behaviour of households. The first part has been validated by CPB experts by showing them the behaviour of the population model over time.

The second part of the model, the moving behaviour of households, has been validated by discussing the underlying assumptions, included concepts and model relations. The CPB deemed our model to be valid for the purpose of this research.

# Model experimentation

So far this document has focussed on the conceptualising, building and testing of a model for determining the effect of household-level decision making on the emergence of a Zipf's Law in city size distribution. The following section will go in depth on the experimental design used to analyze the model's behaviour and the results which came from it.

The model should be considered to be an infinite system. Such a system will not return to its initial state after a predetermined amount of time. Since the model is an infinite system a warm-up time will be required before data can be gathered from the model. The system is expected to reach steady-state after a full life-cycle of the initial households. Since the life-expectancy of a household (not a person) is about 40-60 year, the model warm-up time should be higher than 60 years. Apart from reaching steady state in migrational behaviour, the model also exhibits some transitory behaviour in the initial period with regards to the total population (see figure x). This population system reaches steady state at around 250 years. However, in experimentation it was found that this transitory behaviour does not effect the outcomes heavily and therefore this factor was omitted in determining the warm-up period as to reduce total model run time. Therefore, For this experimentation a warm-up period of 80 years is chosen.

As became clear in section x, there are large uncertainty bands around many of the model parameters. For this study, a multi-variate analysis was set up in order to determine which factor or combination of factors has the largest effect on an emergence of a Zipf's Law.   
  
Due to the large input space a full-factorial experimental design in unfeasible within the set time-constraints. Within these time-constraints it was possible to perform about 300 model runs. In order to still get decent results, latin hypercube sampling was used as a sampling method to ensure a uniform distribution over the input space.  
  
After running the model the individual runs were evaluated on how well they approach a Zipf's Law. This was done by calculating the root-mean-square error of a run compared to a true zipfian distribution with 1 as a Zipf's coefficient.   
  
Analyzing the effects of the change in input parameters on the model behaviour was done by using a decision tree. This decision tree was inferred using an C4.5 Decision Tree inference algorithm supplied by the RPart R Package. In order to avoid overfitting on the data, the tree was pruned optimizing on a lowest cross-validation error. The results of this decision tree are shown in figure x.

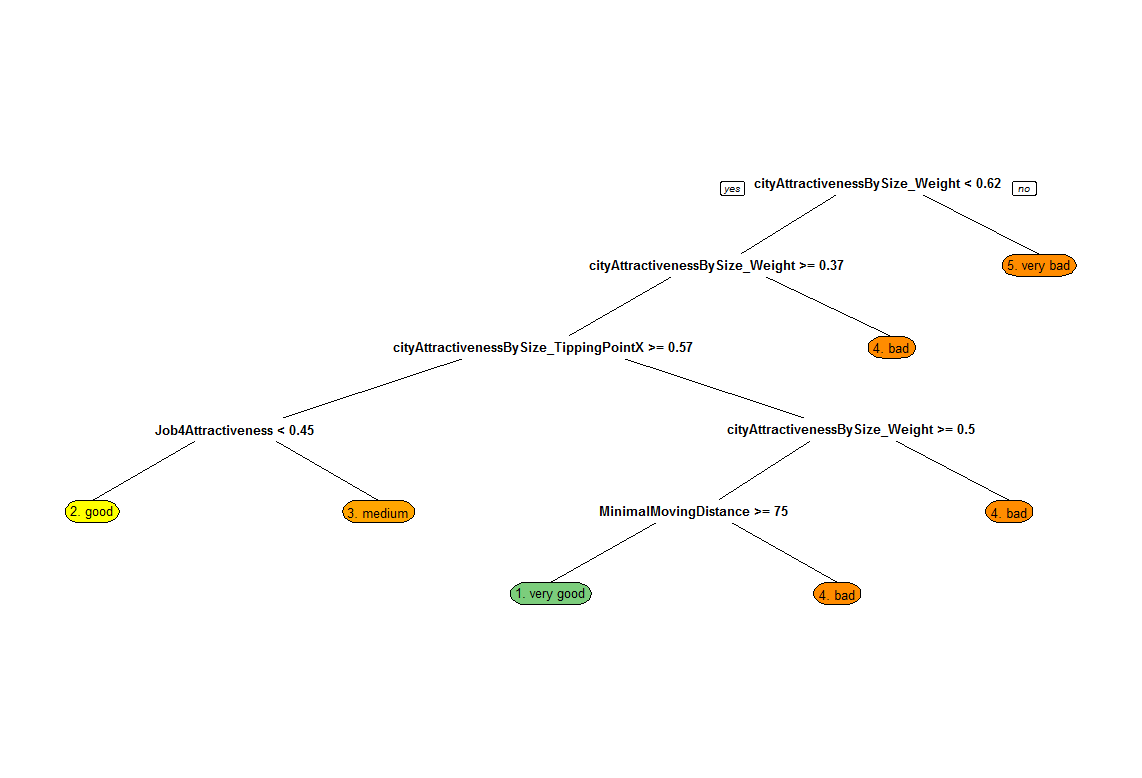


Figure 12: The effects of changes in input parameters visualized in a decision tree

This decision tree shows that a Zipf's law is most likely to emerge when the weight of the city attractiveness by size modifier is between 0.5 and 0.63, and when the people are less inclined to move to cities within 75 patches from them.

This result has been verified by running 20 new independent replications with these settings. From these runs we get an average RMSE of 435 while the original runs had an average RMSE of 1412. This result shows a massive improvement in the emergence of a Zipf's law which leads us to conclude that this moving distance is indeed a large factor in the emergence of a Zipf's law in the model.

# Conclusions

…

# References

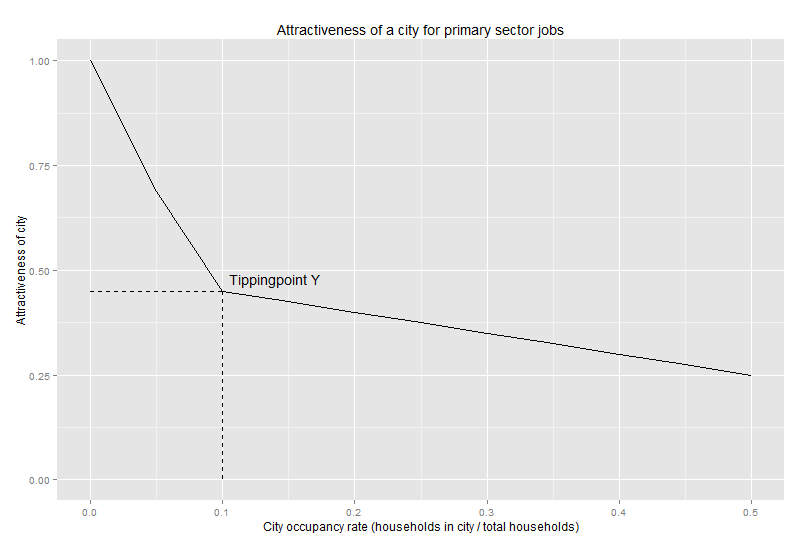
# Appendix A: Job graphs

This appendix shows under which circumstances certain jobs are attractive. Different job types rely on different formulas to calculate the attractiveness of a city for this job type, which results in different ranges in x-axis. Why they differ will be explained for each job type.

## Primary sector jobs

Job 1 and 2 are primary and secondary sector jobs, which attractiveness for a city is determined by the fraction of households in a city. The graph showing the attractiveness of a city for that job type for different fractions is shown in figure 4. The y-axis is the city attractiveness for that job type and the x-axis is the fraction of households in that city. These jobs are coded in such a way that cities with low density are more attractive for these job types.

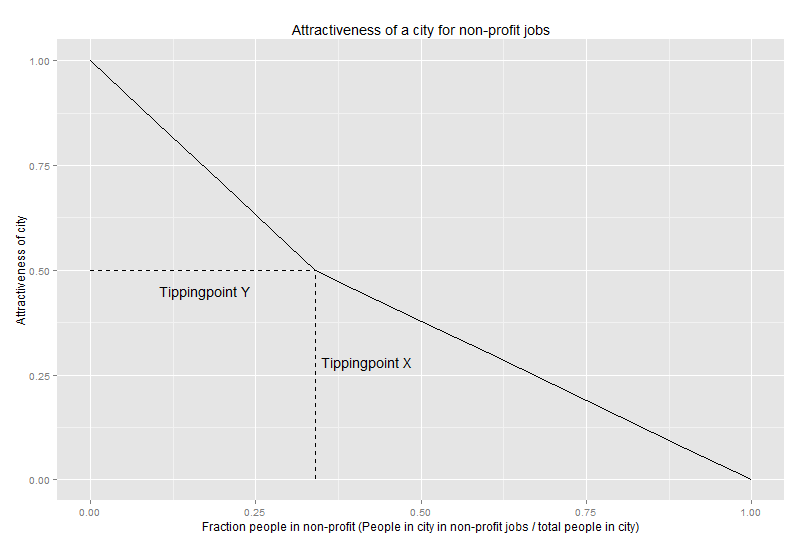
The x-axis ranges from 0 to 0.5 because it depicts the fraction of households in city X from the total number of households in the country. In the base case of the model there are 5 cities, therefore if all households were distributed equally, each city would have this fraction be equal to 0.2 and a city with a fraction equal to 0.5 would be an extremely large city. Therefore a range from 0 to 0.5 takes into account all possibilities.



## Service and non-profit jobs

Job 3 and 6 are service and non-profit jobs, which attractiveness for a city is determined by the fraction people in service / non-profit. These jobs are coded in such a way that all cities need a certain percentage of people performing these jobs compared to the total number of households in that city. The graphs in which these effects are taken into account is shown in figure 5 and 6 respectively. The x-axis shows the fraction of people in service and the y-axis shows the city attractiveness for that job type.

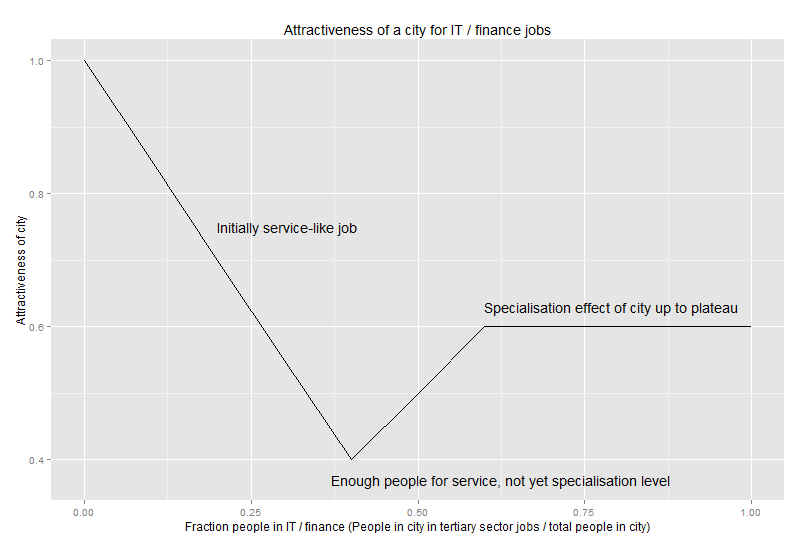
The x-axis has a range from 0 to 1, which is different from the graph in figure 4, because the fraction used to make this graph is different. Here the fraction is not the number of households in city X compared to the total number of households, but the number of people in service or non-profit jobs compared to the total number of households in that city. When this fraction is 0, the attractiveness for this job should be very high as a percentage of the total households in that city is always required to be filled for this job. When this fraction is equal to 1, the city attractiveness for this job type should be equal to 0 so that no more people take on such jobs than is necessary. Therefore the range of the x-axis is from 0 to 1.



## Finance and IT jobs

Finally, job 4 and 5 are finance and IT jobs, which attractiveness grows the more finance and IT jobs are taken in a certain city. This effect is shown in figure 7 where the x-axis presents the fraction of people in finance/IT and the y-axis shows the attractiveness of a city for these job types. The attractiveness is bound at 0.6 when the fraction is high so that these job types don’t cause exponential growth.

The x-axis for this graph shows values from 0 to 0.1. This is because the x-axis represents the fraction of people in finance or IT jobs compared to the number of households in a city. Because the number of IT and finance jobs in total is very low, only 4% of all people, this value in a single city will range from 0 to very high. Because this maximum value is entirely dependent on the type of people that life in a city (more IT/finance people in a city will exponentially bring even more IT and finance people), we have capped the attractiveness at 0.6 as explained before. The x-axis is then capped at 0.1 since the y value is stable anyway.



# Appendix B: Changing model input values

There are several input values and graphs that can be changed to test different scenarios. This appendix will summarize all the values that can be changed.

*BASE INPUT VALUES*

* Number of cities
* Number of households
* Runtime
* Warmup time
* Percentage initial households in and out of cities
* Minimum distance between cities
* Maximum distance between cities

*JOB GRAPHS (all job graphs can be altered)*

All job graphs can be altered, including the attractiveness of the jobless category which is the same for all cities.

*RESISTANCE TO MOVE*

The effect of ‘time since moving’, the number of children in a household and the age of the adult members in a household can be changed.

*ATTRACTIVENESS JOBS COUNTRYSIDE*

The attractiveness of all the job categories in the countryside can be changed.

*MOVING DISTANCE*

The range in which cities are close to the current location of a household and when they are too far away can be changed. The effect of these close by or far away cities on the city attractiveness can also be changed.

# MODEL\_SETUP.nls

## Households in City

*“to setup-households*

*let noOfHouseholdsInCity round (PercHouseholdsInitialInCity \* noOfHouseholds)*

*let noOfHouseholdsOutCity round ((1 - PercHouseholdsInitialInCity) \* noOfHouseholds)”*

This piece of code divides all the households between households in cities and out of cities. The percentage of households that are in cities can now be changed using a slider. This percentage is names “PercHouseholdsInitialInCity”. Initially the value was 0.4 (40%). The slider is can be changed from 0.0 to 1.0.

# MOVE\_CITIES.nls

## Resistance to move based on Time since moving + Children + Age

See Excel sheet named: RTM\_TimeSinceMoving

## Moving Distance modifiers for City Attractiveness

Here we alter the attractiveness of the city based upon the distance from the current location to that city. For cities very closeby (less than 100 patches) people will move less likely, since it is easily commuted daily. For cities very far away (more than 250 patches) people will move less likely since it is too far (from family etc.)

*“let minMovingDistance MinimalMovingDistance*

*let maxMovingDistance MaximumMovingDistance”*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Default** | **Min** | **Increment** | **Max** |
| **MinimalMovingDistance** | 100 | 0 | 10 | 200 |
| **MaximumMovingDistance** | 250 | 200 | 10 | 400 |

*“if distToCity < minMovingDistance [*

*set cityAverage cityAverage - MinDistCityAttractiveness*

*]*

*if distToCity > maxMovingDistance[*

*set cityAverage cityAverage - MaxDistCityAttractiveness*

*]”*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Default** | **Min** | **Increment** | **Max** |
| **MinDistCityAttractiveness** | 0.1 | 0 | 0.05 | 0.3 |
| **MaxDistCityAttractiveness** | 0.1 | 0 | 0.05 | 0.3 |

## Attractiveness Jobs Countryside

*“set jobAttractivenessList lput (list Job1Attractiveness Job2Attractiveness Job3Attractiveness Job4Attractiveness Job5Attractiveness Job6Attractiveness Job7Attractiveness) jobAttractivenessList”*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Default** | **Min** | **Increment** | **Max** |
| Job1Attractiveness | 0.55 | 0.4 | 0.05 | 0.6 |
| Job2Attractiveness | 0.55 | 0.4 | 0.05 | 0.6 |
| Job3Attractiveness | 0.50 | 0.4 | 0.05 | 0.6 |
| Job4Attractiveness | 0.45 | 0.4 | 0.05 | 0.6 |
| Job5Attractiveness | 0.45 | 0.4 | 0.05 | 0.6 |
| Job6Attractiveness | 0.5 | 0.4 | 0.05 | 0.6 |
| Job7Attractiveness | 0.5 | 0.4 | 0.05 | 0.6 |