CLUSTERING AS A RATIONAL DECISION OR A LUCKY ACCIDENT

Herd behaviour as an alternative explanation of the emergence of sectoral clusters using an agent-based model

Author: Gerben van Roekel

Human Geography – Utrecht University

University Supervisor: Dr M.J. Smit
Stadkwadraat Supervisor: C.K. Teoh

Date of submission: 1-8-2020

Abstract:
The formation of sectoral clusters has long been the focus of geographical study (Marshall, 1890, Porter, 1998). Generally, agglomeration externalities are credited as the prime cause behind this process (Delgado, et al., 2014, Diodato, et al., 2018). This theory is however far from uncontroversial (Martin & Sunley, 2003) with many scholars proposing other explanations (Klepper, 2010; Vicente & Suire, 2007). Conventional neoclassical explanations typically lack a perspective on the psychology of firms' decision-making (Berg, 2014). This study attempts to create an alternative perspective on clustering by incorporating herd behaviour as a decision-making strategy as a potential cause of clustering. An innovative agent-based model is used in which cluster formation is constructed from individual relocation decisions. In doing this, it connects individual relocation decisions to macro-level emergent patterns. This simulation shows that imitation has the potential to severely exacerbate the effects of agglomeration economies on clustering. The outcomes of this model challenge the dominant neoclassical view on clustering. An alternative hypothesis is postulated in which the interaction between herd behaviour and agglomeration externalities is proposed as an explanation for cluster formation.

Keywords: Herd behaviour, clustering, agglomeration, agent-based modelling, decision-making, relocation

Universiteit Utrecht

stad kwadraat

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

Over the last few months, I have worked hard to create this thesis and I can say that I am proud of the results. I have learned a lot about the topic, computer simulations and myself. I am also very grateful for the people who have supported me during the writing process. Special words of thanks are devoted to Chee-Key Teoh and Martijn Smit, who were my supervisors from my internship and university respectively. They have both contributed greatly with their advice on content, structure, and design. Their insightful comments on research proposals and draught versions were very valuable. Others whom I want to thank are prof. Eveline van Leeuwen, for her helpful comments on an earlier version of the model, the respondents to the interviews, my colleagues at StadKwadraat, and my fellow students.

1. INTRODUCTION

Suppose you are a company executive with the task to find a suitable new location within the city. How would you go about making this decision? And how would this decision impact the economic landscape of the city? The study presented here revolves around these types of questions. It delves into the individual firm's decision-making processes and assesses the impact they have on the spatial organisation of firms.

The analysis of both relocation decisions and the spatial distribution of companies have been pivotal topics in economic geography. This previous research often presupposes that relocation-decisions are made fully rationally, but deciding managers often exhibit non-economic and non-rational behaviour. Also, they lack information about the multitude of options, as well as the capacity to accurately process this information. Instead of rational analysis, decisions are often based on decision-making shortcuts such as risk-aversion, inertia, and imitation[1]. The role of imitation in decision-making has long been recognized in economics. J.M. Keynes (1935, Ch. 12§5) famously wrote that "it is better to fail conventionally than to succeed unconventionally", meaning that economic decisions can be more about conforming to a norm than actual sound business practices.

Even though this is widely accepted, its consequences on the distribution of firms are still obscure. One could suggest that imitation in relocation choice could lead to 'conventional failure', in which companies cluster on an inferior location. Because economic geography often assumes an analytical rigour which decision-makers do not have in practice, these decision strategies remain relatively under-researched, though some more behaviourally inclined researchers have presented alternative theories which do consider the psychological factors behind relocation decisions (Berg, 2010; Suire & Vicente, 2009). This is not to say that standard economic geography has no value, but that is perspective might be distorted by its assumptions. This strain of research has uncovered many facets of the dynamics of relocation choice and spatial distribution. It has uncovered what companies consider important and why they move. It also investigates the spatial patterns that form as a product of aggregate location choices (Balbontin & Hensher, 2018; Klier, 2006; Nguyen, et al, 2012). These patterns show that companies often cluster together but it is not immediately obvious why these spatial patterns exist in the places where they do. Often these patterns are not directly attributable to local properties but have come to exist as a result of complicated and path-dependent processes (Boschma & Frenken, 2006; Brekke, 2015; Elola, et al., 2012; Vicente & Suire, 2007). This makes predicting the location of geographic concentrations of firms ahead of time very challenging. Consequently, local government initiatives to create such thriving locales with policy incentives, usually have limited success. (Graf & Broekel, 2020; Wolman & Hincapie, 2014). In this study, I try to contribute to the understanding of the formation of clusters, by broadening the perspective on company relocations by introducing the concept of herding behaviour from behavioural theories as an alternative explanation for cluster formation.

This research aims at incorporating decision-making strategies under bounded rationality and uncertainty into the body of research regarding form relocation and clustering. Specifically, the impact of herding behaviour in location choice on the spatial organization of firms is investigated. Traditionally, clustering is explained as an economic force: an observed concentration of firms in a given sector, assuming rational decision-making, can be explained in two ways. Either the cluster location is excessively superior to other places or there are economic benefits to colocation (Malmberg & Maskell, 2002). In this study, I examine whether herd behaviour, theoretically, can function as part of the explanation for the emergence of companies. The research question studied is:

> *To what extent can herd behaviour in location choice processes be a cause for spatial clustering of related office firms on an urban scale?*

There are strong theoretical indications that herd behaviour plays a strong role in economic decision-making, especially in a complex environment with heterogeneous options and limited information (Armstrong & Huck, 2010; Baddeley, 2010). Location choice is such a decision. This study tries to assess how herd behaviour could affect spatial patterns using a computer simulation. The purpose of this exercise is to generate well-founded

[1] For the sake of readability the terms herding, herd behaviour, imitation and mimetic behaviour is used interchangeably and with the same meaning in this thesis.

hypotheses about the formation of spatial clusters and expand the existing theory on clustering with insights from bounded rationality. This study does not prove the presence of herd behaviour in location choice, but instead postulates and makes plausible some hypotheses about the impact of herd behaviour in the formation of spatial patterns. Thus, encouraging further (empirical) research into this topic.
In this research design, I chose to make two confinements to the scope of the research. Firstly, as an object of study, I chose office firms because they do not have a relatively simple production process. Office firms can move relatively easily compared to industrial firms and they can relocate to more places. Secondly, I chose to limit myself to the urban scale, because most relocations happen over small distances and the initial conditions are more equal.

Understanding clustering requires understanding the location choice process of the company as well as the behaviour of the system as a whole. The latter is not trivial as it is often markedly different from the simple aggregation of individual decisions. The micro-level choices of individual companies shape the macrolevel of locational patterns, but the causes of individual decision-making cannot be deduced from the aggregate pattern and neither can the aggregate patterns be induced from individual decision-making. This is because location choices are not made independent of the particular distribution of companies at that point. Researching the effect which herding behaviour on the micro-level can have on spatial patterns on the macrolevel, requires bridging the gap from the micro to macro-level. This is done using an agent-based model (ABM). This is a useful method to describe and explore interactions between individual behaviour and aggregate macro-level patterns, especially in complex social systems (Ajraldi, et al., 2011, Bonabeau, 2002, Zhao, et al., 2011). In ABM, the researcher programs individual behaviour and explores the impact it has on macro-level patterns. The benefits of ABM in this study are fourfold. First, it allows for easy exploration of the dynamics of a system without the necessity of an elaborate empirical study, which would be particularly cumbersome in studying commercial herd behaviour. Secondly, ABM sets up a behavioural model in which boundedly-rational and stochastic decision-making can be easily incorporated. Thirdly, ABM is especially useful in understanding the formation of complex systems out of individual behaviours. Fourthly, it is possible to run models multiple times to see the diversity in possible outcomes from the same starting situation. This study provides a more nuanced understanding of the formation of spatial patterns and the role bounded rationality can play in explaining them. The insights used to formulate an answer to the main research question are established in two sub-questions:

> *1: To what extent could clustering patterns be explained by herd behaviour dynamics instead of or in addition to agglomeration externalities?*

Agglomeration externalities are the most common explanation of spatial clustering in places which are not clearly superior (Delgado, 2014; Diodato, et al., 2018). They relate to the idea that there are concrete economic benefits to spatial proximity to other companies. If the assumption of rationality is abandoned, other non-economic forces might play a role in explaining the formation of clusters. Herd behaviour has the potential to explain concentration because it means that companies have a predisposition towards places with other related companies. For this sub-question, a model is set up to explore to what extent imitation can explain the geographical concentration of firms. In doing this it combines behavioural economic theory and the psychology of decision-making with geographical theories of firm relocation and clustering formation. This question infers how the degree of clustering is affected by herd behaviour and pairs this with the impact of agglomeration externalities.

> *2: How predictable are spatial distribution patterns of firms in a herding-based location choice process?*

This research question delves into the consistency of the model; that is, how much do the outcomes of different runs of the simulation vary. The model is predictable if it consistently produces the same results. It is known that the locations of clustering are particularly hard to predict because clustering tends to be a highly path-dependent process (Maskel & Malmberg, 2007). Going back through history, the steps which lead to the formation of a cluster can be analysed, but ex-ante predicting which patterns will form from a given starting point is very complicated. That means that it is in practice very difficult to assess how likely the current distribution of firms was to develop. This question answers the questions whether the predictability of the model changes if herding behaviour is present. It inquires how large the range of likely outcomes is.

Together the answers to these questions show the effect of herd behaviour on geographical patterns. This is both scientifically and societally relevant. Scientifically, this study broadens the perspective of clustering and adds a behavioural approach. Oftentimes, researchers only review individual relocation decisions (e.g. Bodenmann & Axhausen, 2012; De Bok & Van Oort, 2012; Nguyen, et al., 2014) or the spatial distribution of companies (Delgado, 2014; Diodato, et al., 2018; Duranton & Puga, 2004), without establishing how these interact. This results in an incomplete view of a singular system. This study tries to overcome this by bridging the conceptual gap between micro-level behaviours and macro-level patterns. The current understanding of both levels is somewhat inaccurate. Studies on clustering and relocation tend to ignore how companies make decisions, because rational, economic optimization is assumed (Berg & Gigerenzer, 2010; Boschma & Frenken, 2006), though there are some exceptions (e.g. Berg, 2014; Brunes, 2005; Suire & Vicente, 2009). Assuming rationality makes research easier, but behavioural researchers have established that boundedly-rational decision-making strategies may be crucial in understanding economic phenomena (Armstrong & Huck, 2010; Baddeley, 2010). This thesis tries to adjust the theory on cluster formation by introducing a bounded-rational herd behaviour approach.

The societal relevance of this study relates to its policy implications. Many local governments have tried to establish cluster locations, often with limited success (Graf & Broekel, 2020; Wolman & Hincapie, 2014). This shows in which situations clusters can come to exist and how predicable that is. Thus, it can be used to encourage or discourage certain policies. Because it does not imply rational optimization it can also be used to assess the potential impact of policies which invoke decision-making psychology such as marketing.

The structure of this thesis is as follows. Firstly, the relevant scientific theory is discussed and integrated into a conceptual model. Secondly, this conceptual model is refined into a research method, which is described in detail. This produces the agent-based model. The results of which are analysed in the fourth section. The fifth section includes a conclusion and discussion of the results.

2. THEORY REVIEW

This thesis is founded on a theoretical foundation which is used to design and interpret the model. This foundation is structured in a conceptual model. This conceptual model features the integration of different insights from both neoclassical and behavioural economics, organizational psychology, the geography of clustering and relocation choice studies. The purpose of this section is to create a basis for the development of the agent-based model and a framework for its interpretation.

Firstly, the traditional, neoclassical approach to relocation choice is brought forward. This is then contrasted with theory from behavioural economics and organizational psychology. This perspective is used to formulate a behavioural alternative to neoclassical the location choice theory. The study of individual relocations describes the micro-level dynamics of the system. Besides the individual location choice, other researchers have devoted their work to understanding the patterns these relocations cause. This is the macro-level of the system. It explains why companies tend to concentrate in certain places.

2.1. Neoclassical theory of location choice

Traditionally, location choice studies assume that relocation decisions are a rational optimization issue. It implies that companies gather and process information on all possible locations for all aspects which impact the business operation. Analysis of this information leads the company to the optimal location for their business. This is an application of the assumptions of classical economic theory.

2.1.1. Neoclassical economic theory

Neoclassical economic theory has long been the dominant perspective on economic development. Essential to this theory is the use of mathematical expressions to determine market equilibria in the economic landscape. These models are useful because they can easily project how an economic system works. The neoclassical models used in economics are based on the classical assumptions of the economic discipline. Among these are fully rational actors, complete information, profit maximization and self-interest (Baddeley, 2010; Dugger, 1979). These assumptions are necessary to make the models function. Neoclassical economic models are also often used to explain geographical phenomena, such as relocation and clustering (Boschma & Frenker, 2006; Hassink & Gong, 2019). These attempts are often criticized by geographers, who claim that economic methods in the absence of a geographical perspective are not suitable to research economic-geographical events (Berg & Gigerenzer, 2010; Boschma & Frenken, 2006).

An alternative model which explicitly introduces a spatial perspective in the study of firm locations is Krugman's (1991) New Economic Geography. This theory uses economic and geographical perspectives to explain regional disparities in economic development (Hassink & Gong, 2019). Empirically, this theory is sounder than classical economics (Fingleton & Fischer, 2010) but it still an economic modelling approach featuring many of these classical assumptions (Boschma & Frenken, 2006; Hassink & Gong, 2019; Martin & Sunley, 1996).

This is why many behavioural researchers have expressed their concerns. They argue that even though these assumptions make the models manageable, the also strongly reduce their utility (Berg & Gigerenzer, 2010; Simon, 1955). Nevertheless, many researchers still use some version of the neoclassical assumptions to research clustering and location choice, because using behavioural methods makes it very difficult to develop a general theory of location choice. Assuming some form of rational optimization is necessary for most statistical research methods. The main findings of these studies are discussed in the next section.

2.1.2. Explaining location choice

Under these assumptions, motives for location choice can be induced from organizational and spatial characteristics using statistical analysis on data of commercial relocations. These revealed preference studies are the main body of our current understanding of relocation choices (e.g. Barrios, et al., 2006, Bodenmann & Axhausen, 2012, Brouwer, et al., 2004; De Bok & Van Oort, 2011; Nguyen, et al. 2012; Van Dijk & Pellenbarg, 2000). Revealed preference models investigate actual relocation decision data and try to find correlations between this data and locational characteristics using statistical analysis. If these are found, it is expected that the companies' location choices are produced by the favourable locational characteristics. An alternative research method uses stated preference data (e.g. Elgar et al., 2009, Willigers & Van Wee, 2011). These studies ask firms on what basis they made or would make a location choice. This provides insights into the factors which companies deem important.

Determinants of location choice may be organizational or spatial in nature. There are firm-specific characteristics such as age, size, and growth rate which impact relocation decisions (De Bok & Van Oort, 2011; Nguyen et al., 2012). In general, older and bigger firms are less inclined to move. Growing or shrinking firms are more likely to move to match the size of their accommodation to the size of their enterprise. Both revealed and stated preference studies show similar spatial determinants of location choice for office firms at the urban scale: accessibility, amenities, availability, and price.

Besides the actual office, the prime production factor in an office firm is labour. For other sectors, proximity to materials or customers may be important, but for office firms the dependence on labour is quintessential. This means that a location close to a large, qualified labour pool is highly valued (Bodenmann & Axhausen, 2012; Keeble & Nachum, 2002). Consequently, office firms are often located in cities. Accessibility means that a location close to major mobility infrastructures is also preferred, so that staff can arrive easily by transit, car or bike (Balbontin & Hensher, 2019; De Bok & Van Oort, 2011). Workers are also thought to like places in attractive environments, such as near a natural area or in a vibrant neighbourhood, and near urban amenities. Employees want to be able to enjoy, use and experience the area in their personal time (Glaeser, et al., 2001). Urban location choice is also fundamentally shaped by availability. One can only move to places where an office is available or can be built. Over time areas with high demand for office spaces develop higher densities, either because each workplace gets smaller or expansion takes place by building higher or closer together. As such, the capacity of an area is not static over time (Anas, et al., 1998; Bodenmann & Axhausen, 2012). Places in high demand are also more expensive, which is the next major element in urban location choice (Nguyen, et al., 2012; Van Dijk & Pellenbarg, 2000). Companies typically seek to pay as little as possible. The price effect is converse compared to the previous factors. Companies are not per se attracted to cheap places but drawn away from high prices. Relocating firms balance the qualities of a given location with the associated price. They rarely opt for a very cheap place with no qualities because a low price in itself is not a quality of the land. Additional costs occur with relocation, such as the potential loss of staff. These have to be compensated by the difference in profitability between the new and old location (Brouwer et al., 2004; Van Dijk & Pellenbarg, 2000). Of course, it is possible to identify a multitude of other factors of minor importance or which affect only some companies. Generally, however, the factors mentioned can be considered the most important spatial determinants of office location choice for almost all companies.

2.2. Decision-making in behavioural economics

In the previous section, different perspectives on firm relocation were posited from a neoclassical perspective. The neoclassical economic approach to corporate decision-making has the problem that it does not reveal the internal decision-making mechanisms in the firm: decisions are assumed to be rational optimization questions (Boschma & Frenken, 2006). Within the standard assumption of economic theory, a perfectly rational, well-informed economic agent would evaluate all possible paths of action on the effect that they have on a clear set of defined company goals (typically profits) and then choose the one with the highest rewards (Baddeley, 2010; Cyert & Hendrick, 1972, Simon, 1955). Other branches of the economic discipline challenge this way of thinking. The behavioural theory has less rigid assumptions and does not presuppose rationality, complete information, and profit maximization strategies. In this section, I discuss the various ways bounded rationality and behavioural economics challenge the assumptions and outcomes of neoclassical theory. First, the background of behavioural economics concerning decision-making is briefly discussed and then applied to the relocation decisions in the next chapter.

2.2.1. Behavioural economics

The integration of insights from psychology and management studies regarding decision-making processes into the economic discipline has been a cornerstone in the development of behavioural economics. A pioneering article by Simon (1955) lays the groundwork for incorporating boundedly-rational decision-making into a classical economic framework. This included several new paradigms. Firstly, the scope of action is limited for economic agents: they only review a limited subset of all possible actions. Secondly, humans lack the competency to perform complicated probabilistic and combinatorial computations for accurately predicting the consequences of their actions. Thirdly, decisionmakers look for satisfactory and not optimal solutions. All actions which result in a satisfactory outcome are considered favourable. These theoretical contributions have been deductively and experimentally established by subsequent research (e.g. Conlisk, 1996; Ellison, 2006). Recently Armstrong & Huck (2010) and Baddeley (2010) have reached similar conclusions by empirical research methods. Simon's approach still assumes a singular company which decides in isolation. Further research has expanded on this view, because "[s]ociological forces interact with psychological forces and will affect individual behaviour if groups act in concert without any clear coordinating mechanism" (Baddeley, 2010, pp. 284-5). It is therefore important to discuss not only the decision of one company in an isolated context but also the economic and social context in which a decision is made. Other economists have included elements of interaction in the theories of behavioural economic theory (Conlisk, 1996, Rook, 2006).

2.2.2. Characteristics of boundedly-rational decision-making

Following this behavioural approach, several concepts can be introduced to explain the way corporate actors make economic decisions. These concepts, borrowed from organizational sciences and psychology, relate to the process of decision-making itself rather than any particular outcome.

The portrayal of firm decision-making in neoclassical economics is a simplification which can have far-reaching impacts for analysis. For this study, where firm decision-making is key, firms are portrayed differently. In this section, several factors in corporate decision-making that deviate from neoclassical economics are discussed (cf. Cohen et al., 1976, for an overview of examples and generalizations of the way behavioural economics differs from neoclassical economics).

In this framework, I distinguish several characteristics of firm decision-making processes that deviate from standard rational economic theory, and which are foundational for the decision model presented in this thesis.

1. Managerial myopia

Managerial myopia refers to the phenomenon that managers are overly focused on short-term goals. Myopia can have two sides. First, managers are mostly held accountable for the short-term effects of their actions. Secondly, short-term effects are easier to estimate. Managerial myopia can lead to the advancement of short-term over long-term interests and can, therefore, have detrimental effects on a firm's long-term performance (Edmans, 2009, Larwood & Whittaker, 1977 and Stein, 1988).

2. Satisficing

Managers are satisficers and not optimizers: they are not looking for the best possible solution, but one that fits their requirements (Armstrong & Huck, 2010, Berg, 2014). Companies are not devoting their resources to find the best possible decision, but instead, every option that satisfies all economic and non-economic demands of the firm are considered favourable. Decision-makers have implicit or explicit goals which serve as thresholds in this binary way of thinking.

3. Internal dynamics and managerial intent

A firm is not a single undivided entity. The intent of the decision-makers may not always be in line with expressed company goals (Cyert & Hendrick, 1972). Some economic decisions may be aimed at other goals such as managerial self-interest (Haynes, et al., 2015), ethics (Strudler & Curlo, 1996), social responsibility (Rao & Tilt, 2016), or the advancement of the family in family firms (Chrisman, et al., 2014). Regarding the firm as a single undivided unit is useful in analysis but obscures the role of internal dynamics and the role of managers and non-economic goals.

4. Rigour

Boundedly-rational managers are not able to accurately predict the outcome of their decisions, nor do they try to. Deciding managers most likely do not conduct a rigorous probabilistic analysis into all possible decisions and their associated expected outcomes. They review a small number of potential actions and decide based on their experience and intuition (Lejerraga & Martinez-Ros, 2008; Slovic et al., 2004). The rigorous analysis assumed by neoclassical economics is absent in most regular economic decisions.

5. Risk aversion

If the outcome of an important decision turns out to be adverse, deciding managers may face career consequences, even if that particular outcome was initially unlikely to happen. Deciders might, therefore, opt for a 'safe' decision over a more profitable decision with a higher risk (Milidonis & Stathopoulos 2014).

6. Inertia

In decision-making processes, the default option is usually the current practice. This works in two ways: first companies do not consider new possible courses of action if the current way of doing things still works satisfactorily. Secondly, decision-makers may be inclined to view new ways more sceptically, while having an internal predisposition towards 'doing things the way they are always done' (Alós-Ferrer et al., 2016; Kaplan & Henderson, 2005, Van Witteloostuijn, 1998). This psychological feature can significantly hinder companies to make future-oriented, innovative, and sound business decisions because they instead cling to the old ways.

2.2.3. Imitation

In addition to these characteristics of economic decision making, specific attention is drawn to imitation, as it is the main topic of this thesis. Corporate decisions are not made in isolation. Many other economic actors have made a similar decision previously. Herd behaviour (also known as mimetic behaviour of imitation) is defined by Baddeley (2010) as *"the phenomenon of individuals deciding to follow others and imitating group behaviours rather than deciding independently and atomistically on the basis of their own, private information"* (p. 281). Decision-makers infer from the actions of others which decision is good enough. If many others make the same choice, then it is probably a sound business decision (Baddeley, 2010, Banerjee, 1992, Bikhchandani et al., 1998). Moreover, managers who make the same decision as their rivals are perceived as wise and are not as stringently scrutinized as when they had made a divergent decision. When a decision turned out to be detrimental, the consequences are milder if most other firms made the same mistake (Armstrong & Huck, 2010). Herd behaviour is therefore not only possible but also expected in corporate contexts, especially in decisions which involve many complex and heterogenous options (Armstrong & Huck, 2010).

Early theorists such as Banerjee (1992) and Bikhchandani, et al., (1998) show how limited information supply and sequential decision-making can logically lead to imitation. They show that herd behaviour emerges, and is even rational, when the probability of making a right decision is higher if one does what everyone else does, compared to following one's own information. Many other researchers have shown empirically that imitative behaviour plays a role in a wide range of economic decisions (Cipriani & Guarino, 2009; Koetsier & Bikker, 2017; Luo & Zin, 2011; Seiler et al., 2012), including location choice (Berg, 2014; Suire & Vicente, 2009, Vicente & Suire, 2007).

However, the utility of herd behaviour decreases according to the share of actors engaged in it (Bannerjee, 1992, Bikhchandani, et al. 1998, Lux, 1995). This can sometimes, but not necessarily lead to a more volatile long-term outcome, and the formation of economic bubbles (Lux, 1995; Suire & Vicente, 2009; Zhao et al., 2011). A factor that further complicates matters is the imperfection of herd behaviour itself: imitative behaviour is similar, but rarely identical (Posen et al., 2010). Indeed, humans often take the wrong lessons from observing others. Weizsäcker (2010) finds that "*in situations where the evidence conveyed by others' choice is less clear, the information is not only less valuable, but the participants are disproportionately worse in making the correct inferences"* (p. 2357). As a result, even in if imitation is theoretically justified, actors always appreciate their own information. Herding behaviour is thus often not blindly following others, but a comparison of one's own information with others' actions (Cipriani & Guarino, 2009). It is unlikely that many economic actors base their actions solely on others' behaviour altogether, but it is very plausible that herd behaviour shapes and fashions all sorts of economic behaviour in some way. An actor's decision is structurally formed by the decisions that others have made before his decision, especially in a complex decision-making landscape; if not for the positive impact signalled by others' decision, then for the way these shape an actor's perspective on the possible options.

Behavioural economics shows that companies' actions are not only formed by clear economic motives, but also by the psychology and decision-making capacity of deciding actors.

2.3. A behavioural view on location choice

If the assumption of rational optimization is abandoned, it is possible to identify several different ways in which decision-making can impact relocation. Firstly, companies have a small search radius. The distance to the owner's house is among the strongest predictors of business location choices (De Bok & Van Oort, 2011; Elgar, et al., 2009). This means that the owner's personal preferences interfere with his business operations. They are unwilling or unable to expand their search scope: signs of myopia. Secondly, companies only review a very limited set of possible locations before making their choice (Berg, 2014; Elgar, et al., 2009). Berg (2014) finds that the modal number of considered possible locations is three. Moreover, which sites are reviewed is also a matter of intuition and chance (Berg, 2014). This is not the rigorous optimization that neoclassical theory assumes. Because decision-makers are satisficers, the smallness of the consideration set is not problematic: Any location with a perceived attractiveness above a certain mental threshold is considered good enough (Berg, 2014, Elgar, et al., 2009; Van Dijk & Pellenbarg, 2000). A final crucial effect of decision-making psychology on relocation decisions is the desire to stay put (Van Dijk & Pellenbarg, 2000). This is a product of myopia, inertia and risk aversion. Companies typically only move when it has become almost necessary.

However, the main impact of decision-making strategies on relocation studies comes from herd behaviour. Herding is a common phenomenon in economic decision making. It is present in all types of choices, including location choice. Berg's 2014 study finds that 84% of companies engage in herd behaviour in their relocation search. This does not only affect the relocation of individual companies but on a larger scale, it is imaginable that this can severely alter the economic landscape. Herd behaviour can play a large role in relocation decisions because location choice is a sequential, cumulative decision-making process, where the possible outcomes are heterogenous, the decision is complex, and the decision-makers have limited information (Berg, 2014; Brunes, 2005; Elgar, et al, 2009).
The literature on herd behaviour and its impact on location choice is currently rather limited, though there are some studies which reflect on its relevance. These can be classified into two lines of reasoning. The first, proposed by Berg (2014) says that firms exhibit herd behaviour in the sense that they are more likely to know about and look for possible sites near other related companies. Other locations without a presence of related companies are much less likely to be even considered, regardless of their qualities. The second line of reasoning is put forward by Suire and Vicente in several articles (e.g. Suire & Vicente, 2009; Vicente & Suire, 2007, Vicente, et al., 2007). It revolves around the idea of the 'locational norm'. This means that a location can become the default location in a certain industry. In such a situation deviation from the norm is considered a strategic risk. Companies then infer that the norm location is optimal because so many other companies are there already. Besides, customers and business partners place companies located in this place in higher regard, because they use the firm's location as a proxy for its quality.

A behavioural view on location choice does not result in a deterministic predictive equilibrium model, as neoclassical models do, but instead describes the choice behaviour of decision-makers and the strategies they have. It can, therefore, be a valuable addition to the standard neoclassical relocation models, which discard the effect of the decision-making process.

2.4. Spatial patterns

The collection of individual (re)location decisions creates a distribution of firms across geographic space. Any patterns in the distribution of firms are the result of (re)location decisions and by extension of the factors that shape relocation decisions. However, it is quite common that spatial concentrations of firms do not match these criteria. High concentrations of firms are often located in places that do not make intuitive sense. Many scholars have investigated this topic and come up with explanations (Boschma & Frenken, 2006; Brekke, 2015; Elola, et al., 2012; Vicente & Suire, 2007). How spatial patterns have evolved is the focal topic of evolutionary economic geography. This branch of economic geography investigates why concentrations of economic activity emerge (Boschma & Frenken, 2006). Evolutionary geographers show that, because the existing distribution of firms affects the location choice of moving companies, concentrations of companies do not always form on the logical spots. Places with a concentration of companies seem to attract more companies than similar location without clusters. When a concentration of companies reaches a certain size, it develops a critical mass, which sets the clustering effects in motions. It creates a virtuous cycle where the attractiveness of the cluster increases with its size (Menzel & Fornahl, 2009, Porter, 1998). Before this critical mass forms, the economic landscape might develop in many different directions. After the critical mass has formed, the system locks into a specific path. The accumulation of historical decisions creates a situation which is maintained through future decisions (MacKinnon, 2008; Maskell & Malmberg, 2007). This path-dependency creates a complex, circular-causal system, where the geographic patterns of firm location cannot be trivially traced back to the initial properties of space.

Researchers have come up with multiple reasons why companies tend to concentrate. The most prominent of these revolves around the concept of agglomeration externalities, but this is not without criticism.

2.4.1. Agglomeration externalities: Localization and urbanization economies

Alfred Marshall (1890) observed that certain places have a remarkable concentration of a certain industry, which could not wholly be explained by some natural quality of a place. To explain these concentrations, he introduced the concept of what is now known as Marshallian externalities: In many industries, there are agglomeration advantages which are the product of economies of scale. Sectoral agglomeration allows for sharing of services, a matching labour market, and the spread of tacit knowledge through social interaction (Diodato, et al., 2018; Duranton & Puga, 2004). As a result of this, a company in a specialized location can produce a product with higher quality and lower costs (Diodato, et al., 2018; Marshall, 1890). This means that a spatial concentration in a certain sector may occur if that yields agglomerative benefits. If colocation offers significant economies of scale it is, from a classical economic standpoint, inevitable that clustering eventually happens. As such, Marshallian externalities create industrial agglomerations far larger than any natural quality of the land would justify (Ellison & Glaeser, 1999; Ellison et al., 2010).

There is debate on the question whether these externalities are sector-bound or also straddle across sectoral boundaries. This discussion is framed as localization versus urbanization economies (Beaudry & Schiffauerova, 2009; Glaeser, et al., 1992). Those who propose the localization economies follow Marshall’s line of reasoning and state that agglomerative forces primarily make colocation attractive for firms within the same sector, whereas the theory of urbanization economies, put forward by Jane Jacobs (1969) states that there are major economic benefits of colocation when there is a diversity of sectors in a place, as this provides opportunities for the exchange of ideas (Beaudry & Schiffauerova, 2009; Glaeser, Kallal, et al., 1992). Empirical research into the topic reveals mixed results, with some studies indicating a larger role for localization and others for urbanization economies (Beaudry & Schiffauerova, 2009). In general, however, there are strong indications that agglomeration forces in both or either form strongly affect regional and urban economic development (Delgado et al., 2014).

2.4.2. Criticism and alternatives

Even though agglomeration externalities are widely used as an explanation for geographic clustering, they are far from uncontroversial. There has been significant criticism on everything from its definition, its research methodology to its empirical foundations (Duranton, 2011; Martin & Sunley, 2003). The concept is typically not very well defined, and the theory is difficult to prove empirically (Malmberg & Maskell, 2002; Markusen, 2003). Malmberg & Maskell (2002) consider the theory of agglomeration externalities logical, because:

> *[L]ocalized clusters do exist and this may legitimately make us assume that such a spatial structure is in some sense efficient or rational. […] At the same time, it has turned out to be extremely difficult to identify empirically the mechanisms that are supposed to account for its existence.* (p. 442)

They find that persuasive prove for agglomeration is lacking. Others such as Boschma (2005), Elgar and Miller (2009) and Gordon and McCann (2005) have also indicated that the impact of agglomeration externalities on the spatial organization of firms is considerably overestimated. Following this debate, alternative explanations have formed. One such alternative is the spin-off theory, which is popular in evolutionary economic geography (Boschma & Frenken, 2006; Buenstdorf & Klepper, 2007; Klepper, 2010). This theory states that successful entrepreneurs almost always have prior experience in their industry. Because people leave a company and start their own related business in the same area, clusters can form. Moreover, these spin-off companies tend to outperform non-spinoff companies because they inherit successful practices from the parent company.

Another alternative is the locational norm, discussed previously. In this concept, clustering is not the result of agglomerative externalities, or any real economic benefit whatsoever, but stems from the psychological inclination to follow the crowd (Suire and Vicente, 2009). In this view, clustering emerges following a 'locational cascade': one company's management makes an independent location choice, which is a signal to others that this is a superior location. Other companies may locate there as well because this locale is apparently good enough for other companies. This creates a default location for the sector. Managers and entrepreneurs making a location choice would need a strong independent signal for another location to deviate from this norm. This process could potentially lead to a sectoral concentration on a location which is not exceptionally advantageous.

2.4.3. Understanding and predicting clusters

Regardless of the exact dynamics, it has been well established that there are forces that drive companies closer together beyond the general attractiveness of a place. It has however proven difficult to develop a convincing general theory to describe this process. Multiple overlapping and contradicting theories attempt to explain (part of) the concentrating forces (Breschi & Malerba, 2001). These theories generally lack the empirical foundation to be general, all-encompassing explanations of clustering, in such a way that they cannot explain the formation of each cluster accurately. Instead, studies often rely on case-by-case examinations or general, non-empirical reasoning (Maskell & Malmberg, 2002; Martin & Sunley, 2003). Explaining cluster formation, therefore, requires extensive study of each case. Developing a general theory of why clusters formed in the places where they have, is quite hard. Predicting where clusters will form given certain spatial characteristics is of another order of magnitude (Graf & Broekel, 2020; Maskell & Malmberg, 2007). In the absence of a general, predictive theory, ex-post analysis of cluster formation on a case-by-case basis is, therefore, the method of choice in explaining clustering processes.

This complexity regarding clustering and location choice dynamics makes it difficult to create effective policies on local economic development and explain why government attempts to artificially create clusters often fail (Graf & Broekel, 2020; Wolman & Hincapie, 2014). Though in the right circumstances, enhancing clusters can be effective, but only if there is a good economic basis to implement such policies (Sydow, et al., 2010).

2.5. Conceptual model

The purpose of this thesis is to contribute to the unravelling of the dynamics between firm locations choices and spatial patterns of economic activity. This can be separated into two distinct levels of aggregation. On the micro-level, this is the decision of individual companies, whereas the macro-level refers to the spatial distribution of companies and its impact on the economic landscape. The theory overview above provides a basic overview of the relocation decision of companies and the associated geographical patterns.

2.5.1. Microlevel behaviour

Neoclassical economics assumes that, given a set of inputs, firms always make the same, optimal decision about a specific economic goal. Considering decision-making as simply a rational computation of the best possible solution is a simplification that does not do justice to the way companies make decisions. Managers and entrepreneurs decide boundedly rational. They do not make decisions randomly or irrationally, but the rationality of their consideration is severely limited by the imperfection of the human mind and the scarcity of information. Decision-makers take short-cuts and make errors. Concerning location choice, this means that managers do not perform locational optimization, but a simplistic evaluation on just a few locations and a just a few elements they deem important. Thus, the location choice is formed by the locational demands companies have, but also by the process of deciding.

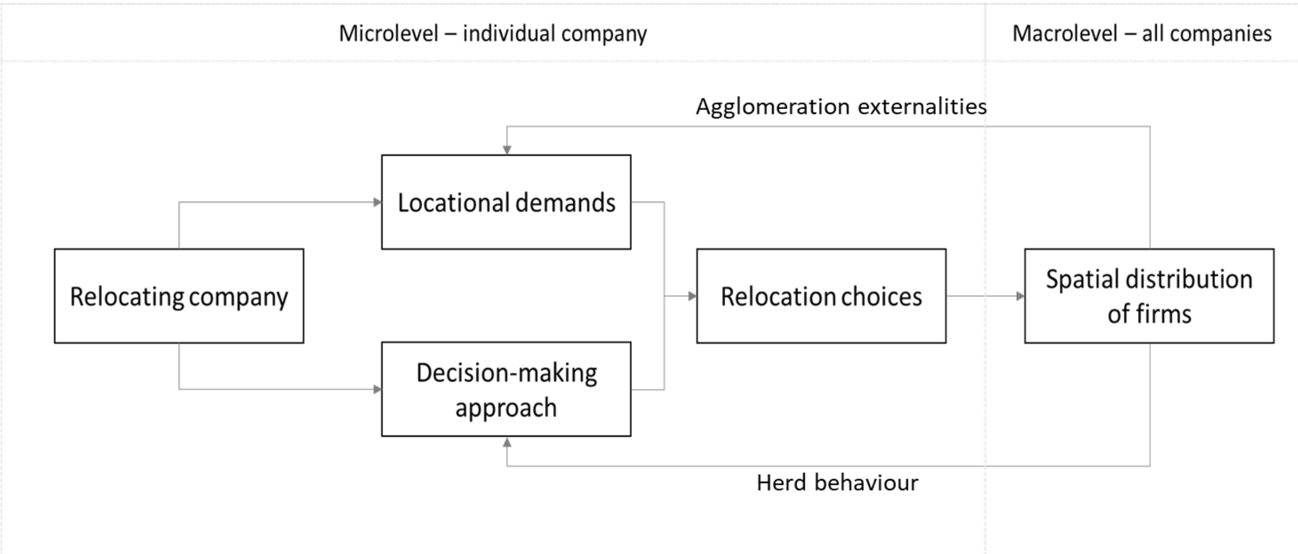
2.5.2. Macrolevel

Companies tend to be clustered in specific places. Sometimes this is because these places offer great business opportunities. In some other situations, the intrinsic qualities of the location do not seem to justify the number of businesses observed. There have been several suggested reasons for this. Many of these relate to the benefits of concentration and the presence of path-dependency. When concentrations of companies increase in size, they also grow in attractiveness as a place for relocation. However, the precise dynamics of this process have not yet been fully uncovered. The heterogeneity of space makes it difficult to make good comparisons between places. These macro-level patterns are the product of the accumulation of micro-level decisions. However, these macro-level patterns also in turn affect the microlevel decisions. It is therefore not possible to consider the macrolevel to be simply the addition of independent individual decisions. Because of the feedback effects of the macrolevel into the microlevel, a virtuous cycle is created in which the attractiveness of a location increases with its size, and its size increases with its attractiveness.

2.5.3. Conceptual model

This is structured into a conceptual model which shows the integration of the individual concepts described in this chapter. Figure 2-1 shows this conceptual model.

Figure 2-1: Conceptual model



This model shows a company which makes a relocation decision. This choice is formed by two factors: how the decision is made (decision-making approach), and on what basis the decision is made (locational demands). This choice than contributes to the macrolevel pattern of the spatial distribution of firms. This distribution, in turn, further affects the locational demands and the decision-making approach, through agglomeration externalities and herd behaviour respectively, creating a circular causality or feedback loop. This model is further developed into an algorithmic operation model in the next chapter.

3. METHODOLOGY

The conceptual model described above is researched using an agent-based model: a computer simulation in which the individual actors' behaviour is precisely programmed using a small number of behavioural rules. In the simulation, the actors interact with each other and the environment and the resulting macro-level patterns is investigated. These patterns are not modelled itself but stem from individual behaviour. In this chapter, I first explain what ABM is and why it is useful in this study. Secondly, I discuss the decisions and assumptions made in the design of the model. Thirdly, I explain the operation of the model in detail. Fourthly, the way the data is generated and analysed is put forward. Fifthly, I describe the validation process of this research.

3.1. Agent-based modelling

Scientific modelling is a way to report one's understanding of reality. Because reality is usually rather complicated, a model is a stylized formulation of the main relations and assumptions. A scientific model is a simplified and codified representation of the real world for purposes of both research and understanding. Some models simply describe what happens in a process and why, whereas other models also make predictions about what might happen. Developing an accurate predictive model on cluster formation has proven difficult (Maskell & Malmberg, 2007). The model presented here is not aimed at precisely describing relocation choice processes, but instead at showing the importance of incorporating herd behaviour in predictive clustering models.

Using Equation-Based Models (EBM) is a well-established practice in science. A phenomenon is then represented as a set of interacting equations. These equations can be used to model how a phenomenon develops over time, such as predicting equilibrium values. Agent-Based Modelling (ABM) is different from EBM in that it does not model the system as a whole, but instead, it models the behaviour of individual actors. ABM programs behaviour using computer algorithms instead of mathematical equations. It is particularly useful in complex systems with many distinct actors. Complex uncoordinated system patterns can be developed using simple behavioural rules on the individual level (Wilensky & Rand, 2015).

In this study, the ABM is used to model the location choice behaviour of individual companies and see how that impacts spatial patterns of location choice. The objective is to explore whether and how decision-making strategies can lead to spatial clustering. This means that instead of using empirical data on herd behaviour and clustering, a computer simulation is made, which shows the general insights from the conceptual model.

For a thorough general explanation for the use and relevance of ABM as a scientific research method see Wilensky and Rand (2015), and Torrens (2010) for a discussion of its use in the geographical sciences.

3.1.1. Fundamentals of ABM

Agent-based modelling is a research method and a tool for understanding the dynamics of complex systems. ABMs focus on exploring the dynamics between micro and macro-level behaviour. ABM is associated with such systems because it separates individual and system-level behaviour. In their seminal introduction to ABM, Wilensky and Rand (2015, pp. 32-36) identify five fundamental principles of ABM:

1. ABMs are computational. They are computer simulations and therefore require the researcher to frame agents' behaviour in terms of computational algorithms.
2. ABMs are based on the structure of natural thinking and not of mathematics. An ABM is formulated in such a way that it logically follows a line of reasoning of an agent. This makes ABMs easy to understand and the dynamics of the model less obscure.
3. ABMs are useful in situations where complexity is present. This means that aggregate behaviour does not logically follow from the aggregation of individual behaviour. The complexity of a situation makes it difficult to tie individual actions to system-wide patterns. The patterns emerge from the individual behaviour, without any actors specifically aiming to create the pattern.
4. ABMs match the real-world phenomenon closer than traditional EBMs. There is more room for heterogeneity, discreteness, and variation than there is in EBMs. As such it is possible to model a phenomenon more precisely.
5. A fifth principle is the absence of determinism. In most EBMs there is only one outcome for a given set of inputs: the specific causation was inevitable. ABMs are typically stochastic, meaning that there is

random variation involved in the process. Small variations in the behaviour of agents in combination with path dependency in the process, can lead to a model in which outcomes of runs may be vastly different.

These general principles highlight when an ABM is a useful method, namely when a system consists of clear individual parts whose actions create a coherent pattern, but the pattern is not directly attributable to any particular individual behaviour. The individual behaviour can vary among different actors but needs to systematic enough to be computationally described. ABMs are useful when individual behaviour is important and differentiated, but a complex pattern emerges. The relationship between individual location choice and cluster formation is such a phenomenon.

Wilensky and Rand (2015) describe eight uses for ABM, which can be used in different situations: description, explanation, experimentation, providing sources of analogy, communication and education, providing focal objects or centrepieces for scientific dialogue, as thought experiments and for prediction (p. 28). As for this study, I mostly focus on the former three. Description refers to the idea that an ABM can describe how a complex system functions, and which factors are important. Explanation means that an ABM can be used to explain why and how certain phenomena come to be. Experimentation is an approach in which the researcher tries to see what happens if a system would behave in a certain way. If a few simple behavioural rules can accurately represent a complex phenomenon, that may be a good indication that a system works in a certain way.

ABM is used here to study the formation of sectoral clusters in a behaviourally focussed manner. The ABM is useful for several reasons. First, the behavioural modelled outline is heavily focussed on decision-making strategies. A large component in this is the behavioural steps that a company undertakes in choosing. Empirical research only typically uses the outcome of this process as their data, namely company relocation data. This obscures the role of the decision-making process. ABM is useful in this context because the decision-making can be precisely modelled. This modelling happens strictly on an individual basis. Because only behaviour of the actors is modelled, the patterns that are observed, emerge from the model itself. This allows for the exploration of the relationship between decision-making strategies and macro-level geographical patterns. In a traditional data-driven design this would not have been possible, because it is not feasible to gather a lot of data on these decision-making processes. Moreover, ABM allows the researcher to make multiple runs of the same model and compare them to explore whether multiple different outcomes are possible. Besides, it is possible to compare different scenarios and see which best fits reality. This ABM follows the simple behavioural rules laid out in the conceptual model. As with any model, the ABM used here is a simplified and stylized version of reality. Making this model requires several fundamental assumptions. In the following section, I go through these assumptions and justify the choices I have made.

3.2. Model Assumptions

Essential in every agent-based model is simplicity. A model that is too complicated and has too many parameters becomes uninterpretable. It is therefore imperative to simplify everything as far as possible (but no further). The model contains many elements which are important in the model functioning but do not relate the core process in study. These elements are reduced to unidimensional objects that only do one thing. The purpose of designing an ABM is not to model reality precisely, but to model one core process in the simplest way possible.

Figure 3-1: Three-phase model

3.2.1. Three-phase model

The decision process, of course, varies wildly from company to company. It is not a coherent, consecutive mechanism. Instead, it is an iterative, erratic and fluid process. However, this is simplified into a model consisting of three distinct consecutive phases: selection, evaluation and decision.

These three phases are defined using both the behavioural and neoclassical insights about the way location choice works. In the selection phase companies search for potential new locations, which are called targets in this study. This incorporates the importance of consideration sets in location choice, as stressed by behaviouralists. This is the phase in which herd behaviour can occur. In the evaluation phase, the targets are evaluated on the locational characteristics. This is a behaviourally constrained neoclassical optimization. The third phase is the decision. In this phase, the company decides whether it stays put or moves to the best location from the evaluation phase. This three-phase model is a stylized adaption of the phasing models of Townroe (1973) and Louw (1996) as cited in Van Dijk and Pellenbarg (2000), using behavioural concepts from Berg's (2014) description of the location choice process. Even though selection, evaluation and decision happen in reality, they are neither distinct not consecutive. Decision-making is a flexible and iterative process, which does not happen in a fixed and structured manner. This is exactly the property of the system that makes it hard to predict and standardize. The three-phase model is a simple and comprehensible way to approximate the decision-making process using basic behavioural rules, which can be algorithmically defined. The three-phase model provides simplicity and clarity while also maintain a stylized form of real-life location choice. The foundations of the three-phase model are explained in more detail in a subsequent section. First, several meta assumptions are discussed, which show some quintessential facets of the model.

1. Markov property:

This three-phase model has the Markov property. Markov chains are processes of discrete sequential transitions in which the probability of each possible event depends solely on the current state of the system, regardless of how that particular state came to be (Maltby, et al., n.d.). This means that only the current distribution of firms affects the relocation decision in this model. Path dependency implies that past and future situations are of little impact on a company's decision and the options of each actor are therefore directly shaped by the current state of the system. In general, assuming 'memorylessness' may not be truthful, but it is a reasonable and useful assumption. It would be possible to program actors with a memory, so that, for instance, the probability of moving would be shaped by the time since their last move. Adding elements like these would make the model more complicated and is not considered relevant in answering the research question. Adding a component of memory is left as a suggestion for further research.

2. Consecutive

The model presented here is nicely structured in consecutive procedures. In reality, decision-making happens in a process that is much more erratic, iterative and parallel. Companies do not follow a strictly ordered decision-making approach and they decide in parallel with other companies. However, to make a model there needs to be a structure. ABMs function properly when they consist of a few simple rules. It is, therefore, crucial to reduce the chaos in real-life situations to a few comprehensible structured concepts, which still describe the essence of the real-world system. The assumption of consecutiveness is followed throughout the model. However, at many points during the model elements of chaos are introduced to mimic the effects of the chaotic process in the real world.

3. Unidimensional and static

Because simplicity is of the essence in ABM, everything that is not immediately pertinent to the core process of the model is reduced to a rudimentary object. This means that both the actors and the environment are modelled in such a manner that they barely reflect real-life situations. In a good model, only the most important things are present to keep the model intelligible. This means that some quite important elements are considered static objects. The number of companies in a simulation remains the same over time: no companies exit and enter. This simplification leaves out an important cause of industrial organization (Klepper, 2010). However, the model aims to investigate the dynamics of herd behaviour in commercial location choice, and not entry- and exit dynamics. This is just one of the many examples in which complex processes are reduced to unidimensional static objects.

4. Stochastic

A stochastic process is a process that is partly or completely shaped by random chance events. Stochasticity is very common in ABM. It results in models which are not deterministic, meaning that the course of the model is not fixed. Stochasticity is also associated with bounded rationality. Because boundedly rational decision-making does not follow a strictly logical course. Stochasticity is introduced in the model to signal events happening by random change, internal variation, chaotic decision-making and erring. This partly compensates for the variation which has been removed from the model under the other assumptions.

3.3. Model design

In the following section, the model design is precisely described. The location choice algorithms are also visualized in diagrams. The algorithm is written from the perspective of the deciding company. The full code for the model can be found in appendix 6.1.

There are multiple software packages for designing ABMs, but the most widely used is the NetLogo software, developed by Northwestern University. In this thesis, version 6.1.1 is used to develop the simulation. The model operates using a 'set-up' and 'go' command. The set-up command sets the initial variables and distributions. This is mainly laid out in the actors and environment paragraphs. The go procedure defines the actual mechanism of the simulation. It does three main things: it defines when the model stops, it calls the calculation of some measurements and it produces the actual relocation behaviours. The model stops after each company has had thirty rounds. This is an arbitrary cut-off point, but after thirty rounds the model has stabilized considerably. In each round, every company gets a 'turn' in sequential order. They reconsider their location one after the other and not all at once. The order in which companies are called is different in each round. The number of rounds can be varied. First, the set-up of the environment and actors is described, followed by each phase of the model. The expert interviews conducted for this study are used to create a model that represents the experience of experts in the field.

3.3.1. Environment

The environment of the NetLogo software is built of squares called patches. The environment is set up in a 5x5 grid, where each cell block represents one work location in a given urban area. Each cell is then built up of 25 patches so that the total number of patches in the simulation equals 625. Because the relocation process is studied on the urban scale, the distance and orientation of different patches relative to one another are considered unimportant. It is assumed that the position of a patch in the city and the distance to other work locations does not matter. The smaller the area considered, the more truthful this assumption becomes.

Each patch has three properties regarding its quality as a business location. These are attractiveness, price, and agglomeration. These variables are important in as far as deciding companies consider them important. Even if agglomerative externalities do not exist in practice, they matter in location choice if deciding companies think they exist. The former property denotes all factors shaping location choice which do not depend directly on the distribution of firms. This encompasses a plethora of divergent elements, many of these are dependent on the distribution of firms indirectly. This applies to transportation infrastructures, residential locations, standards of living etc. In this model, however, these exogenous variables are considered fixed in time and independent of the model. Also, for interpretability and computational ease, all exogenous variables are grouped into one container for an aggregate measure of *ex-ante* locational quality. This is the attractiveness of a location if no other firms would exist.

The latter two both depend on the location of other companies. Occupation, O, is defined simply as the number of companies which are located on each patch. Agglomeration is calculated as the number of companies of the same sector as the choosing company within an arbitrarily set two patch radius. This means that agglomeration externalities in this model are localization and not urbanization economies. This has a practical reason. Using both types of agglomeration at once would cloud the outcomes because this would create two intertwined

process which would be difficult to disentangle in the analysis. It is therefore imperative to use either type. In this case, the use of localization economies is preferred because herd behaviour also happens on a sectoral basis. Localization economies are most comparable to this dynamic.
Prices are set using a supply and demand approach, in which more popular places are more expensive. This measure is not attempted to reflect actual market dynamics but is a simple way to address the fact that high levels of occupation signify high demand and thus higher price. Price levels start at an initial price level of P_w, which is set at 100. Each patch has capacity C which is set at 2. This means that patches with up to 2 companies on them have a price of 100. After that, prices start to rise as space gets scarce and extra office space needs to be created. There is no particular reason to set P_w and C at these values. Regarding the initial price it is fair to say that even if there is little demand for a product, there are still costs associated with the production, maintenance etc. Moreover, prices are only used relative to each other. The capacity is just a construct used to set a price. The idea is that at some point there is no unused land at a given patch. If more companies want to locate there, the price of land goes up. For this model, a deliberate choice is made to give no fixed maximum to the occupation of a single patch, because there is no real limit to the number of offices spaces that can be present in a given area. Building higher, denser, and smaller can increase the number of available workplaces nearly indefinitely (Koster, et al., 2013). The price, p, of a patch j is then given by the equation:

$$p_j = \begin{cases} p_w & , \quad if\ O_j \leq C \\ p_w \times \left(1 + \dfrac{1}{C}\right)^{O_j - C} & , \quad if\ O_j > C \end{cases}$$

This means that each patch has a price that increases when a new company arrives there. The price a company expects to pay for a location is equivalent to the price for $O + 1$. The model works with the assumption that all companies in the same place pay the same price. This may not be entirely realistic but allowing for differentiating in prices for companies in the same place, makes interpreting the effect of the price variable on the model too complicated.

To make a comparable score for attractiveness, price and agglomeration properties, they are scaled from 0 to 100 in which 0 marks the lowest possible value and 100 marks the highest. Concerning the exogenous attractiveness, it is assumed that no patch has all or none of the desired attributes, but that there is some variation among the patches. For sake of structure, I assume that the cells in the 5x5 frame are ordered according to this attractiveness, q_α, and that all patches within each cell have equal values. The first cell on the top left then gets q_α of 75 with a 1.5 drop for each cell, so that the worst cell, bottom right, has a q_α value of 37.5. The scores for agglomeration, q_β, and price, q_γ, for a patch j are defined by the equations:

$$q_{\beta\, j} = \frac{A_j^*}{\max A} \times 100$$

$$q_{\gamma\, j} = \frac{p_j - \min p}{\max p - \min p} \times 100$$

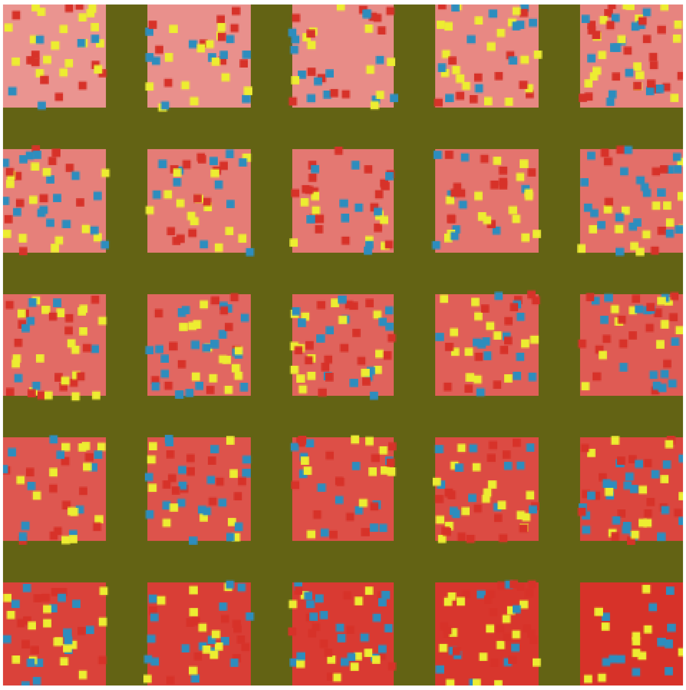
In which A_j^* denotes the number of firms within a two patch radius in a specific industry and $\max A$ the largest amount of companies within a two patch radius of any patch in the simulation. These measures are crude and simplistic constructs and are not meant to be accurate representations of the real-life processes.

In this manner, each patch has a location in a cell and three attributes: an unchanging value for attractiveness, a price which changes over time and an agglomeration metric which changes over time and according to the sector of the choosing company.

3.3.2. Actors

Within this environment, companies choose their location on one of the patches. Because there are 625 patches with a capacity of 2, the number of companies is set at 1,250. The companies are defined as simple behavioural agents with one specific function: relocation. The companies are visualized as coloured squares. The colour indicates to which sector a company belongs. There are three undifferentiated identical sectors. These are relevant in herd behaviour and agglomeration externalities. All agents have the same simple behavioural strategies for location choice, but they have several firm-specific variables which introduce variation into the stock of companies. Companies value different locational characteristics differently. These deviations are dubbed sensitivities in the thesis. The values of these sensitivities are all drawn from a normal distribution with a mean of 1 and a standard deviation of 0.25. These values do not change over time. The sensitivities relate to the importance companies attach to certain variables, relative to the average.

Figure 3-2: Intial situation of companies and environment



Companies also have a propensity to move. This means that a company moves if the potential location is at least a certain percentage better than the current location. This is an expression of the costs associated with moving and inertia in decision-making. Research suggests that some companies are much more likely to move then other companies, specifically younger and smaller companies (De Bok & Van Oort, 2011; Nguyen et al., 2012). Therefore, this threshold can vary quite strongly among companies. This propensity is drawn from a normal distribution with a mean of 0.6 and a standard deviation of 0.25.

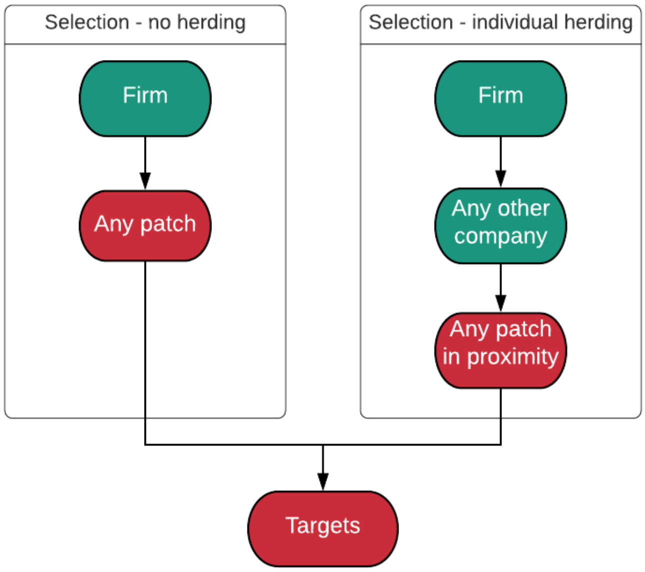
The numbers used here do not bear any significance. They are a tool to create some diversity in the stock of companies.

The choice to have identical sectors is deliberate. It allows for easier interpretation of the model because it is more obvious what causes the emergence of the spatial pattern. It is also a way to model conservatively. For if any sectoral segregation appears in the distribution of firms, this happens despite the sectors having identical demands. This enhances the explanatory power of the model because, for a situation in which sectoral differences exist, these patterns would be even stronger. Also for reasons of conservativeness, companies start in the least ordered way possible: they get assigned random locations as a starting point. In general, this is of course not the case. It would also have been possible to let each company generate its first location in the same manner that they decide about relocation. This is not done for two reasons: 1) this would be more open to pattern formation. It is imperative to avoid any elements that can enhance pattern formation, which are not strictly the process under review 2) there are strong theoretical indications that first location and relocation choices are fundamentally different (Elgar, et al., 2009).

Figure 3-2 shows the way the environment and actors look before the model starts.

3.3.3. Selection phase

Figure 3-3: Selection framework



The assumption behind the selection phase is that companies do not consider all available locations. Instead, they form a consideration set, which is a subset of all available locations. This idea is not only commonsensical, but also substantiated by academic research (Berg, 2014; Van Dijk & Pellenbarg, 2000), and expressed in the expert interviews. Respondent A mentioned that on the local level "*many companies [decide about relocation] somewhat impulsively. Few companies do that very extensively*". In this model companies thus only review a small number of locations.

What determines which locations make it into the consideration set, is a complicated issue and dependent on a multitude of commercial, geographic, and personal factors. Companies seldom perform an elaborate search routine (Berg, 2014; Elgar & Miller, 2009; Van Dijk & Pellenbarg, 2000). Instead, at least on the local level "a surprising number of business locations are apparently discovered by chance, while entrepreneurs are involved with unrelated business activities or during leisure time" (Berg, 2014, p. 1700), similar to what Respondent A expresses in the quote above. Consequently, the formation of the consideration set is assumed to be a random selection. This is to say, the process of target selection is so erratic and unpredictable that it may as well be random. This smaller the geographical scale level, the truer this assumption becomes. The choice of a location within a neighbourhood is more random than the location within a country.

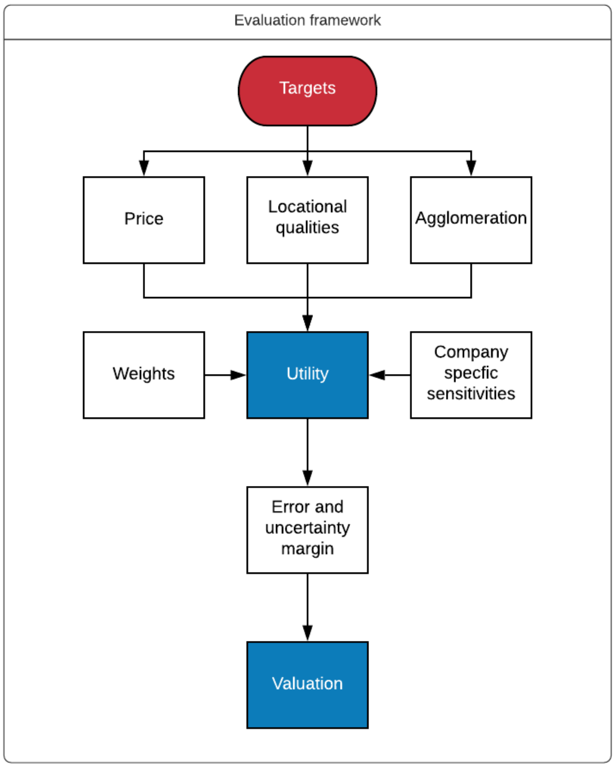
Alternatively, herd behaviour might impact target selection. Herd behaviour can be a factor in the selection phase because companies are likely to look for and/or to encounter an available location in a place they know is used in their sector (Berg, 2010; Vicente & Suire, 2007). Herd behaviour is defined as any type of behaviour in which the probability that locations are considered increases with the number of related companies in the area, regardless of the qualities of that place. Herd behaviour is thus strictly separated from any agglomeration externalities. The selection of targets can happen in two ways. Either a company picks a random patch or it picks a random other company from the same sector and then picks a random location within two patches of that company. In the first situation, all patches have an equal probability of being considered. In the second situation, the probability depends on the number of companies from a given sector in the vicinity.

A further assumption needs to be made regarding the size of the consideration set. This, of course, varies wildly among companies, but in general, it is relatively small. Berg (2014) finds a mode of three locations for local relocation decisions. Many companies have a larger consideration set, but others only ever consider one location. To make a good comparison between companies the size of the consideration set is set at three. The consideration set of all companies must be of equal size because the number of herded and random targets is important in the analysis. In this model, I investigate all scenarios from none to all targets being selected by herd behaviour. These scenarios are called base (no herding behaviour), H1 (one target selected by herd behaviour, two by chance), et cetera. Figure 3-3 shows the algorithm for the selection phase.

3.3.4. Evaluation phase

The evaluation phase consists of a rational and analytical optimization based on a select few criteria. The core of the evaluation phase thus stems from neoclassical theory, but in a constrained manner: the number of both options and criteria is small. The evaluation is simple and one-dimensional: an addition and weighing of scores for some variables and the best option is chosen. Decision-making scholars may rightfully point out that this is not how decisions are made. To compensate for this, the decision-making is modified to include elements of bounded-rationality, idiosyncrasy, and stochasticity, to approximate a real decision-making process. This crucial phase of the model is thus a behaviourally constrained neoclassical approach, and not a behavioural model (Berg & Gigerenzer, 2010). The algorithm for the evaluation framework is visualized in figure 3-4.

Figure 3-4: Evaluation framework



The inputs of the evaluation phase are the locational properties of the targets selected in the selection phase. These properties have scores q_α, q_β and q_γ for attractiveness, agglomeration and price respectively. The calculation of these scores is explained in the environment section, Environment3.3.1. The scores are multiplied by the idiosyncratic sensitivities and the weights. The weights indicate the relative importance of a variable compared to the attractiveness of a location. The relative importance of these variables is not known, so several assumptions are made in this respect. Price is not the focal object of this study, so the weight is just set at half the importance of the locational qualities. In the validation section, some other values are tested to see whether this changes the dynamics of the system.

The weight for agglomeration externalities is crucial in this study. It is one of the central objects of study in this thesis because it designates the importance of agglomeration externalities in the model. This weight can be turned on and off. When it is off, agglomeration externalities do not play a role in that version of the model. When it is on, it is randomly drawn between 0 to 1. This shows how incremental changes in the importance of agglomeration externalities affect clustering dynamics. Using the variables, sensitivities and weights as inputs it is possible to calculate a utility U for each firm(i)-patch(j) combination:

The final utility of a place j for a company i is then created as follows:

$$U_{ij} = \frac{q_{\alpha\, j} \times W_\alpha \times s_{\alpha\, i} + q_{\beta\, ij} \times W_\beta \times s_{\beta\, i} - q_{\gamma\, j} \times W_\gamma \times s_{\gamma\, i}}{W_\alpha \times s_{\alpha\, i} + W_\beta \times s_{\beta\, i} + W_\gamma \times s_{\gamma\, i}}$$

In which W denotes the weight for each variable and s denotes the sensitivity for each variable for each company. The final valuation V_{ij} is then calculated by adding an error margin. This reflects a company's inability to make accurate judgements about this optimization question, both regarding the estimation of the values for the variables as well as in the comparison of the selected targets. This error margin, ε_{ij}, is randomly drawn from a normal distribution with a mean of 1 and a standard deviation of 0.25. In the valuation of the current patch ε_{ij} the standard deviation is half as large at 0.125. This indicates that companies have a more accurate view of the qualities of their current location than of potential future locations.

$$V_{ij} = U_{ij} \times \varepsilon_{ij}$$

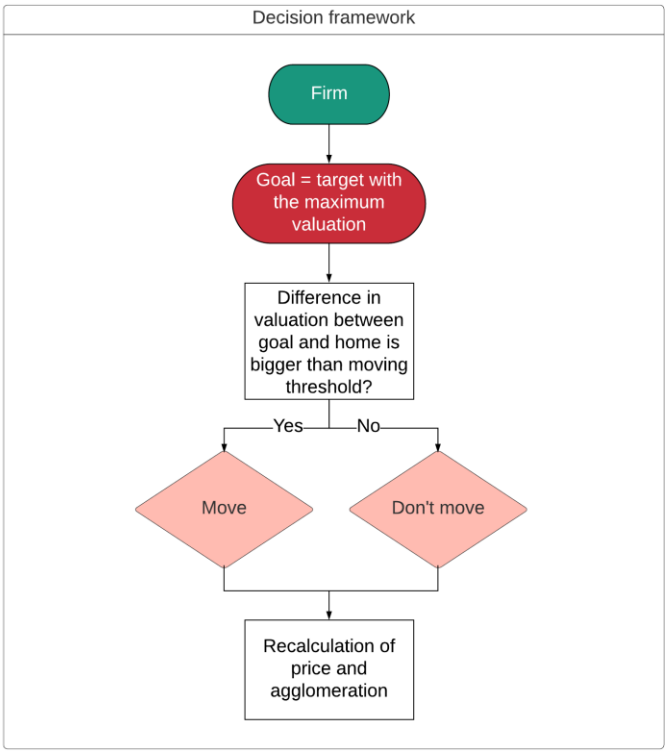
3.3.5. Decision phase

The decision phase assumes that a company has found a patch that they might consider for relocation. This is the target with the highest V_{ij} of the consideration set. They then decide between moving to that place or staying put. This decision is made based on two key premises. The first is that the company makes an honest comparison between the current and alternative location based on the valuations found in the evaluation phase. The second premise is that companies have a strong inclination to stay in their current location. This has two reasons, namely the costs and risks associated with moving and the forces of inertia in decision-making. The company only moves when this threshold is overcome. The decision is therefore dependent on the perceived utility of the current and alternative locations and the idiosyncratic propensity to move threshold, ζ. The decision to move thus takes place if,

$$V_A - V_H > V_H \times \zeta_i$$

In which V_A and V_H denote the valuation of the alternative and current location, respectively. If the company decides it will move, it changes its location in the simulation. This instigates a process in which variable scores for locations are recalculated. If this is completed, or if the company decides to stay, the next company is called. The algorithm for the decision phase is visualized in figure 3-5.

Figure 3-5: Decision framework



3.4. Analysis

The analysis of the model is a descriptive analysis of the impact of different variables on clustering. For the answering of the research questions, the model will be run many times under different parameters to investigate what impact they have on relocation decisions.

3.4.1. Data generation

The data is generated by running the model many times over. Each run of the model generates three observations, one for each sector. The parameters for herding and agglomeration are not fixed. Each dataset contains data for four herd behaviour scenarios with one through three targets selected this way (dubbed scenarios base, H1, H2, and H3) and data with and without agglomeration externalities. When there are agglomeration externalities, the weight is randomly drawn between 0 and 1.

The study will investigate how these different scenario impact clustering. To do this, clustering needs to be measured. This is done in two ways, to make sure that the measurement is valid. The first method is an intuitive, simple measurement. It calculates the percentage of companies whose closest neighbour is in the same sector. When there is no spatial clustering with three sectors this measure is expected to be ⅓. When companies are completely segregated according to their sector this is 1. This measure is called segregation.

The second measure is Ellison and Glaeser's (1997) dartboard approach. This shows how strongly a distribution deviates from a distribution of randomly thrown darts on a map where the size of each cell is proportionate to the share of total companies in that cell. This score, γ, is calculated as follows:

$$\gamma = \frac{G - H}{1 - H} = \frac{\sum_{i=1}^{M}(s_i - x_i)^2 - H}{1 - H}$$

Table 3-1: Excerpt of the dataset

Run	Turns	Firms	Capacity	Sector	Herded targets	Agglomeration	Weight price	Weight agglomeration	segregation	Gamma	Modes
7	30	1250	2	2	0	false	0.5	0	0.31	0.04	4
68	30	1250	2	3	2	false	0.5	0	0.39	0.09	10
98	30	1250	2	3	1	true	0.5	0.97	0.99	2.03	1
150	30	1250	2	3	2	true	0.5	0.19	0.65	0.87	3
160	30	1250	2	1	3	true	0.5	0.44	0.91	1.81	6
165	30	1250	2	1	3	true	0.5	0.98	1.00	2.00	3

In this formula, H refers to the Herfindahl index, which is a measure of the distribution of market shares among companies. The one-dimensional companies have equal sizes, so this index defaults to $\frac{3}{N_C}$, where N_C is the number of companies in the simulation. G is the sum of the squared difference between the share of companies in a sector that are in a particular cell (s_i), and the total share of companies that are in that cell (x_i), for all cells M. Gamma is calculated for each sector separately. Lastly, the dataset includes the rank of the cell on which most companies in that sector are located: the *modal* cell. The cells are ranked on attractiveness. This includes the following data: the model parameters, the cluster measures, and the modal cell. Data are reported for each sector which means that each run results in three rows in the data set. Also, for each cell, the number of occupying companies were counted.

Table 3-1 contains a random excerpt of the results showing the structure of the data, without the counts per cell.

3.4.2. Data analysis

The analysis of this data shows how the measures of clustering respond to the introduction of herd behaviour and agglomeration externalities. The purpose of this is to investigate whether herd behaviour can partially or completely explain the same degree of clustering as agglomeration externalities, and whether or to what extent herding changes the predictability of the model. Because this thesis is somewhat exploratory in nature, the analysis does not have the form of a rigid statistical analysis but it is a descriptive analysis. This analysis investigates the effect of the number of herded targets and the importance of agglomeration on the degree and location of clustering. The outcomes of this analysis are visualized in several graphs and tables that show how clustering changes with varying parameters. There is no statistical analysis because, given enough runtime, any slight change makes the model significantly different, but the effect size can be rather small[2]. Even though not reported as such, all presented results are also statistically significant.

For the first research question, the effect of introducing herd behaviour and/or agglomeration externalities on the cluster metrics is analysed. This is done by running the model with different levels of herd behaviour and agglomeration externalities and observing changes in the cluster metrics. The second research question regards the consistency of the model. A predictable model would lead to the highest number of companies on the cells with the best locational qualities and would result in similar degrees of clustering in each run of the model.

3.5. Validation

When using agent-based modelling, it is important to think about the validity of the results. In other words, whether the outcomes of the model reflect real-world processes. The validity of this model is substantiated in three ways. Through theoretical grounding, in the research design and analysis, and with expert interviews. The behavioural model that is proposed here is firmly rooted in existing scientific knowledge. Behavioural economic concepts of decision-making, the fundamentals of location choice, as well as the theory regarding cluster formation have been used. By using scientific theory as a background for the model design the credibility of the model is enhanced. All behavioural steps are based on insights from this literature. This does not mean that the data this model produces are immediately true, but it shows that the basic dynamics of the model are stooled on pre-existing understanding. In the analysis, the findings of this thesis are continuously mirrored with the

[2] To show this by example: with a 150 observation per group, the clustering of the H1 scenario without agglomeration is significantly different from H0 with a p-value smaller than the smallest number R can process ($p < 2 \times 10^{-16}$), despite the fact that the difference between these scenarios is among the smallest in this study. (example from chapter 4.1.2)

literature to establish the connection with the scientific understanding of the relocation and clustering phenomena.
Secondly, the validity of the model is a central point of the design. This takes form in two ways. The model is designed conservatively. This means that elements which might have reinforced the clustering but are not the focus of this study were purposefully obscured. Conceptualizing the sectors with identical locational demands, and the random initial distribution of companies are ways in which this is done. It leads to less sectoral clustering than may be theoretically justified, but it makes sure that clusters do not form for other reasons than those which are under review. Even though these assumptions are not realistic, they strengthen the validity of the model because they show that clustering emerges even with these constraining assumptions. Robustness is another way of testing the validity of the model internally. This means that the model dynamics occur under a wide variety of parameters and not just in one specific combination. If the reported outcomes occur in every setup, it shows that the results are pertinent. Therefore, the analysis will also include versions of the model under different parameters.
Lastly, three expert interviews are conducted to establish the connection between this thesis and actual location choice practices. The interviewees work in local government or consultancy and are experts in commercial relocations. The first expert works at a municipality in the Dutch province of Zuid-Holland as a project manager for business parks. He is amongst other things responsible for issuing new business lots to companies. As a result, he has extensive knowledge about how companies make location choices and why they move to his municipality. The second expert is a consultant at a large real estate firm, where he advises local governments in the development of new business parks. He has both international and domestic experience and knows what strategies are used to create a successful business park. The third expert is a consultant at a different real estate company, where he counsels companies in finding the right place for their business. The insights from these interviews are used to interpret the results of the model. If their statements signal similar dynamics as the model, this is a good indication that the model represents an actual, real-world process. The interviews are discussed to establish whether the outcomes of the modal fit the image these experts have about clustering and relocation. Transcripts of the interviews[3] are attached in appendix 6.4.

These three approaches enhance the credibility and validity of the results. This does however not mean that the results precisely measure the real world. The findings of this study show the potential impact of herding behaviour in location choice in a simple and constrained environment and are aimed at analysing the dynamics of spatial organization. The results need to be viewed on an abstract level. The agent-based model does not produce any particular values which can be considered predictive for real-world situations. Instead, they contribute to the understanding of the formation of clustering patterns in the spatial distribution of firms in general. Because the model is set up very crudely, the results need to be viewed at the level.

[3] All interviews were conducted in Dutch. Whenever a respondent is quoted, the quotation was translated into English by the author.

4. RESULTS

It is possible to investigate a multitude of processes from this model, but I only look at the phenomena for which the model is designed. This means looking at spatial patterns, and how they come to be. To answer the research questions, the focus will be on the different conditions under which clusters form and, the role of herd behaviour and agglomeration externalities in these processes.

4.1. Model outcomes

The purpose of this section is to compare herding in the selection phase to agglomeration in the evaluation phase. To do this I ran four versions of the model: a base scenario without both herding and agglomeration, a model with agglomeration externalities, a model with herd behaviour, and a model with both.

4.1.1. Base scenario

Before exploring the impact of herding and agglomeration on spatial patterns, it is necessary to first establish what the reference situation looks like. In this version, neither herd behaviour nor agglomeration externalities exist. The outcome of this situation is quite predictable. There is no sectoral clustering: gamma and the segregation metric are around their minimums (see table 4-1). This means that companies are no more likely to be close to related companies than they are to any other company. The variation between the models is also small: both the standard deviation and range are small. Regarding the distribution of companies, it is unsurprising that the number of companies in a cell is proportional to its attractiveness. Figure 4-1 shows a boxplot for each cell, showing the number of occupying firms at the end of different runs. The range of possible outcomes here is also small: the distribution is rather stable across runs. There are some small variations between different runs, but this is quite small, and the model is rather consistent.

4.1.2. Agglomeration

The version of the model with the inclusion of agglomeration externalities is more interesting than the base scenario. Agglomeration externalities are the most common explanation for the development of clusters in economic geography. The precise importance of agglomeration is not known, but for this study I assume that it is somewhere between absent and as important as the qualities of the location: weight for agglomeration is between 0 and 1, meaning between 0% and 100% of the weight for attractiveness. Varying this weight between 0% to 100% produces figure 4-2. This figure shows how clustering responds to agglomeration externalities. Even though clustering responds relatively strongly to agglomerative externalities, the association is not linear. Up to about 20% importance, agglomeration causes just a minor increase in gamma, however with higher levels of agglomerative externalities clustering rises rapidly, until it starts to flatten off at the higher end. A second observation is the increased range of possible outcomes for higher levels of agglomeration. The more important agglomeration is, the larger the possible variation in the degree of clustering. At the high end, most values fall into the 1 to 1.75 range. This implies that the model becomes less and less consistent when agglomeration becomes more important.

Table 4-1: Base scenario metrics

Base scenario metrics (N = 150)	Mean	St. dev.	Minimum	Maximum
Gamma	0.03	0.01	0.01	0.07
Segregation	0.34	0.02	0.30	0.38

Figure 4-1: Distribution of firms across cells in base scenario (N = 75 per cell)

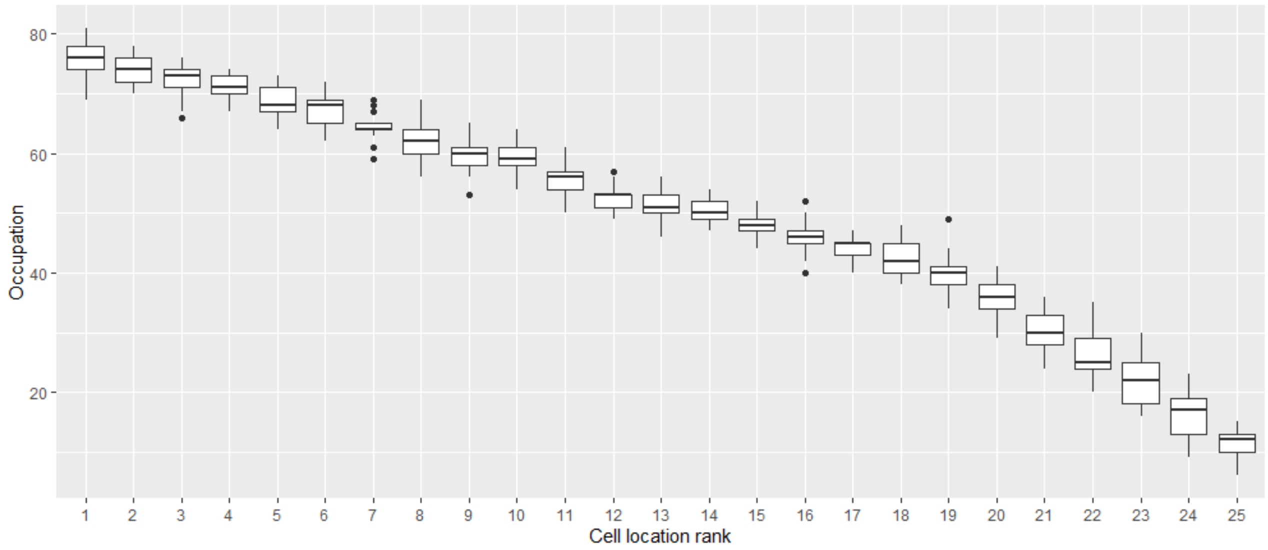
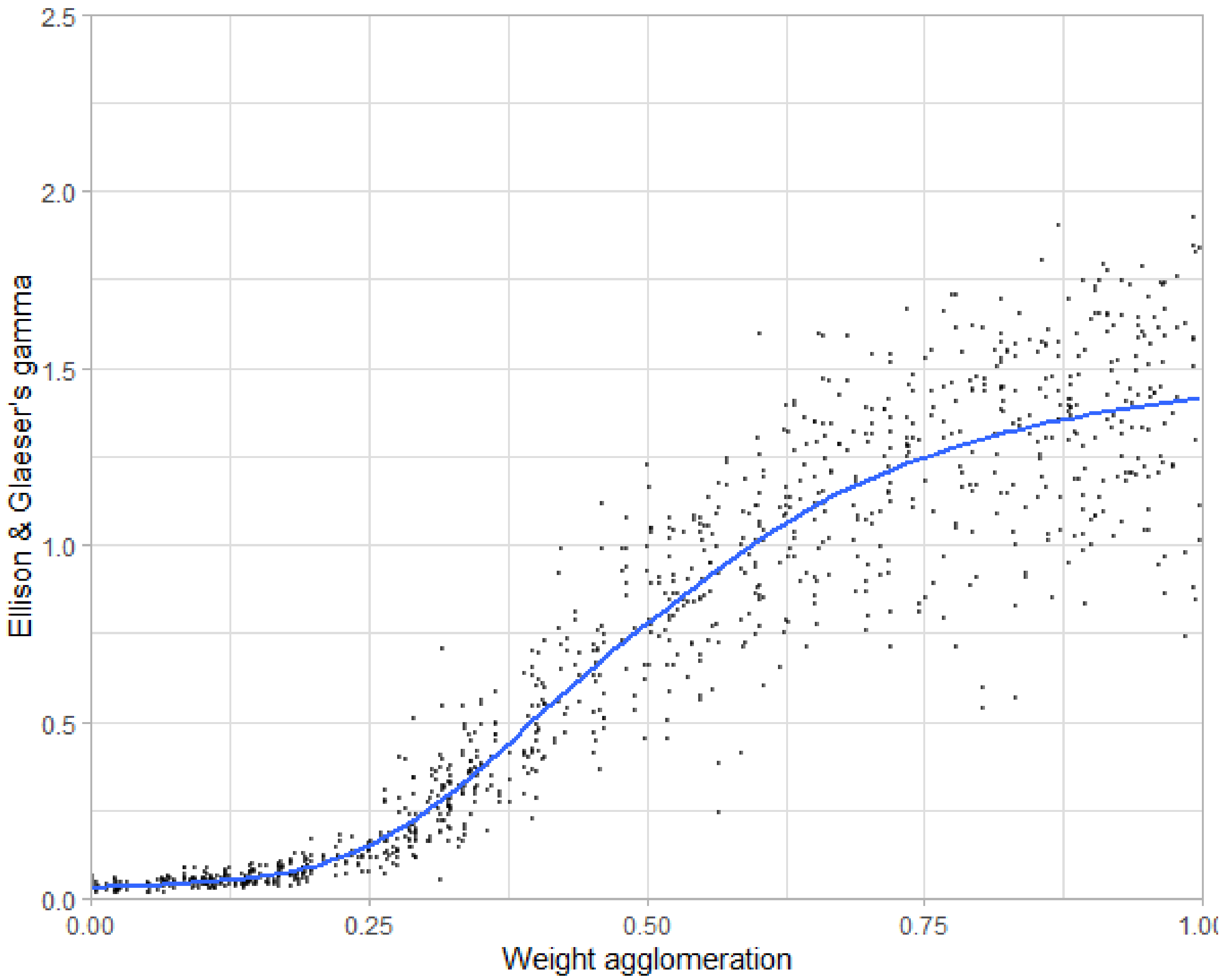


Figure 4-2: Sectoral clustering as a product of agglomerative externalities without herding (N = 1050, includes a generalized additive model)



4.1.3. Herd behaviour

Introducing agglomeration externalities in the model enhances the degree of clustering. In this section, I delve into the question of whether similar effects can be reached by introducing herding behaviour into the base model. In these scenarios, the companies are more likely to consider locations which are close to related companies. Consequently, the probability of being considered depends on the number of companies in its proximity. This manner of selection leads to a slight increase in the sectoral clustering of companies. The effect of herd behaviour on clustering can be seen in figure 4-3: with the increase of imitation in the simulations clustering increases, but the model also becomes more inconsistent. The figure shows a violin plot with Gamma on the y-axis, separated by the herding scenarios. Horizontally a Gaussian kernel density of is plotted. This plot shows that clustering increases when herding increases, but also that the observations are more dispersed. This means that the degree of clustering in the model becomes less predictable. These data make clear that herd behaviour has the potential to change the spatial distribution of firms as a result of the changed manner of forming the consideration set, even if the actual evaluation of locations happens in the same manner as the base scenario. The large difference between the H2 and H3 scenarios can be explained by the fact that in H3 a sector can by definition never return to a cell when there are no related companies in that cell. These metrics for clustering, show that herd behaviour can create the same degree of clustering as moderate agglomeration externalities up to a weight of about 30%. To show this, the metrics for herd behaviour are mapped to certain brackets of agglomerations weights in table 4-2. This shows that modest amounts of clustering can be explained by either herding behaviour or agglomeration externalities.

It should, however, be noted that the degrees of clustering produced by herding are very minor. In fact, looking at the visual output of the model it is hard to discern which variant was produced by which scenario. In the base scenario, a plot was produced to show that the distribution of firms across the cells was logical and stable. This plot included a boxplot for each cell. The same plot is reproduced in figure 4-4, but with the data for the H2 scenario and the matching bracket of agglomeration externalities (17-23%) included. This shows that the distribution is almost identical. The inclusion of herd behaviour or agglomeration externalities does not considerably change the distribution of the firms across the cells. Basically, the only noticeable differences are 1) the fact that these alternative scenarios tend to produce a slightly more skewed distribution, in which more companies are located in the most attractive cells compared to the base scenario, and 2) that the H2 scenario produces slightly more outliers. This means that the model can occasionally result in outcomes which are rather divergent from the median, but this does not happen all too often. The former effect means that both agglomeration and herding behaviour create benefits mostly for the patches which are already the most attractive. Because herding does not affect the evaluation of a patch, one could, at least in this boundedly rational model, deduce from this observation that imitation improves the search routine of companies, at least in this boundedly rational model. The latter effect implies that herd behaviour has the potential to produce more unlikely outcomes, because of the cumulative effect of imitation.

Regardless of these two observations, the difference in the degree of clustering in these models is negligible. Figure 4-5 shows the outcomes of a run of the base and the H2 scenario. For each sector, the gamma scores are given, and these are relatively average for their respective scenarios. H2 produces slightly clustered outcomes, but it is rather hard to find this out without statistical measurements.

Table 4-2: Comparison of herding and agglomeration scenario's regarding clustering (Ellison & Glaeser's gamma)

	mean	sd	IQR	Min	Max	Segregation
Base	0.03	0.01	0.01	0.01	0.07	0.34
H1	0.05	0.02	0.02	0.02	0.09	0.35
A 5-15%	0.05	0.02	0.02	0.02	0.12	0.35
H2	0.09	0.03	0.04	0.04	0.19	0.39
A 17-23%	0.09	0.03	0.04	0.04	0.18	0.38
H3	0.23	0.07	0.08	0.08	0.46	0.48
A 27-31%	0.24	0.10	0.13	0.11	0.51	0.44

Figure 4-3: Mean and distribution of gamma in different herding scenarios (N = 150 per group)

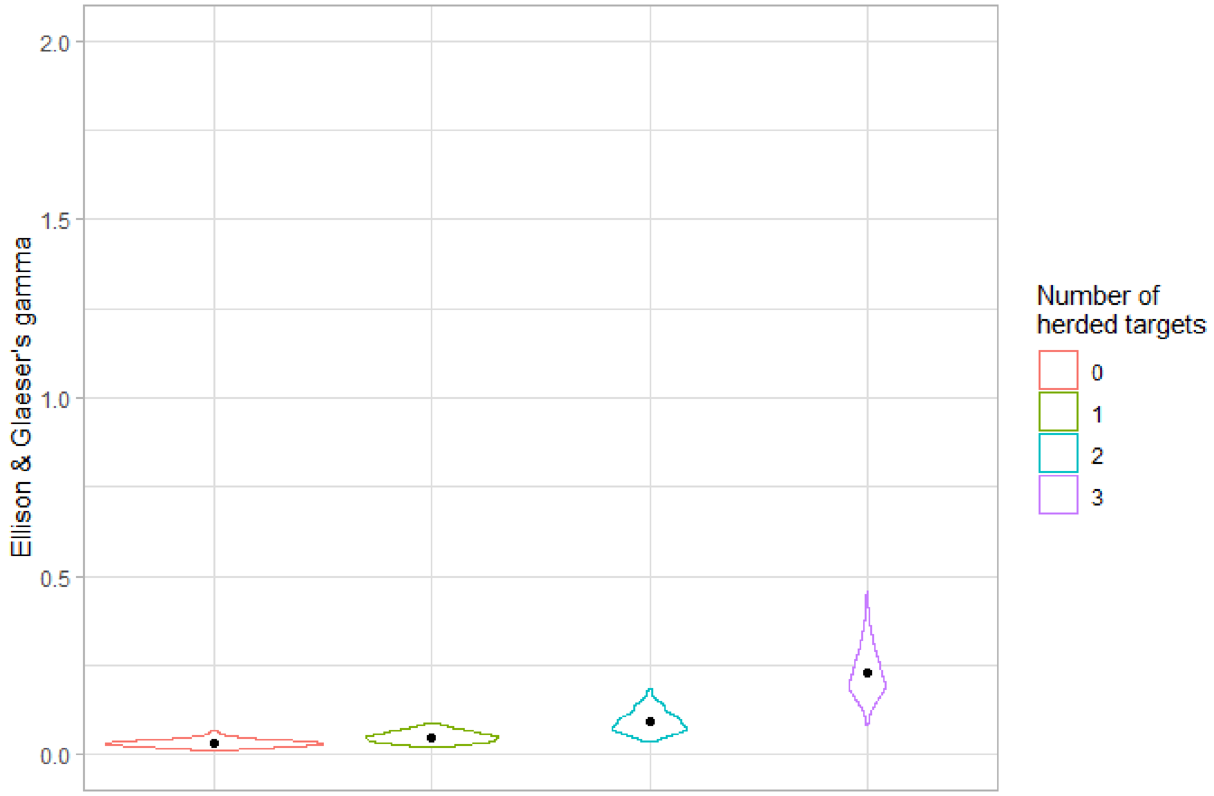


Figure 4-4: Distribution of firms across cells in three different scenarios.

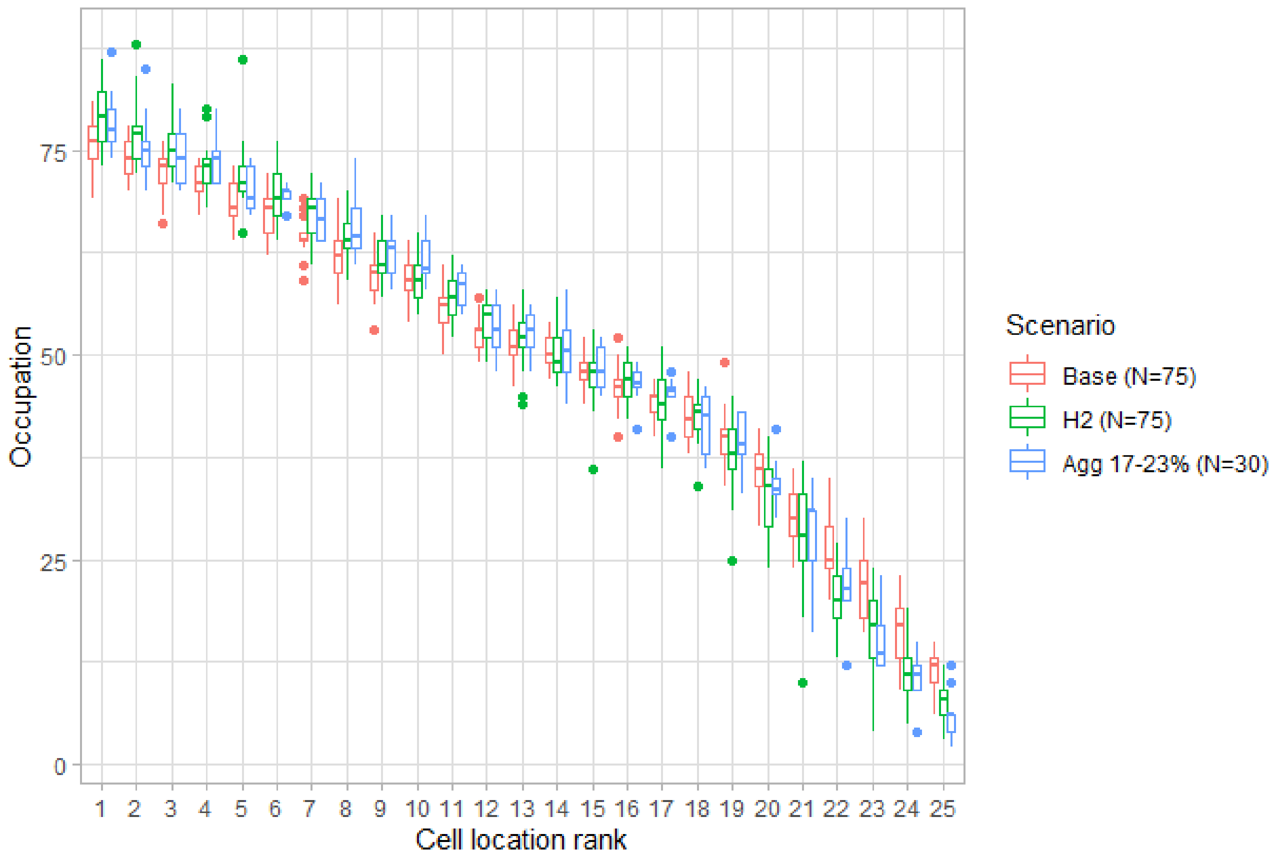
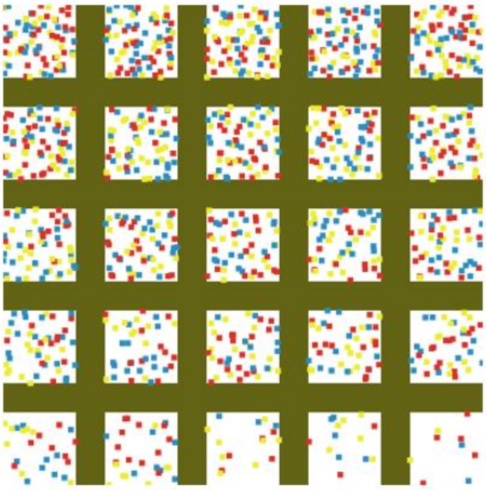
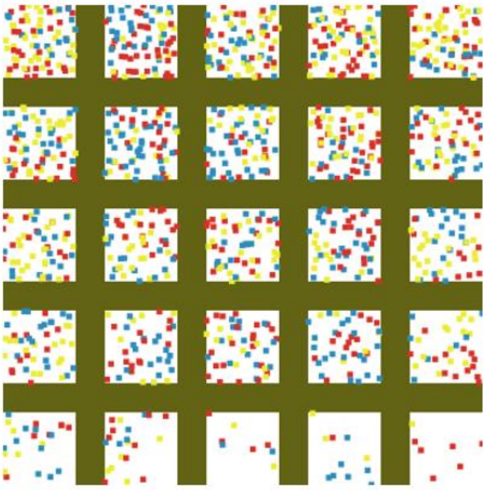


Figure 4-5: Model view: two example runs from scenarios base and H2





Scenario H0: Gamma = 0.02, 0.03 and 0.02 respectively for blue, red and yellow.

Scenario H2: Gamma = 0.10, 0.05 and 0.09 respectively for blue, red and yellow.

4.1.4. Combined

Herd behaviour can only explain a negligible degree of clustering on its own. However, a combination of herd behaviour and agglomeration externalities is rather powerful in explaining clustering. Figure 4-2 showed the effect of agglomeration on clustering. This figure is expanded with data from the herding scenarios to create figure 4-6[4]. This visualization shows that the introduction of herd behaviour in this system fundamentally changes the dynamics of the system. Noticeable in this graph are:

1) As established in the previous paragraph, the starting point of the herd behaviour scenarios is higher (the level of clustering in the absence of agglomeration externalities).
2) The tipping point occurs at a lower weight of agglomeration externalities in the herding scenarios. The weight of agglomeration externalities necessary to generate strong growth in clustering is lower with more imitation.
3) The herding scenarios all plateau at the maximum degree of clustering[5]. Regardless of the amount of herding, all imitation scenario plateau at this value, whereas the scenario without herding, stays considerably lower.
4) Between the tipping point and the plateauing, the different herding scenario graphs have more or less the same slope, whereas the curve of the scenario without herding is considerable flatter.
5) Herd behaviour does not lead to an increased range of possible outcomes. The herd behaviour scenarios all converge to one stable outcome for higher values of agglomeration. At lower values of agglomeration, the spread in the distribution of values is relatively similar for all four groups.

These observations show that the introduction of herd behaviour fundamentally alters the behaviour of the system. The actual amount of imitation in the system changes the model slightly (it more or less moves the curve to the right), but the effect of the three herding scenarios on the degree of clustering is relatively similar. If herding is completely absent, on the other hand, a substantially different outcome appears.

Another interesting approach is to look at where clusters come to exist. In general, companies are expected to prefer the most attractive location, which are thus the most logical places for clusters. Because the cells are ranked according to their attractiveness, cell 1 is the most attractive and cell 25 is the least attractive. To visualize the distribution of clusters, I selected all sectors with gamma greater than 1 and plotted which cell contained the largest number of companies of that sector (the modal cell). This is shown in figure 4-7. This shows the weight

[4] Figure 4-6 is plotted with outliers removed. These outliers are one specific type of case. In some circumstances two sectors create a cluster on the same cell. This creates a gamma of circa 0.4 for these cells. 24 observations were removed, because they had a disproportionate effect on the best-fit line.

[5] The value of gamma = 2.0 is reach edwhen each sector has its own cell, where all of their companies are located. Higher levels of gamma can only be reached if this condition is (almost) true for one sector, but not for other sectors. This is because the calculation of gamma depends on the ratio between the share of companies of a specific sector in a cell and the share of all companies in that cell.

for agglomeration on the x-axes and is facetted according to the different herding scenarios and coloured according to the degree of clustering. The red lines represent the generalized additive model as a best-fit line. This figure shows where sectors have clustered in different runs. The plot has a lot of information, which allows for several observations:

1) The more attractive a cell is the higher the probability of cluster formation in a cell. 21% of modal locations is on cell 1, 17% on cell 2, 11% on cell 3 et cetera.
2) The distribution of modal location is rather stable: the range and the average remain virtually constant under varying circumstances. Changing the weight of agglomeration does increase the degree of clustering, but it does not change the location of the cluster. The mean rank of modal location stays more or less consistent. Besides, the spread and average of the rank of the modal location are also more or less the same in all herding scenarios. The rank of the modal location also does not say anything about the degree of clustering. Clusters on highly ranked cells are on average no more or less clustered than clusters on a lowly ranked cell. In fact, the rank of the modal location does not show a strong correlation with any of the measures involved.
3) It that reconfirms the combination of herd behaviour and agglomeration externalities strongly increases the degree of clustering in a simulation.

Figure 4-6: Sectoral clustering as a product of agglomerative externalities with herding (N = 3726, includes a generalized additive model, outliers removed)

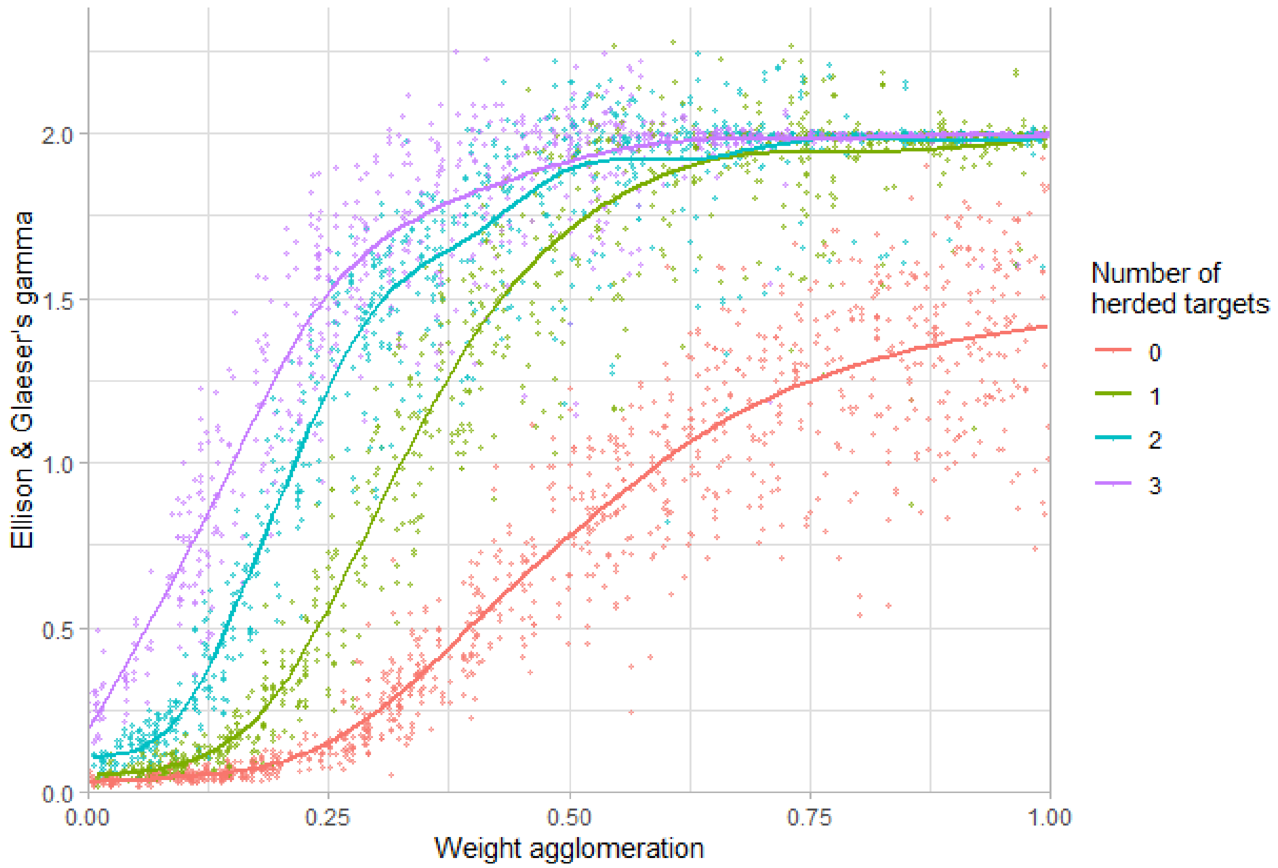
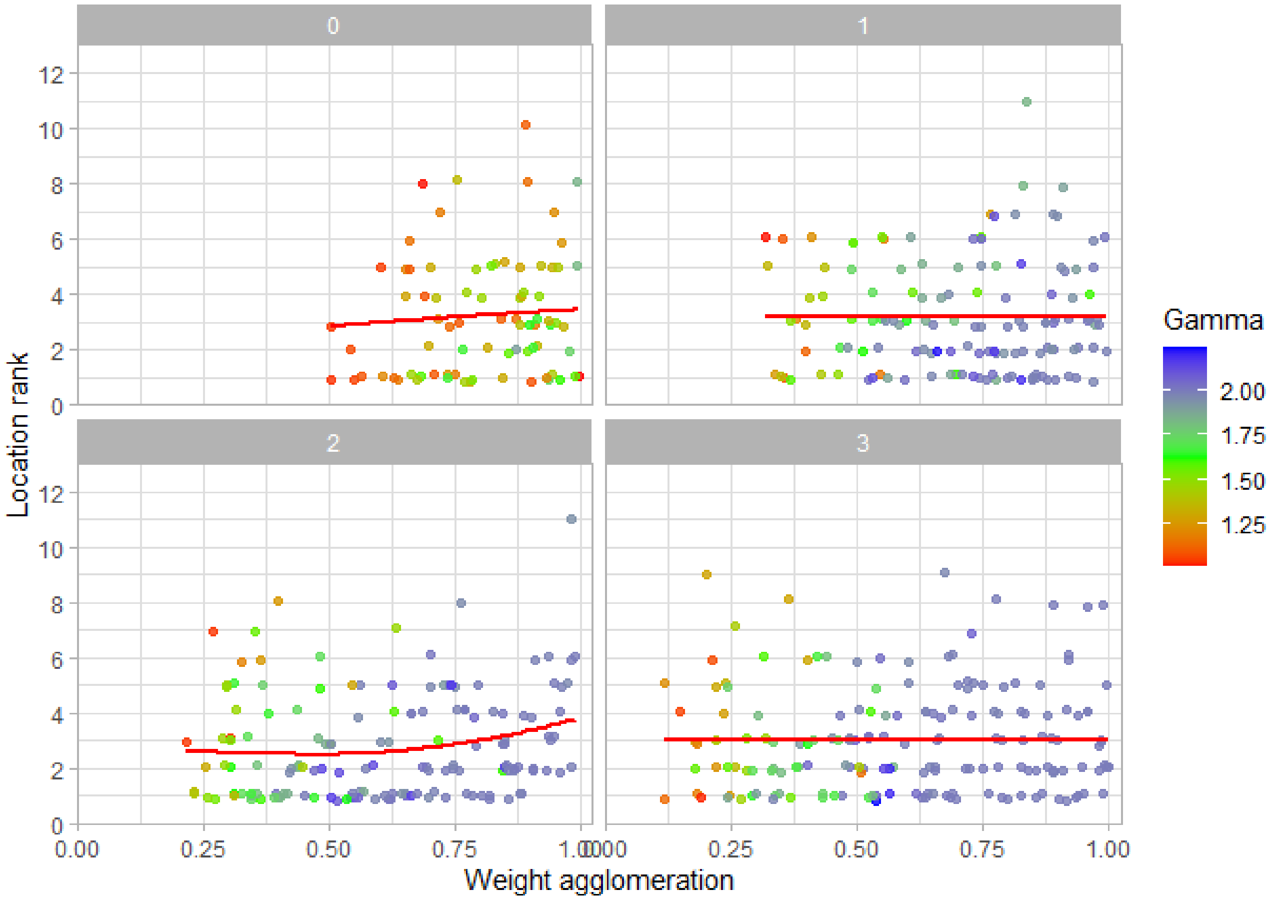


Figure 4-7: Modal cluster location rank by herding and agglomeration (N = 900). Points jittered vertically for visibility.



4.2. Validation

Before anything meaningful about the real-world implications of these data can be said, it is important to assess the validity of the results. That means whether outcomes are applicable outside of the specific context of the model. In the research design, this was done by specifically modelling the simulation on previous scientific knowledge. In this chapter, the validity is assessed in two ways. First, I show that the reported outcomes are not the product of a particular unsubstantiated parameter set-up, but are indeed distinctive properties of the model. I do this by investigating alternative parameters. In the second section, I connect the findings from the model with the real-world experience of three experts. This is done to establish whether the results here are indicative of actual geographical phenomena.

4.2.1. Robustness

Robustness refers to the idea that the results from an agent-based do not only occur only under a particular set of parameters but are indeed general properties of the model. Changing the parameters in the model, of course, affects the outcomes, but slight changes in these supporting parameters are not supposed to change the fundamental behaviour of the model. In the model design, I introduced several random parameters. These were not supported by any empirical evidence. One of these decisions is to stop each run after thirty timesteps. One could argue that the results drawn here are just a snapshot of an ongoing process and do not reflect any definitive outcomes. Though this is true, the amount of change that occurs from step to step decreases rapidly after the first few steps. In fact, thirty steps are more than enough to create a stable view of the spatial distribution of companies. Increasing the runtime of the model leads to an increase in clustering, but the actual patterns do not change much further. With 30 steps the direction of the outcomes is clear. In real life, there are always changes in this distribution, but there are no exogenous developments in the model. This results in a stable long-term equilibrium. To test for this stability, the model was also run with forty instead of thirty steps per run,

which results in more or less the same outcomes (see Appendix 6.2). Notably, the degree of clustering is consistently somewhat higher in the variant with forty rounds. This is because the clusters have developed further. One could infer from this that cluster formation is inevitable provided that the model runs long enough but this is not the case. The variations of the model were run for extended periods of time and no substantial clustering occurred. One version included 20% weight for agglomeration and no herding. This variant ran for 1250 rounds and resulted in gammas of 0.12, 0.14 and 0.18. A second version included 10% weight for agglomeration and one herded target. This variant ran for 2800 rounds and resulted in gammas of 0.08, 0.09 and 0.14. This shows that thirty rounds are probably enough to establish the dominant patterns. The number of runs is large enough to make meaningful conclusions about equilibrium effects. The clusters which have formed might develop further, but new clusters are unlikely to form.
Besides, two methods of measuring clustering are used. Both methods yield very similar outcomes. For the sake of space, the graphs are only produced for gamma but using the segregation metric results in almost identical visualizations. Gamma is the preferred metric here because it has a basis in the literature on clustering. The fact that both measurements show the same patterns indicates that these patterns are not the product of measurement sensitivities but reflect actual spatial distributions.

In addition to the model with a longer runtime, I established three alternative models. These include two models in which the weight for price, number of companies, patch capacity, and number of steps are changed. This shows how different types of parameters affect the created patterns. Herein the weight for price is especially important as it countereffects the agglomeration externalities. This means that when price becomes more important, higher levels of agglomeration weights are necessary to overcome the burden of price and create clustering. A third alternative model relates to the differences between the cells. In the original model, the attractiveness of cells scaled down with steps of 1.5: the attractiveness of each cell was 1,5 lower than the previous cell. In this alternative model, this step was expanded to 2,5. This model shows what happens if the various cells are less similar.

These alternative setups create rather different results, but in essence, support the same main findings. In these alternative scenarios, two main distinctions are noticeable. In the association between herding and agglomeration, the tipping point can occur at any weight of agglomeration depending on the particular parameters. The range and mean of the modal location rank also vary with the parameters. They do not change in different herding scenarios, but they can be impacted by agglomeration externalities in some situations. These differences are interesting but do not deny the outcomes of the model. The observations made in the previous paragraph are also supported by these alternative models, even though the particular outcomes are different. These differences do not negate the principal findings, because the purpose is to look for general patterns and not particular findings. The particular values a model produces do not mean much outside of the model context, but the general patterns do. Consequently, these alternative models validate the observations from the main model. Results from the comparison between the initial and alternative models are shown in appendix 6.2.

4.2.2. Interviews

Another way to substantiate the results from the model is the use of the interview data. In these interviews, three expert participants shared their vision on firm relocation choice and clustering. This was first used to build the model but their observations are also used to validate the model. The interviews do not directly reflect on the model but show the experiences of experts in the field of commercial relocation. They bring forward three elements which can be used to show that the outcomes of the model are in line with first-hand experience. The interviews were based on a topic list (shown in Appendix 6.3), which includes the topics of discussion. Herd behaviour was not explicitly part of the topic list to allow respondents to introduce their personal vision on relocations and clustering. The respondents knew however that herd behaviour was the topic of this thesis. The statements from the experts are used to discuss, firstly, why clusters form, secondly, where clusters form, and thirdly, the role of non-economic factors in this process. The full Dutch transcripts from the interviews are presented in appendix 6.4. Quotes presented here were translated from Dutch by the author.

The experts all attributed clustering to real-world benefits. Clusters form because colocation offers operational benefits to companies. Respondent B, for instance, says that "*the main reason for locating in a cluster is that it is specialized. Specialized in labour, suppliers, and materials*". Respondent C adds the importance of a network: "*Companies are relocating to Amsterdam because their network is there. Business networks are really important location factors for office firms*." The three respondents can easily sum up a list of factors which makes a cluster attractive. There are certain benefits of colocation. Some of these are quite obvious, such as specialized labour and materials. Other factors are a bit fuzzier, such as the presence of a network, or the image of a business park.

The question of where clusters form is also interesting because it is not always predictable. Respondent C uses the example of the Zuidas in Amsterdam as an example of how a prime business park emerged almost coincidentally:

> The Zuidas emerged, basically, because ABM Amro was located there, more or less by accident. Historically, the entire financial sector was in the city centre, but at some point, there was no more space, so they went looking for an alternative location. Amsterdam Zuid was a logical option; it was already a wealthy area of course. There were some banks and lawyers there. Even further south, ABM Amro and the predecessor of ING were located at what is now the Zuidas, but back then it was an empty field with some sports facilities on the ring road. They built their offices there. At some point, the court also relocated there, which meant that large legal firms, which used to be in [Amsterdam] South and the city centre moved to the Zuidas, close to the court. […] and then came business in legal and financial services coming to there. Thus it became the prime business district of the Netherlands. Earlier it was just legal and financial, but at some point, other headquarters came, which dealt with legal and financial services.

This example clearly shows the path-dependency and natural development of clusters. This is also something that is affirmed by other respondents. Clusters form more or less naturally: first, a few companies come and then step by step it develops into a functioning cluster. Respondent B mentioned that this is precisely why he warned governments that asked him how to develop a high-tech cluster: "*You cannot build up anything when there is not already something there. It needs to have grown naturally, there must be a certain basis. That is important in cluster formation."* Clusters may not develop very predictably, but that does not mean that they will develop anywhere. They emerge because of a combination of locational qualities, network effect and some luck. Respondent B also mentioned that the branding and marketing of a location towards a particular sector can be effective, provided that the basic locational demands are fulfilled and there is some presence of that sector in the area.

In contrast to many scientific articles, the respondents also assign importance to the role of non-economic factors. Especially respondent A, who mostly works with small companies moving over short distances, experiences that a lot of companies do not make a very thorough decision, because *"entrepreneurs are really busy with their core business. For all side tasks, including making a good consideration for relocation, they do not give themselves the time. […] [B]roadly speaking many entrepreneurs make decisions like that very impulsively."* He describes one instance in which an entrepreneur from a neighbouring business park called him to ask about a plot which was on sale. This entrepreneur had noticed a billboard alongside the motorway and wanted to buy the plot because he was located in a nearby city and needed room for expansion. He did not do research and had not considered other locations. Respondent A is critical of this approach to relocation:

> [M]oving is rather impactful and brings a lot of costs. They are decisions you only make once and then you are in that new place for 10 years on average. So, many companies do this a little impulsively, few companies do take this up very extensively and when they do, it is usually with a broker.

Respondent A states that in his experience, companies do not behave fully rational. Similar viewpoints are expressed by the other respondents. Respondent C for instance experiences that companies relocate to the Zuidas because they have the feeling that they "*are second class financial service firms"* when they locate somewhere else in Amsterdam. He identifies a psychological factor in clustering, which he considers analogous with choosing a restaurant:
"[W]*hen you walk into a restaurant and nobody is eating there, you do not feel comfortable to sit there. However, when there would be three or four people there, you do consider it comfortable. That is a psychological factor which also works for companies. Besides the network effects, they feel comfortable being close to companies they are in contact with, with whom they do business*". This observation reconfirms the behavioural notion of non-rationality in decision-making and also identifies a potential role for herding behaviour in the explanation of clustering.

Even though the model data and the interview results do not immediately overlap because they cover different levels of abstraction, some distinct parallels can be drawn. Firstly, in the interviews primarily agglomeration externalities, but also psychological factors such as imitation are brought up as explanations for the formation of clusters. This confirms the notion presented in the model that agglomeration and herding together create the strongest clusters. On the other hand, the observation that clusters can only form on location with the necessary qualities, such as amenities and accessibility is also present in the model. Lastly, both the model and the interviews show that the formation of clusters is somewhat unpredictable and erratic. One cannot in advance tell where clusters will develop.

These interviews show that the outcomes of the model produced in this study roughly correspond to the actual processes of firms relocations. The overarching patterns in the model are thus substantiated by real-world experiences. Because the model outcomes and the interview show similar phenomena, the credibility of the model is enhanced.

4.3. Discussion

In this section, I interpret the model data and use them to formulate answers to the research questions posited in the introduction. It is important to keep in mind that this thesis does not prove nor disprove the presence of agglomeration externalities or herd behaviour. Instead, the ABM investigates to what extent these processes can explain cluster formation and/or consistency, provided they exist. In this section, the outcomes from the model are connected with the interview data to formulate and answer to the research questions within the framework of the theory discussed.
To make a comparison, the base scenario is discussed first. This base results in rather predictable results: there is no clustering and the number of companies in each cell is proportionate to its attractiveness. The only way in which companies affect each other location choice is through price. This is equivalent to a situation in which companies make their relocation decision regardless of the locations of other companies. This very trivial and commonsensical. This is because the variant of the model is supposed to be a reflection of the real world and an uncontroversial starting place. This base scenario is thus useful as a foundation for the herding and agglomeration scenarios.

4.3.1. Question 1

To what extent could clustering patterns be explained by herd behaviour dynamics instead of or in addition to agglomeration externalities?

The standard explanation for the formation of sectoral clustering is the presence of agglomerative externalities. The model data indeed show that the more important agglomeration is, the more clustering occurs. However, the association between the importance of agglomeration externalities and clustering is not linear. There appears to be some sort of tipping point after which clustering rapidly increases with higher values of agglomeration weight. This tipping point reflects the moment where individual groups of companies are large enough to set the clustering process in motion. As Menzel and Fornahl (2009) and Porter (1998) have described, clusters need a 'critical mass' to create the centripetal force that pulls companies towards it. Also in the interviews, this critical mass and pre-existing network of companies are identified as necessary for the development of a cluster. With small values of agglomeration, this critical size is never reached, so profound clustering does not occur. For agglomeration values higher than this tipping point, the degree of clustering rises because it creates a feedback loop in which each move to the cluster makes the cluster more attractive for other companies. The more important agglomeration externalities become, the smaller the concentrations of companies necessary to spur the creation of clusters. This critical mass creates a situation in which the probability of a new company entering a cluster is much larger than the probability of a company leaving for one sector but not for the other sectors. This happens as a result of a combination of four forces: locational attractiveness, random chance events, agglomeration weight and price. If a place has good locational qualities, it attracts businesses. Initial small groups of companies come together because of the quality of a location or because they just happen to be close together. This makes these places slightly more attractive to companies in the same sector because of agglomeration externalities. If this effect is strong enough, it keeps drawing in more companies from that sector, while increasing prices drive out companies from unrelated sectors. The way clustering works in this model is broadly equivalent to the way it is described in traditional cluster studies, such as Maskell & Malmberg (2007) and Menzel and Fornahl (2009).

An alternative approach in explaining cluster formation is herd behaviour. The results here show quite decisively that herding behaviour, as it is defined in this study, does not cause any substantial clustering. It is important to note here that herd behaviour in this instance only means that companies are more likely to encounter or consider an available location if there are other related companies located in the vicinity. Herding does not imply that they consider these places more attractive. This process slightly enhances the degree of clustering in the system, but this does not have the power to explain the emergence of sectoral clusters.

However, when imitation is included in addition to agglomeration economies, it works as a catalyst. Herding causes the tipping point in the model to occur at much lower levels of agglomeration weights. Moreover, it creates a much stronger clustering effect after the tipping point. Herding increases the power of the feedback loop, which results in a considerably higher degree of clustering. Imitation does not by itself radically alter the economic landscape, but it interferes with the feedback loop of agglomeration externalities and firm concentrations. This happens in such a way that the critical mass for cluster formation emerges more rapidly and that stronger clusters emerge.

As a result, the probability of new companies entering a cluster location rises even when it is still rather small. This dynamic lowers the level of agglomeration externalities that is necessary to spur clustering. This observation in itself is not standard in economic geography, but it is a logical extension of different bits of knowledge: psychological factors impel companies to seek places close to other companies and economic forces create benefits for economic concentration. It is therefore not illogical to hypothesize that herding behaviour magnifies agglomeration externalities; a hypothesis supported by this simulation. In the absence of herding, agglomeration externalities needed to be many times more important to explain the formation of a cluster. The development of clusters in this simulation is similar to Vicente & Suire's (2007) concept of the locational norm. They show that default locations for a sector can emerge, because of the centripetal force of imitation, even if that particular location is not disproportionately beneficial for that sector.

Thus, if concentration is for, any reason, considered a positive asset of a location and if companies are more likely to consider locations with many related companies, this creates a powerful cumulative, circular-causal development out of which strong clusters can emerge, even if the herding and agglomeration are relatively mild. Thus herding creates strong clusters, but only if it functions in tandem with some agglomeration externalities.

In a more general form, it shows that the way in which the consideration set is formed has the potential to fundamentally alter how the relocation system works if it interacts with the evaluation of potential locations.

4.3.2. Question 2

How predictable are spatial distribution patterns of firms in a herding-based location choice process?

This question does not relate to the actual distribution of companies itself, but to the consistency of the patterns that form across runs. If a given model setup consistently produces the same outcomes, then the model is very predictable. If, on the other hand, the results vary from run to run the model is less predictable. If such variation occurs, this means that the bounded-rationality and random-chance events in the model create different paths in which the system can develop. To answer this research question, I consider two facets of this issue:

1) how consistent is the degree of clustering?;
2) how consistent is the location of clustering?.

The answer to the first question is relatively clear. The degree of clustering can be very stable. This occurs when clustering is either virtually absent or really strong. There are model setups in which clustering never occurs and setups in which complete sectoral clustering always happens. However in the situations in between the two, the deviation of individual runs from the average can be quite substantial. This indicates that the degree of clustering in a run can be markedly different even when it is produced under the same circumstances. The differences between the different herding scenarios in this regard are not very notable, which indicates that this variation in clustering is just a result of the behavioural constraints that were put on the model.

The answer to the second question may be slightly more complex. Here we see some predictability in the model. Clearly, unattractive locations never contain a cluster. The probability that a cluster develops on a location increases with the attractiveness of the location. However, the precise location of an individual cluster is not predictable, especially when the attractiveness of the locations is similar. In these situations, clustering can occur on moderately ranked locations, while superior locations remain relatively empty. This means that a cluster will not always develop on the best location, but it never develops in an inadequate place. This vision on herding is also shared by respondents to the interviews, who stress that companies will never cluster on places where the basic locational demands such as accessibility are not met.

The distribution of clusters across different locations does not seem to be affected by either herding behaviour externalities. This means that predicting where clusters will emerge is difficult, regardless of herding or agglomeration. In this respect, it matches existing scientific knowledge about cluster formation. The forces of path-dependency create unpredictable outcomes in clustering patterns. This means that it is impossible to ex-ante determine whether and where clusters will form (Maskell & Malmberg, 2007), but it is possible to rule out certain places. In effect, this is what the respondents in the interview say. They can identify prerequisites for cluster formation, but whether clustering does occur on a given location is something time will tell.

Herd behaviour would not make the model less or more predictable. It shows that even in a situation with an analyst who has all information, predicting clustering and cluster locations is not possible, because of the stochasticity in firm relocation decisions, and the cumulative, circular-causal effects of relocation on firm distributions. This is however not a result of imitation but is a property of the decision framework itself.

5. CONCLUSION

This study has investigated the development of clusters in different scenarios using an agent-based model. It seeks to combine insights from different branches of economic geography to develop a more complete view on the way clusters form. The main insights from behavioural and neoclassical economic geography have been combined to form a general model of relocation choice. This model is used to approximate how different elements of the decision process affect clustering. One of these elements is the economic rationale behind clustering: agglomeration economies are widely credited as the driving force behind clustering. This supposes that companies decide to relocate to a cluster because this yields economic benefits for them.

Others have criticized this approach because it assumes a psychology which decision-makers do not exhibit. One of the alternative explanations is that companies concentrate because of psychological rather than economic reasons. This study challenges the traditional view on clustering and posits that the psychological forces of imitation can play a decisive role in the formation of clusters in an urban context. This study suggests that ignoring the psychology of decision-making may cause considerable overestimation of the importance of agglomeration externalities in the dominant scientific discourse. This study investigated what effect herding behaviour can have on the distribution of companies.

> *To what extent can herd behaviour in location choice processes be a cause for spatial clustering of related office firms on an urban scale?*

Herd behaviour in this study is defined as any type of behaviour that increases the probability that locations with a concentration of related companies are considered, regardless of the qualities of that place. This study shows that this type of behaviour cannot explain the emergence of large clusters, but that the imitation can strongly exacerbate the power of agglomeration externalities. Clusters start to form after a critical mass is reached. Herd behaviour creates a system in which this critical mass is reached easier. Imitation can, in tandem with agglomeration externalities, explain why clusters form. This happens in such a way that the power of agglomeration externalities necessary to explain the formation of clusters is much lower in the situation in which herd behaviour is involved. If companies exhibit imitation, clusters can form with relatively little agglomeration externalities. This challenges the dominant way of thinking in economic geography. It shows that in theory agglomeration externalities might be overvalued in analysis if analysts ignore psychological forces like herd behaviour. The conclusions of this thesis bring together several different criticisms on the standard view on cluster formation. The current understanding of clusters through agglomeration externalities is not uncontroversial with many scholars raising questions on the scientific rigour of the theory (Maskell & Malmberg, 2002). Researchers such as Boschma (2005), Gordon & McCann (2005) and Klepper (2010) show empirically that agglomeration externalities cannot completely explain the formation of clusters, while others have stressed the role of imitation in decision making (Berg, 2010; Vicente & Suire, 2009). the presence of imitation in relocation decisions is expectable because companies make boundedly rational decisions, have limited information and are susceptible to influence from others. This study suggests that a combined effect of small amounts of both herding and agglomeration externalities has the potential to create sectoral clusters, even if both effects are quite limited. It thus shows that ignoring the psychological side of decision-making can lead to substantial overestimation of the importance of agglomeration externalities and a misunderstanding of the decision-making processes involved in relocation.

One could expect that the erratic nature of herding would increase the unpredictability of the model and cause cluster formation on illogical locations, but this is not the case. The locations of clusters are not less predictable when imitation is present in a simulation. In all simulations, the probability of a cluster forming on a location increases with the attractiveness of that location. Herd behaviour does not change this.

This study does not prove the existence of herd behaviour in the real world. Instead, it shows its potential power in explaining the formation of sectoral clusters, if companies exhibit imitation. Whether herd behaviour exists in location choice is not a topic for this study, but previous research gives reason to hypothesize that it plays a role in relocation decisions, as it does in other economic processes. It is, however, unclear to what extent imitation can explain concentrations of companies.

The results from this thesis challenge the dominant theory about cluster formation in economic geography because it shows that agglomeration economies may be overvalued if imitation is ignored. This gives reason for further empirical research into the role of herd behaviour and other psychological factors in relocation decisions. It should here be noted that the model is designed with two important constraints to its universality. It is designed based on office firms relocating on an urban scale. Office firms can relocate rather easily compared to industrial firms and can relocate to many more places because they need less space and are less impactful for the surrounding areas. This model only reviews relocations within an urban area. This means that the distance between the targets and the current location is not crucial. Besides, the assumption of a random selection of the consideration set only holds on lower levels of scale.

The model is not empirical, but it presents a credible view of decision framework companies use. It is, of course, a simplification, but it is firmly rooted in a scientific framework. Moreover, the outcomes the model reflect the way clusters form in real life. That does not mean that the model is an accurate representation of the real world. It must be read and interpreted within the context of the definitions and assumptions of the model. The assumptions used in this model have been outlined. These make the model manageable and interpretable, but they also simplify reality. Many important factors are left out of consideration in this study to create a general view. This includes things like company size and age. An influential idea by Klepper (2010) regarding the role of entry and exit in the formation of clusters is left outside of the consideration. Moreover, the method used here is useful in exploring and comparing the dynamics of multiple different scenarios, but it is also vulnerable to biases and misrepresentations. Nevertheless, the model provides a credible hypothesis for the emergence of sectoral clusters based on a computer simulation of individual relocation decisions.

5.1. Implications for research and policy

The lack of empirical testing and the coarseness of the model design makes it difficult to formulate definitive conclusions but gives reason for further investigation into this topic. This study demonstrates that herd behaviour can have a large effect on the spatial distribution of firms, but it does not prove that it does. Consequently, several different suggestions for subsequent research come from this thesis. Firstly, further empirical testing is necessary into the presence of herd behaviour in real-life location decisions, and its effect on the distribution of companies. Secondly, this research shows that herd behaviour can affect location choice, but there are several other ways in which companies making decision shortcuts. Well established in, for instance, the role of inertia in the cluster life cycle (Hervas-Oliver & Albors-Garrigos, 2014; Maskell & Malmberg, 2007). Because rational optimizing is still the most prevalent theoretical assumption in relocation studies, other behavioural processes like this may be underexposed. Thirdly, the model itself can be further elaborated. The impact of entry-and-exit dynamics may be relevant, as well as differentiating sizes, researching spatial levels other than the urban scale, other companies than offices, etc. This model is intentionally kept simple. Including all these extra elements would render it unintelligible. That does however not mean that these factors are irrelevant.

Besides suggestions for further research, this study also presents several implications for policymakers. Firstly, it shows that neither herd behaviour nor agglomeration externalities can cause clusters to form in unattractive places. If governments want to create a vibrant business environment, they should first and foremost try to make that place very attractive. When this is the case, they should look at companies in sectors with a strong presence in the area. Attracting companies from totally unrelated industries would still be difficult. This study shows the importance of considerations sets in relocation choice. If psychological factors are important in decision-making, governments may also use marketing and placemaking strategies to attract businesses, besides a focus on economic advantages. This might be effective in pulling companies in, provided all basic locational demands are met. However, this study also shows that cluster formation is inherently unpredictable, which makes cluster-based economic policies always risky for governments since its outcomes will be unsure.

BIBLIOGRAPHY

Ajraldi, V., Pittavino, M., & Venturino, E. (2011). Modeling herd behavior in population systems. *Nonlinear Analysis: Real World Applications*, *12*(4), 2319–2338. https://doi.org/10.1016/j.nonrwa.2011.02.002

Alós-Ferrer, C., Hügelschäfer, S., & Li, J. (2016). Inertia and decision making. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00169

Anas, A., Arnott, R., & Small, K. A. (1998). Urban spatial structure. *Journal of Economic Literature*, *36*(3), 1426–1464. https://www.jstor.org/stable/2564805

Armstrong, M., & Huck, S. (2010). Behavioral economics as applied to firms: A primer. In *Oxford University Research Archive*. Office of Fair Trading. https://ora.ox.ac.uk/objects/uuid:b238866d-08d8-4618-9923-b7a1128b8522/

Baddeley, M. (2010). Herding, social influence and economic decision-making: Socio-psychological and neuroscientific analyses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1538), 281–290. https://doi.org/10.1098/rstb.2009.0169

Balbontin, C., & Hensher, D. A. (2019). Firm-specific and location-specific drivers of business location and relocation decisions. *Transport Reviews*, *39*(5), 569–588. https://doi.org/10.1080/01441647.2018.1559254

Banerjee, A. V. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics*, *107*(3), 797–817. https://doi.org/10.2307/2118364

Barrios, S., Görg, H., & Strobl, E. (2006). Multinationals' location choice, agglomeration economies, and public incentives. *International Regional Science Review*, *29*(1), 81–107. https://doi.org/10.1177/0160017605281516

Beaudry, C., & Schiffauerova, A. (2009). Who's right, Marshall or Jacobs? The localization versus urbanization debate. *Research Policy*, *38*(2), 318–337. https://doi.org/10.1016/j.respol.2008.11.010

Berg, N. (2014). Success from satisficing and imitation: Entrepreneurs' location choice and implications of heuristics for local economic development. *Journal of Business Research*, *67*(8), 1700–1709. https://doi.org/10.1016/j.jbusres.2014.02.016

Berg, N., & Gigerenzer, G. (2010). As-if behavioral economics: Neoclassical economics in disguise? *History of Economic Ideas*, *18*(1), 133–165.

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1998). Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades. *Journal of Economic Perspectives*, *12*(3), 151–170. https://doi.org/10.1257/jep.12.3.151

Bodenmann, B. R., & Axhausen, K. W. (2012). Destination choice for relocating firms: A discrete choice model for the St. Gallen region, Switzerland. *Papers in Regional Science*, *91*(2), 319–341. https://doi.org/10.1111/j.1435-5957.2011.00389.x

Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences*, *99*(Supplement 3), 7280–7287. https://doi.org/10.1073/pnas.082080899

Boschma, R. A. (2005). Proximity and innovation: A critical assessment. *Regional Studies*, *39*(1), 61–74. https://doi.org/10.1080/0034340052000320887

Boschma, R. A., & Frenken, K. (2006). Why is economic geography not an evolutionary science? Towards an evolutionary economic geography. *Journal of Economic Geography*, *6*(3), 273–302. https://doi.org/10.1093/jeg/lbi022

Brekke, T. (2015). Entrepreneurship and path dependency in regional development. *Entrepreneurship & Regional Development*, *27*(3–4), 202–218. https://doi.org/10.1080/08985626.2015.1030457

Breschi, S., & Malerba, F. (2001). The Geography of Innovation and Economic Clustering: Some Introductory Notes. *Industrial and Corporate Change*, *10*(4), 817–833. https://doi.org/10.1093/icc/10.4.817

Brouwer, A. E., Mariotti, I., & van Ommeren, J. N. (2004). The firm relocation decision: An empirical investigation. *The Annals of Regional Science*, *38*(2), 335–347. https://doi.org/10.1007/s00168-004-0198-5

Brunes, F. (2005). Overbuilding in office markets, a behavioural approach to investigate possible explanations. In *Overbuilding in office markets: Are behavioural aspects important?* (pp. 28–68). Royal Institute of Technology (KTH).

Buenstorf, G., & Klepper, S. (2010). Why does entry cluster geographically? Evidence from the US tire industry. *Journal of Urban Economics*, *68*(2), 103–114. https://doi.org/10.1016/j.jue.2010.03.005

Chrisman, J. J., Memili, E., & Misra, K. (2014). Nonfamily managers, family firms, and the winner's curse: The influence of noneconomic goals and bounded rationality. *Entrepreneurship Theory and Practice*, *38*(5), 1103–1127. https://doi.org/10.1111/etap.12014

Cipriani, M., & Guarino, A. (2009). Herd behavior in financial markets: An experiment with financial market professionals. *Journal of the European Economic Association*, *7*(1), 206–233. https://doi.org/10.1162/jeea.2009.7.1.206

Cohen, A. R., Gadon, H., & Miaoulis, G. (1976). Decision making in firms: The impact of noneconomic factors. *Journal of Economic Issues*, *10*(2), 242–258.

Conlisk, J. (1996). Why Bounded Rationality? *Journal of Economic Literature*, *34*(2), 669–700.

Cyert, R. M., & Hedrick, C. L. (1972). Theory of the firm: Past, present, and future; an interpretation. *Journal of Economic Literature*, *10*(2), 398–412.

De Bok, M., & Van Oort, F. (2011). Agglomeration economies, accessibility and the spatial choice behavior of relocating firms. *Journal of Transport and Land Use*, *4*(1), 5. https://doi.org/10.5198/jtlu.v4i1.144

Delgado, M., Porter, M. E., & Stern, S. (2014). Clusters, convergence, and economic performance. *Research Policy*, *43*(10), 1785–1799. https://doi.org/10.1016/j.respol.2014.05.007

Diodato, D., Neffke, F., & O'Clery, N. (2018). Why do industries coagglomerate? How Marshallian externalities differ by industry and have evolved over time. *Journal of Urban Economics*, *106*, 1–26. https://doi.org/10.1016/j.jue.2018.05.002

Dugger, W. M. (1979). Methodological differences between institutional and neoclassical economics. *Journal of Economic Issues*, *13*(4), 899–909. https://doi.org/10.1080/00213624.1979.11503711

Duranton, G. (2011). California dreamin': The feeble case for cluster policies. *Review of Economic Analysis*, *3*(1), 3–45.

Duranton, G., & Puga, D. (2004). Micro-foundations of urban agglomeration economies. *Handbook of Regional and Urban Economics*, *4*, 2063–2117. https://doi.org/10.1016/s1574-0080(04)80005-1

Edmans, A. (2009). Blockholder trading, market efficiency, and managerial myopia. *The Journal of Finance*, *64*(6), 2481–2513. https://doi.org/10.1111/j.1540-6261.2009.01508.x

Elgar, I., Farooq, B., & Miller, E. J. (2009). Modeling location decisions of office firms: Introducing anchor points and constructing choice sets in the model system. *Transportation Research Record: Journal of the Transportation Research Board*, *2133*(1), 56–63. https://doi.org/10.3141/2133-06

Elgar, I., & Miller, E. J. (2009). How office firms conduct their location search process? An analysis of a survey from the Greater Toronto area. *International Regional Science Review*, *33*(1), 60–85. https://doi.org/10.1177/0160017609331398

Ellison, G. (2006). Bounded Rationality in Industrial Organization. In R. Blundell, W. K. Newey, & T. Person (Eds.), *Advances in Economics and Econometrics Theory and Applications, Ninth World Congress* (pp. 142–174). Cambridge University Press.

Ellison, G., & Glaeser, E. L. (1997). Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy*, *105*(5), 889–927. https://doi.org/10.1086/262098

Ellison, G., & Glaeser, E. L. (1999). The geographic concentration of industry: Does natural advantage explain agglomeration? *American Economic Review*, *89*(2), 311–316. https://doi.org/10.1257/aer.89.2.311

Ellison, G., Glaeser, E. L., & Kerr, W. R. (2010). What causes industry agglomeration? Evidence from coagglomeration patterns. *American Economic Review*, *100*(3), 1195–1213. https://doi.org/10.1257/aer.100.3.1195

Elola, A., Valdaliso, J. M., López, S. M., & Aranguren, M. J. (2012). Cluster life cycles, path dependency and regional economic development: Insights from a meta-study on Basque clusters. *European Planning Studies*, *20*(2), 257–279. https://doi.org/10.1080/09654313.2012.650902

Fingleton, B., & Fischer, M. M. (2010). Neoclassical theory versus new economic geography: competing explanations of cross-regional variation in economic development. *The Annals of Regional Science*, *44*(3), 467–491. https://doi.org/10.1007/s00168-008-0278-z

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Shleifer, A. (1992). Growth in Cities. *Journal of Political Economy*, *100*(6), 1126–1152. https://doi.org/10.1086/261856

Glaeser, E. L., Kolko, J., & Saiz, A. (2001). Consumer city. *Journal of Economic Geography*, *1*(1), 27–50. https://doi.org/10.1093/jeg/1.1.27

Gordon, I. R., & McCann, P. (2005). Innovation, agglomeration, and regional development. *Journal of Economic Geography*, *5*(5), 523–543. https://doi.org/10.1093/jeg/lbh072

Graf, H., & Broekel, T. (2020). A shot in the dark? Policy influence on cluster networks. *Research Policy*, *49*(3), 103920. https://doi.org/10.1016/j.respol.2019.103920

Hassink, R., & Gong, H. (2019). New Economic Geography. *The Wiley Blackwell Encyclopedia of Urban and Regional Studies*. https://doi.org/10.1002/9781118568446.eurs0222

Haynes, K. T., Josefy, M., & Hitt, M. A. (2015). Tipping point: Managers' self-interest, greed, and altruism. *Journal of Leadership & Organizational Studies*, *22*(3), 265–279. https://doi.org/10.1177/1548051815585171

Heijman, W. J. M., & Schipper, R. A. (2010). *Space and economics: An introduction to regional economics: Vol. 7 Mansholt Publication Series*. Wageningen Academic Publishers.

Hervas-Oliver, J.-L., & Albors-Garrigos, J. (2014). Are technology gatekeepers renewing clusters? Understanding gatekeepers and their dynamics across cluster life cycles. *Entrepreneurship & Regional Development*, *26*(5–6), 431–452. https://doi.org/10.1080/08985626.2014.933489

Hospers, G.-J., & Beugelsdijk, S. (2002). Regional cluster policies: Learning by comparing? *Kyklos*, *55*(3), 381–402. https://doi.org/10.1111/1467-6435.00192

Huang, A., & Levinson, D. (2011). Why retailers cluster: An agent model of location choice on supply chains. *Environment and Planning B: Planning and Design*, *38*(1), 82–94. https://doi.org/10.1068/b36018

Jacobs, J. (1969). *The Economy of the cities*. Random House.

Kaplan, S., & Henderson, R. (2005). Inertia and Incentives: Bridging Organizational Economics and Organizational Theory. *Organization Science*, *16*(5), 509–521. https://doi.org/10.1287/orsc.1050.0154

Keeble, D., & Nachum, L. (2002). Why do business service firms cluster? Small consultancies, clustering and decentralization in London and southern England. *Transactions of the Institute of British Geographers*, *27*(1), 67–90. https://doi.org/10.1111/1475-5661.00042

Keynes, J. M. (1936). *The general theory of employment, interest and money*. Macmillan Cambridge University Press.

Klepper, S. (2010). The origin and growth of industry clusters: The making of Silicon Valley and Detroit. *Journal of Urban Economics*, *67*(1), 15–32. https://doi.org/10.1016/j.jue.2009.09.004

Klier, T. H. (2006). Where the Headquarters Are: Location Patterns of Large Public Companies, 1990-2000. *Economic Development Quarterly*, *20*(2), 117–128. https://doi.org/10.1177/0891242405286158

Koetsier, I., & Bikker, J. A. (2017). *Herding Behaviour of Dutch Pension Funds in Sovereign Bond Investments*. De Nederlandsche Bank. https://www.dnb.nl/binaries/Working%20Paper%20No.%20569_tcm46-362997.pdf?2020021113

Koster, H. R. A., van Ommeren, J., & Rietveld, P. (2013). Is the sky the limit? High-rise buildings and office rents. *Journal of Economic Geography*, *14*(1), 125–153. https://doi.org/10.1093/jeg/lbt008

Krugman, P. (1991). *Geography and trade*. Leuven University Press/MIT Press.

Larwood, L., & Whittaker, W. (1977). Managerial myopia: Self-serving biases in organizational planning. *Journal of Applied Psychology*, *62*(2), 194–198. https://doi.org/10.1037/0021-9010.62.2.194

Luo, B., & Lin, Z. (2011). A decision tree model for herd behavior and empirical evidence from the online P2P lending market. *Information Systems and E-Business Management*, *11*(1), 141–160. https://doi.org/10.1007/s10257-011-0182-4

Lux, T. (1995). Herd Behaviour, Bubbles and Crashes. *The Economic Journal*, *105*(July), 881–896. https://doi.org/10.2307/2235156

MacKinnon, D. (2008). Evolution, Path Dependence and Economic Geography. *Geography Compass*, *2*(5), 1449–1463. https://doi.org/10.1111/j.1749-8198.2008.00148.x

Malmberg, A., & Maskell, P. (2002). The elusive concept of localization economies: Towards a knowledge-based theory of spatial clustering. *Environment and Planning A: Economy and Space*, *34*(3), 429–449. https://doi.org/10.1068/a3457

Maltby, H., Pakornrat, W., Jackson, J., Hernandez, A., Williams, C., Lin, C., & Khim, J. (n.d.). *Markov Chains*. Brilliant. Retrieved June 11, 2020, from https://brilliant.org/wiki/markov-chains/

Markusen, A. (2003). Fuzzy concepts, scanty evidence, policy distance: The case for rigour and policy relevance in critical regional studies. *Regional Studies*, *37*(6–7), 701–717. https://doi.org/10.1080/0034340032000108796

Marshall, A. (1890). Industrial organization continued: The concentration of specialized industries in particular localities. In *Principles of Economics*. Macmillan and Co. https://www.econlib.org/library/Marshall/marP.html?chapter_num=25#book-reader

Martin, R., & Sunley, P. (2003). Deconstructing clusters: chaotic concept or policy panacea? *Journal of Economic Geography*, *3*(1), 5–35. https://doi.org/10.1093/jeg/3.1.5

Martin, R., & Sunley, P. (1996). Paul Krugman's geographical economics and its implications for regional development theory: A critical assessment. *Economic Geography*, *72*(3), 259. https://doi.org/10.2307/144401

Maskell, P., & Malmberg, A. (2007). Myopia, knowledge development and cluster evolution. *Journal of Economic Geography*, *7*(5), 603–618. https://doi.org/10.1093/jeg/lbm020

Menzel, M.-P., & Fornahl, D. (2009). Cluster life cycles--dimensions and rationales of cluster evolution. *Industrial and Corporate Change*, *19*(1), 205–238. https://doi.org/10.1093/icc/dtp036

Milidonis, A., & Stathopoulos, K. (2014). Managerial incentives, risk aversion, and debt. *Journal of Financial and Quantitative Analysis*, *49*(2), 453–481. https://doi.org/10.1017/s0022109014000301

Nguyen, C. Y., Sano, K., Tran, T. V., & Doan, T. T. (2013). Firm relocation patterns incorporating spatial interactions. *The Annals of Regional Science*, *50*(3), 685–703. https://doi.org/10.1007/s00168-012-0523-3

Porter, M. E. (1998). Location, clusters, and the "new" microeconomics of competition. *Business Economics*, *33*(1), 7–13.

Posen, H. E., Lee, J., & Yi, S. (2012). The power of imperfect imitation. *Strategic Management Journal*, *34*(2), 149–164. https://doi.org/10.1002/smj.2007

Rao, K., & Tilt, C. (2016). Board composition and corporate social responsibility: The role of diversity, gender, strategy and decision making. *Journal of Business Ethics*, *138*(2), 327–347. https://doi.org/10.1007/s10551-015-2613-5

Rook, L. (2006). An Economic Psychological Approach to Herd Behavior. *Journal of Economic Issues*, *40*(1), 75–95. https://doi.org/10.1080/00213624.2006.11506883

Seiler, M. J., Lane, M. A., & Harrison, D. M. (2012). Mimetic herding behavior and the decision to strategically default. *The Journal of Real Estate Finance and Economics*, *49*(4), 621–653. https://doi.org/10.1007/s11146-012-9386-8

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, *69*(1), 99. https://doi.org/10.2307/1884852

Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2004). Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality. *Risk Analysis*, *24*(2), 311–322. https://doi.org/10.1111/j.0272-4332.2004.00433.x

Stein, J. C. (1988). Takeover threats and managerial myopia. *Journal of Political Economy*, *96*(1), 61–80. https://doi.org/10.1086/261524

Strudler, A., & Curlo, E. (1996). The moral assessment of corporate decision making. *Philosophy and Public Policy*, *16*(3/4), 1–5. http://ojs2.gmu.edu/PPPQ/article/viewFile/245/174

Suire, R., & Vicente, J. (2009). Why do some places succeed when others decline? A social interaction model of cluster viability. *Journal of Economic Geography*, *9*(3), 381–404. https://doi.org/10.1093/jeg/lbn053

Sydow, J., Lerch, F., & Staber, U. (2010). Planning for path dependence? The case of a network in the Berlin-Brandenburg optics cluster. *Economic Geography*, *86*(2), 173–195. https://doi.org/10.1111/j.1944-8287.2010.01067.x

Torrens, P. M. (2010). Agent-based models and the spatial sciences. *Geography Compass*, *4*(5), 428–448. https://doi.org/10.1111/j.1749-8198.2009.00311.x

Van Dijk, J., & Pellenbarg, P. H. (2000). Firm relocation decisions in The Netherlands: An ordered logit approach. *Papers in Regional Science*, *79*(2), 191–219. https://doi.org/10.1007/s101100050043

Van Witteloostuijn, A. (1998). Bridging behavioral and economic theories of decline: Organizational inertia, strategic competition, and chronic failure. *Management Science*, *44*(4), 501–519. https://doi.org/10.1287/mnsc.44.4.501

Vicente, J, Dalla Pria, Y., & Suïre, R. (2007). The ambivalent role of mimetic behavior in proximity dynamics: evidence from the French 'Silicon Sentier.' In J. Suriñach, R. Moreno, & E. Vayá (Eds.), *Knowledge Externalities, Innovation Clusters and Regional Development* (pp. 61–91). Edward Elgar Publishing Limited.

Vicente, J., & Suire, R. (2007). Informational cascades versus network externalities in locational choice: Evidence of 'ICT clusters' formation and stability. *Regional Studies*, *41*(2), 173–184. https://doi.org/10.1080/00343400601108424

Wang, W., Chen, Y., & Huang, J. (2009). Heterogeneous preferences, decision-making capacity, and phase transitions in a complex adaptive system. *Proceedings of the National Academy of Sciences*, *106*(21), 8423–8428. https://doi.org/10.1073/pnas.0811782106

Weizsäcker, G. (2010). Do we follow others when we should? A simple test of rational expectations. *American Economic Review*, *100*(5), 2340–2360. https://doi.org/10.1257/aer.100.5.2340

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

Willigers, J., & van Wee, B. (2011). High-speed rail and office location choices. A stated choice experiment for the Netherlands. *Journal of Transport Geography*, *19*(4), 745–754. https://doi.org/10.1016/j.jtrangeo.2010.09.002

Wolman, H., & Hincapie, D. (2014). Clusters and cluster-based development policy. *Economic Development Quarterly*, *29*(2), 135–149. https://doi.org/10.1177/0891242413517136

Zhao, L., Yang, G., Wang, W., Chen, Y., Huang, J. P., Ohashi, H., & Stanley, H. E. (2011). Herd behavior in a complex adaptive system. *Proceedings of the National Academy of Sciences*, *108*(37), 15058–15063. https://doi.org/10.1073/pnas.1105239108

Software packages

Kassambara, A. (2020). ggpubr: 'ggplot2' Based Publication Ready Plots [Software Package]. https://CRAN.R-project.org/package=ggpubr

R Core Team (2020). R: A language and environment for statistical computing [Software Package]. R Foundation for Statistical Computing. https://r-project.org/

Wickham, H. (2016). Elegant Graphics for Data Analysis [Software Package]. Springer-Verlag.

Wickham, H., & Bryan, J. (2019). readxl: Read Excel Files [Software Package]. https://CRAN.R-project.org/package=readxl

Wickham, H., François, R., Henry, L., & Müller, K. (2020). dplyr: A Grammar of Data Manipulation [Software Package]. https://CRAN.R-project.org/package=dplyr

Wickham, H., & Henry, L. (2020). tidyr: Tidy Messy Data [Software Package]. https://CRAN.R-project.org/package=tidyr.

Wilensky, U. (2019). NetLogo, [Software Package]. Version 6.1.1. Center for Connected Learning and Computer-Based Modeling. Northwestern University. http://ccl.northwestern.edu/netlogo/

6. APPENDICES

6.1. Model Code

[Omitted]

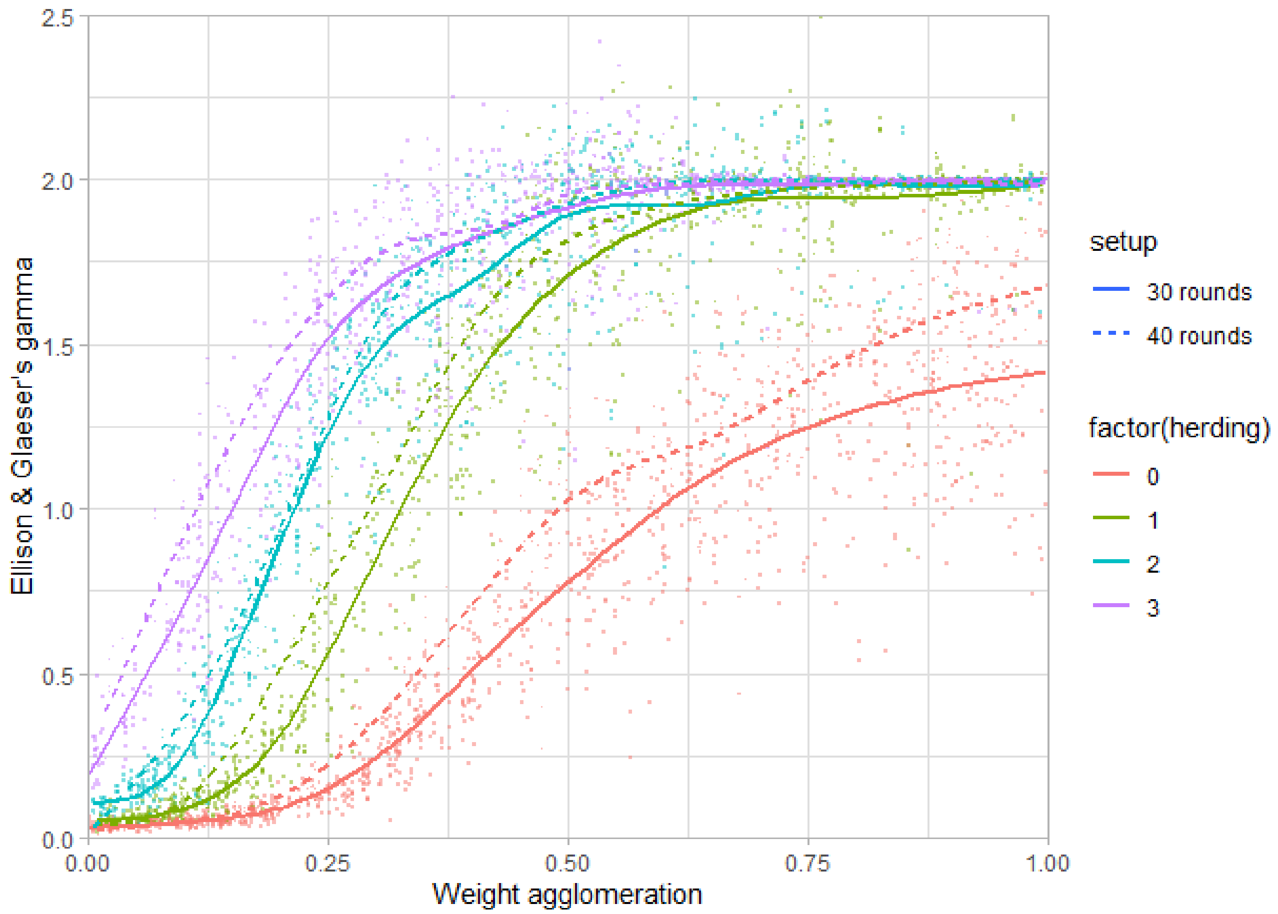
6.2. Validation models

This appendix shows the results of multiple alternative models. The graphics are made in the same way as the ones produced in section 4.1.4.

6.2.1. Number of rounds

This alternative model is generated by adjusting the number-of-turns parameter from 30 to 40.

Figure 6-1: Forty rounds model overlaid on figure 4-6 (N = 954)



6.2.2. Alternative model 1

This model is created the same parameters as the main model, but with the following adjusted parameters:

- Number-of-firms = 2000
- Number-of-turns = 20
- Weight-price = 1
- cap = 5

Figure 6-2: Sectoral clustering as a product of agglomerative externalities with herding (N = 400, includes a generalized additive model, outliers removed). Alternative model 1.

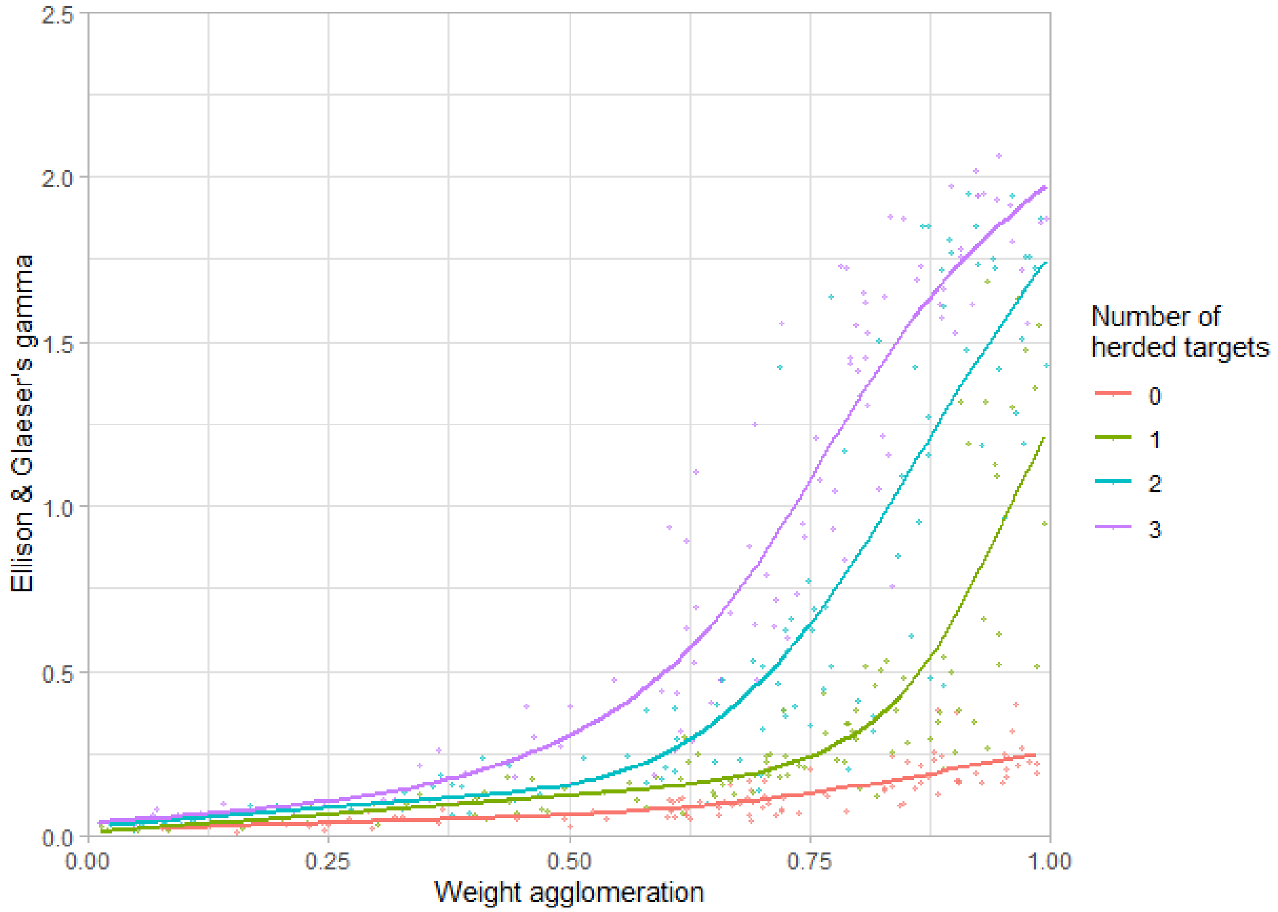
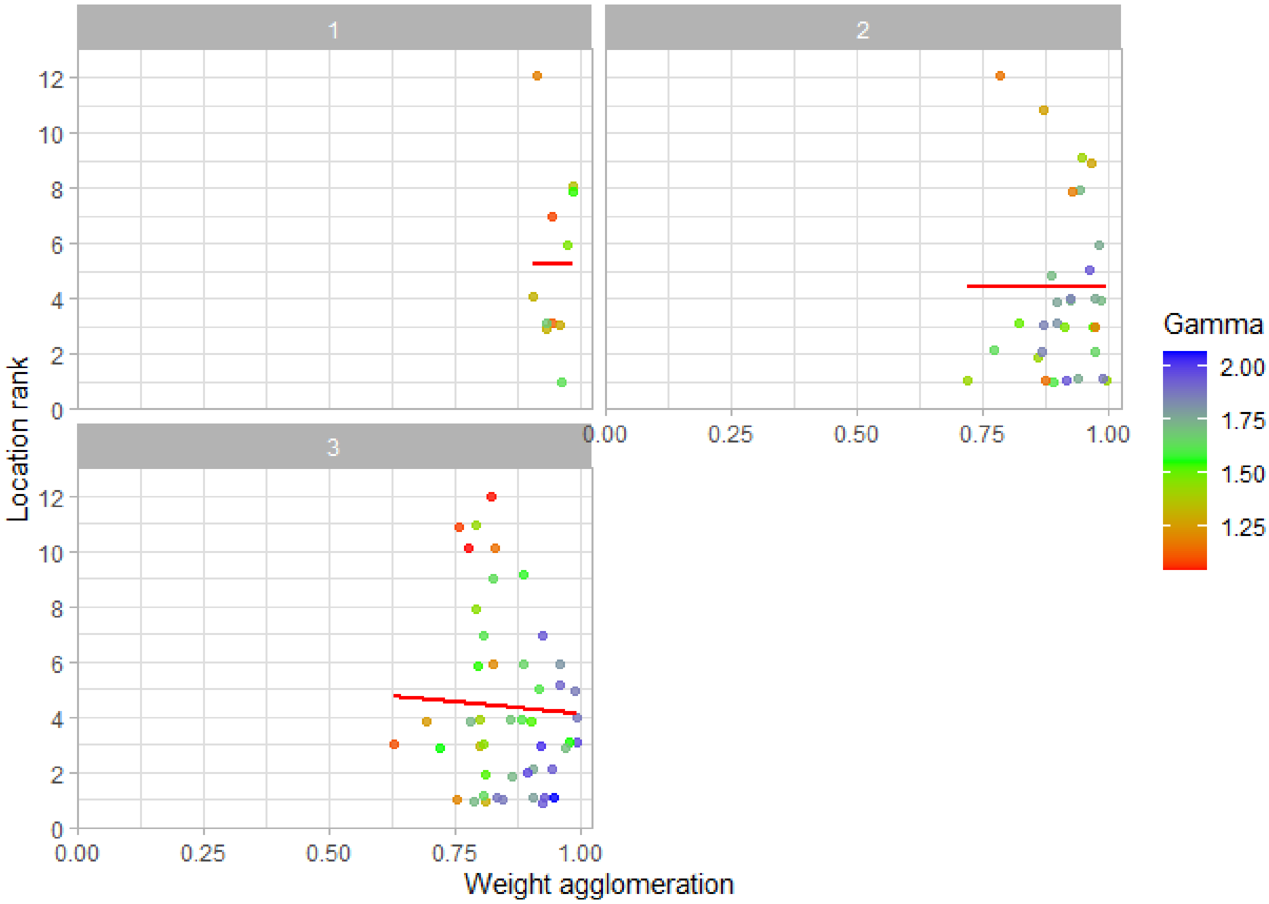


Figure 6-3: Modal cluster location rank by herding and agglomeration (N = 90). Points jittered vertically for visibility. Alternative model 1.



6.2.3. Alternative model 2

This model is created the same parameters as the main model, but with the following adjusted parameters:

- Number-of-firms = 1000
- Number-of-turns = 20
- Weight-price = 0.2
- cap = 3

Figure 6-4: Sectoral clustering as a product of agglomerative externalities with herding (N = 361, includes a generalized additive model, outliers removed). Alternative model 2.

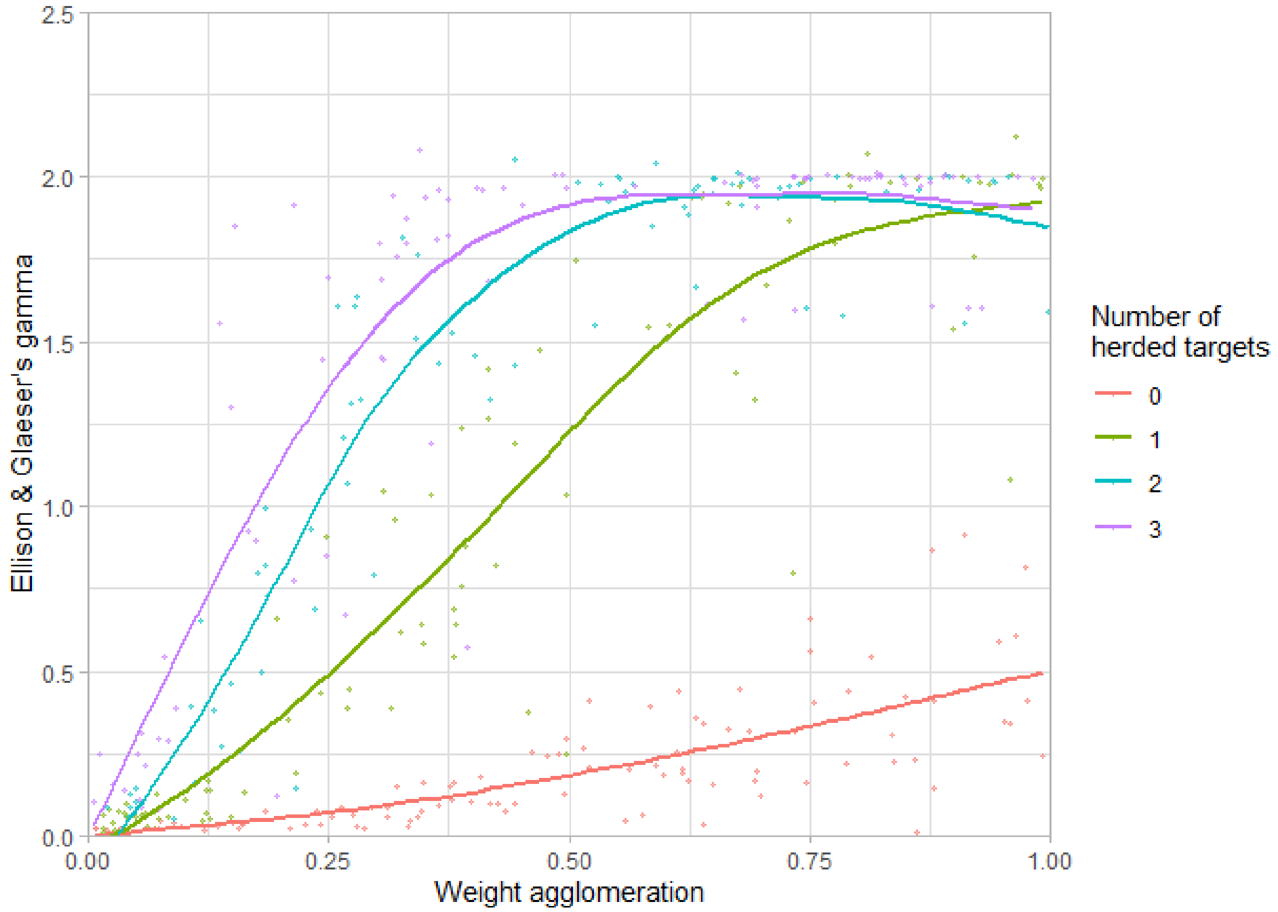
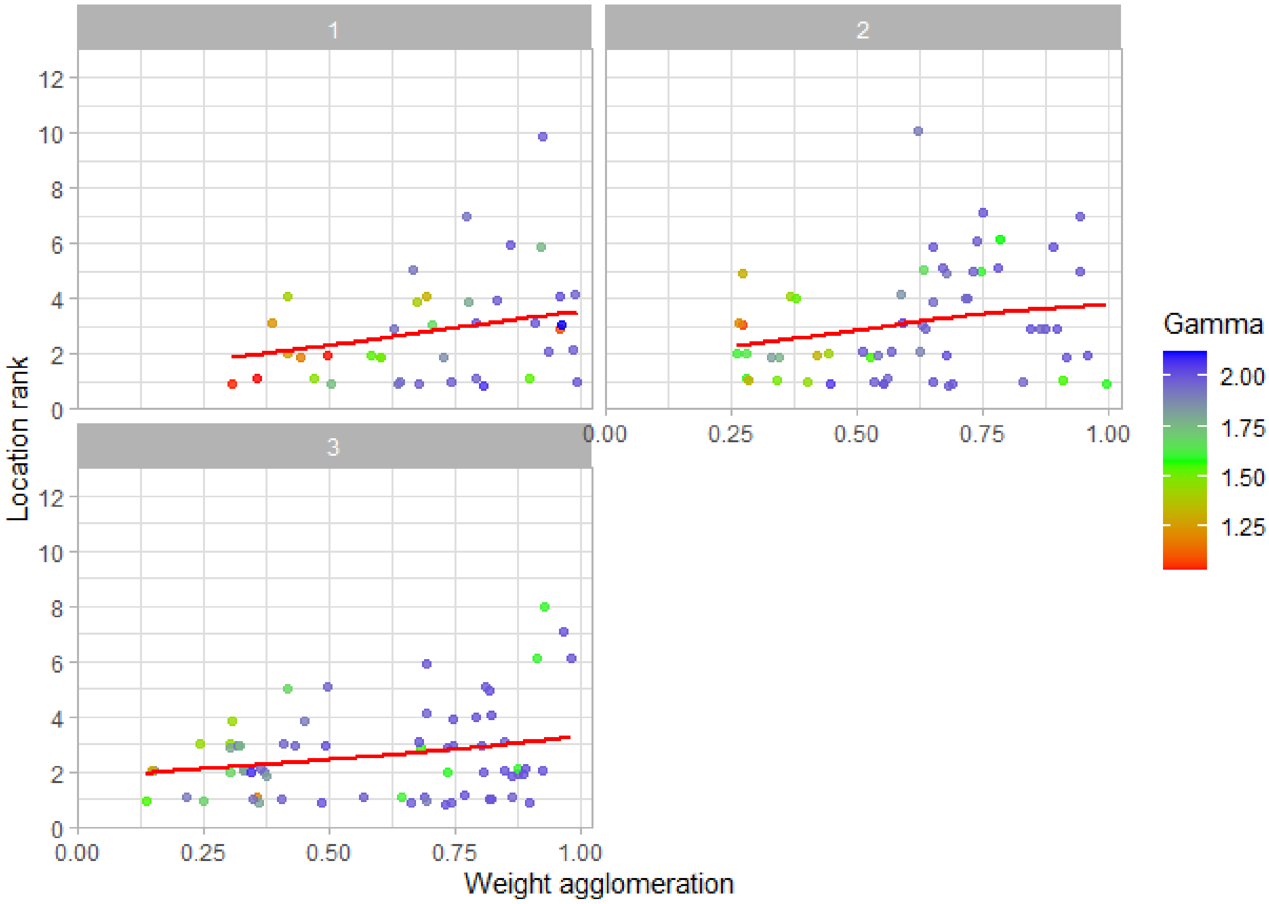


Figure 6-5: Modal cluster location rank by herding and agglomeration (N = 166). Points jittered vertically for visibility. Alternative model 2.



6.2.4. Differences between cells

This model is created the same parameters as the main model, but with one adjustment. The line `ask sites [set qa 75 - 1.5 * cell]` in the code is changed to `ask sites [set qa 75 – 2.5 * cell].` This makes the differences in attractiveness between the cells larger.

Figure 6-6: Sectoral clustering as a product of agglomerative externalities with herding (N = 792, includes a generalized additive model, outliers removed). Differences of attractiveness between cells scale by 2.5.

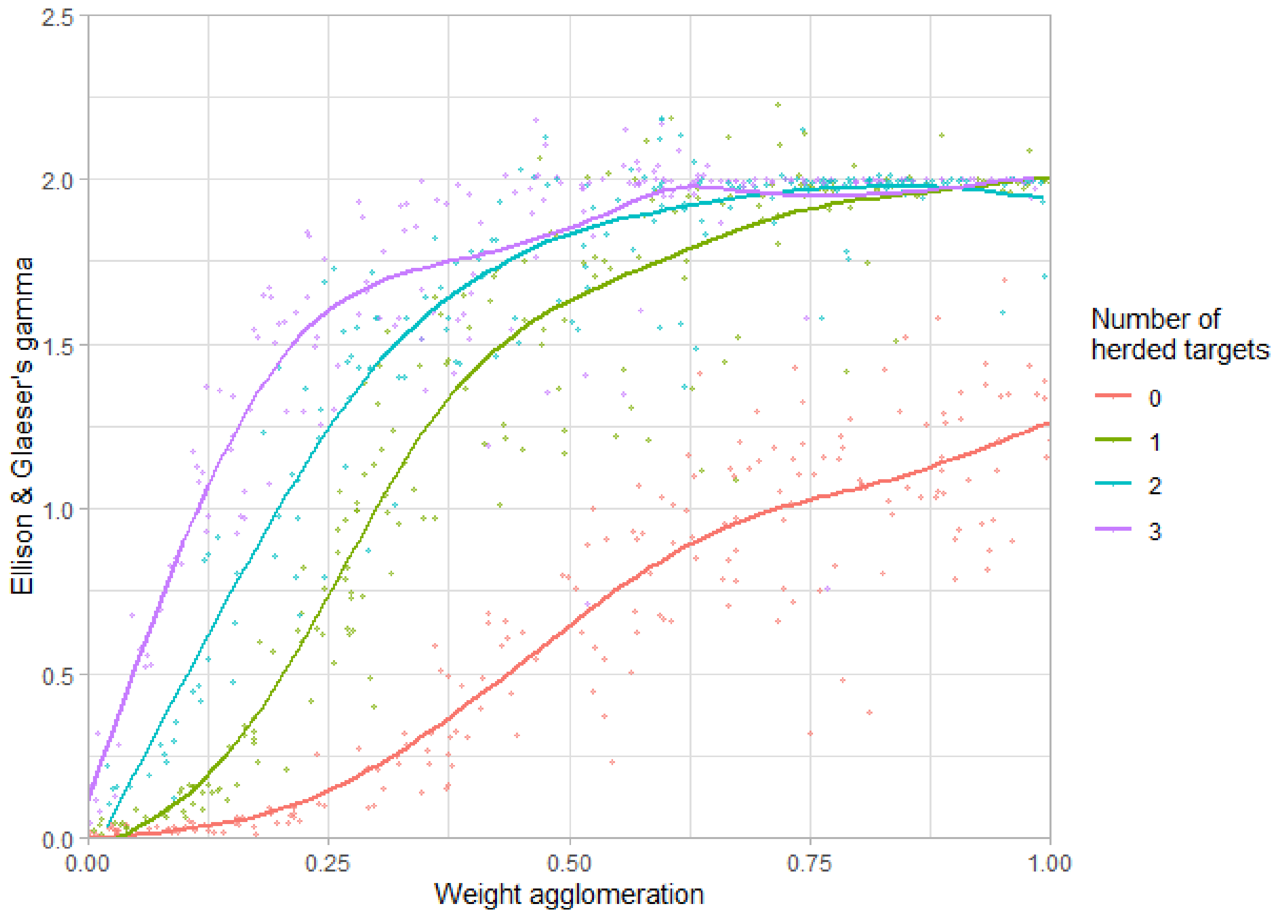
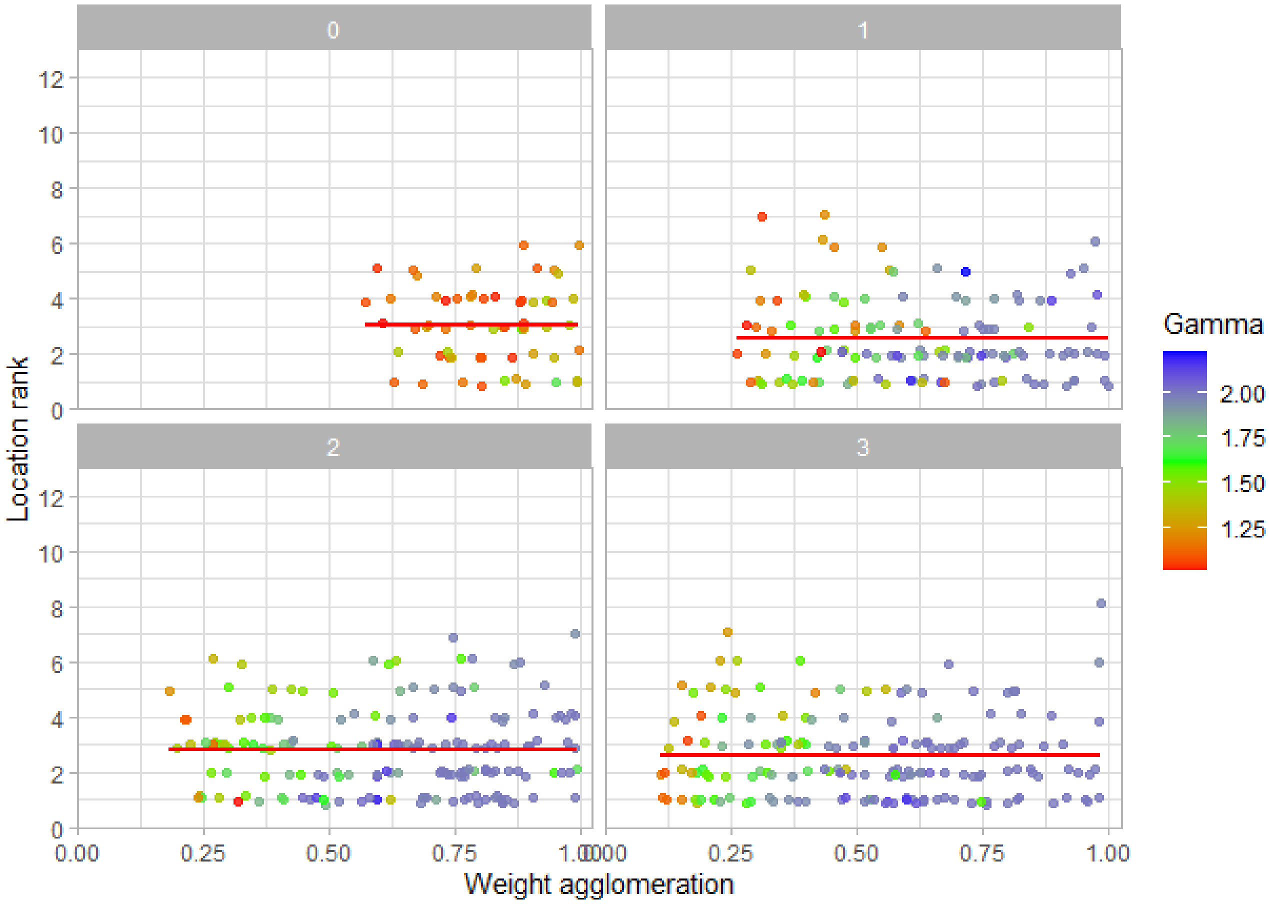


Figure 6-7: Modal cluster location rank by herding and agglomeration (N = 523). Points jittered vertically for visibility. Differences of attractiveness between cells scale by 2.5.



6.3. Topic list

[Omitted]

6.4. Interview Transcripts

[Omitted]