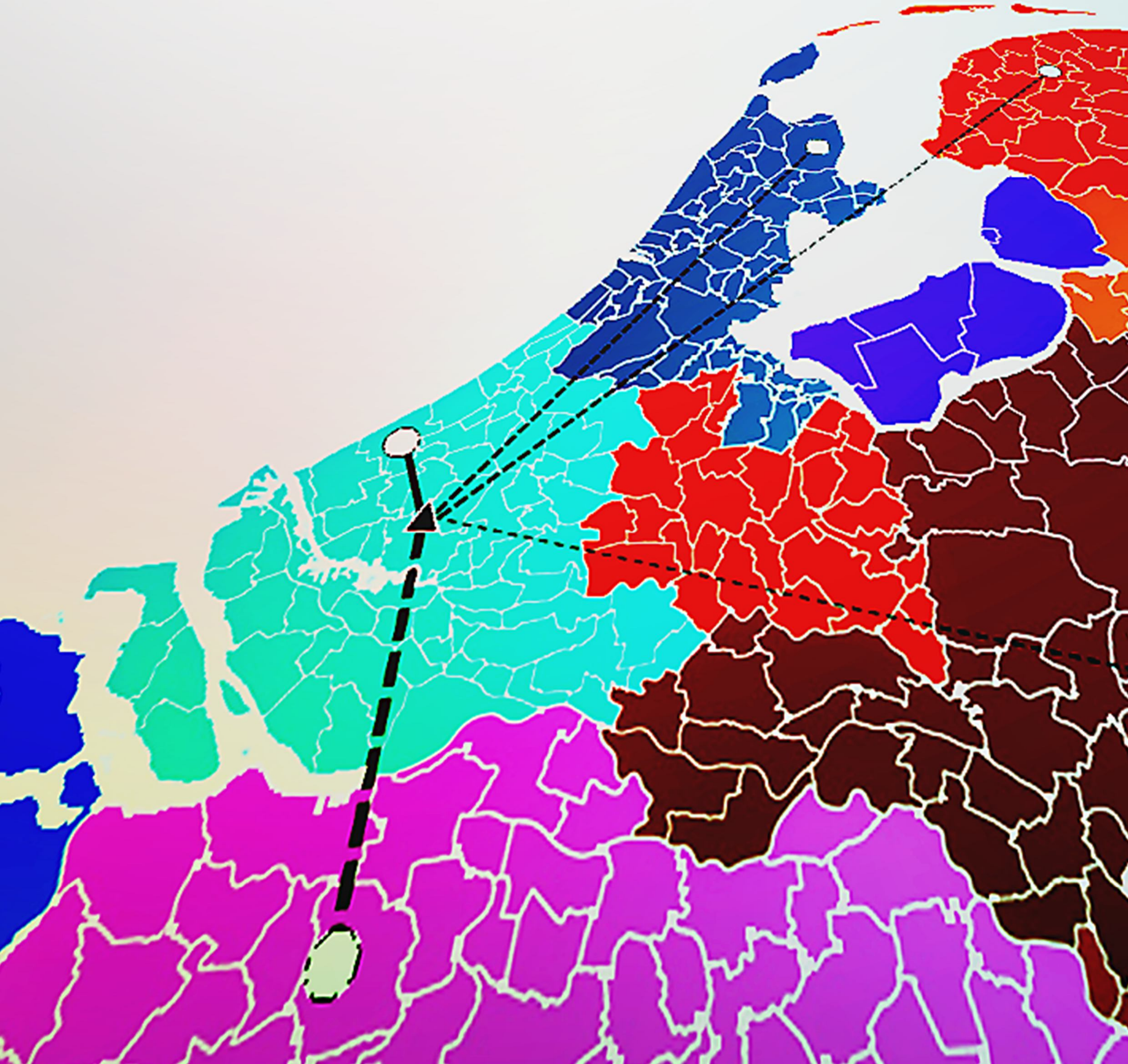
The Role of Geographical Proximity and Industrial Relatedness in Dutch Domestic M&As



Utrecht University
Faculty of Geoscienes
Heidelberglaan 2
3508 TC Utrecht
The Netherlands

Universiteit Utrecht

Master thesis

Research Master in Human Geography and Planning

The Role of Geographical Proximity and Industrial Relatedness in Dutch Domestic M&As

Nils Ellwanger
21.04.1985, Filderstadt
Student number: 344 48 72
nilsellwanger@gmail.com

Primary supervisor

Prof. Ron Boschma
Professor in Regional Economics
Faculty of Geosciences
Utrecht University

Secondary supervisor

Dr. Tom Broekel
Post-Doc
Faculty of Geosciences
Utrecht University

Utrecht, 25.07.2011

Preface

With completion of the research project presented in this master thesis the research master programme "Human Geography and Planning" at Utrecht University is coming to an end. It was an intensive and instructive period of two years in which I have learned a lot about different fields of geography and about how to conduct research in social sciences. Ron triggered me to delve into economic geography, what finally lead to the work in hands. My research project is based on a scientifically little exploited commercial database on M&As. The available data have allowed to construct measures of geographical proximity and industrial relatedness and induced me to research their impact on the formation of M&As. I have also investigated the post-acquisition performance of M&As by using accounting data as well as the role of social proximity by board member data of another database. However, the data coverage was not sufficient to complete these ideas and therefore I focused on the ideas that are presented in this thesis.

The research was conducted between August 2010 and June 2011. About five months of this period were spent on exploring, cleaning and preparing the data. I had never expected that this would take that much time. Nevertheless, there was enough time left to do the actual research and to write this report. The ones who do not want to read the whole thesis are referred to the article based on this thesis. I would like to thank my supervisors Ron and Tom for providing guidance and trust during the time of conducting the research and writing this report. I am also warmly thankful to all my fellow students. We had many fruitful intellectual discussions and a very good time together.

Nils

Summary

Partner selection by humans does not occur randomly. The same holds for the behaviour of companies when it comes to target selection in mergers and acquisitions (M&As). The master thesis in hand shows systematic biases in the behaviour of bidding companies. M&As refer to economic transactions in which two companies consolidate, either by acquiring the assets of the other or by creating a new entity. Many disciplines study this phenomenon and focus primarily on the deal itself and the situation after its completion. M&As are very complex events and every transaction is different and has specific consequences. As unique as every deal is as unique are the events that take place prior to the deal announcement. However, deal announcements are always associated with M&A formation, which is the selection of a suitable company to acquire or to merge with, by the bidding company. This notion of M&A formation has largely been neglected in M&A research. Whereas motives to acquire another company as well as the characteristics of bidders and targets are soundly identified not much is known about which two particular companies actually announce a deal and which do not, or in other words, which is the target selected by a particular bidder.

Economic transactions between two or more actors, such as trade, collaborations or investments, are often associated with geographical proximity. In fact, there are good reasons why geographical proximity influences M&A formation: information advantages, familiarity, localization effects and strategic reasons. Another important predictor of M&A formation is industrial relatedness, which is mainly due to the potential realization of synergy effects and due to other strategic decisions, such as the acquisition of competitors. Considering domestic M&A deal announcements I argue that the more two companies are geographically proximate and the more they are industrially related the more likely a deal is announced. Speaking from the perspective of the bidding companies, it is shown that a home bias and industrial relatedness bias inheres in their target selection behaviour.

Evidence on these biases is sparse and this research mainly builds on a very few studies that show the existence of home bias in domestic U.S. deals. Although these studies provide strong evidences for the existence of a domestic home bias, they suffer from some drawbacks. First, the minimum and maximum distance to potential targets as well as the number of potential targets was not considered. This neglect of the individual choice set of every bidder lead to biased results. Second, the home bias was only estimated on average, leaving the individual home bias for each single company being unobserved. Third, instead of deal announcements solely actual deals were included. Fourth, control variables were sparsely used and also the interaction with other predictors was only studied to a limited extent.

This research aims at resolving these shortcomings and to extend the research on home bias, industrial relatedness bias and M&A formation. Resolving the shortcomings of previous studies helps to gain more valid results. A mayor advance of this research is the development of individual bias measures that take into account the number of potential targets as well as their characteristics. As the measures are standardized the extent of the biases can be compared between different countries. New is also the research site of the Netherlands, in which, due to its size, geographical proximity might be suspected of having little significance. By controlling for localization effects it is shown to what extent the location of certain industries affects the home bias. Striking is the evidence on regional home biases. It is shown that a home bias even exists in deals in which the bidder selects a target from its own municipality, COROP region or province.

The empirical tests are in principal based on a comparison of company dyads that actually announced an M&A transaction and company dyads for which no deal was announced, but for which potentially a deal could have been announced. The data depict domestic mergers and acquisitions in the Netherlands announced between 2002 and 2008. Geographical proximity is on the one hand measured by the co-location of both company headquarters on the same regional level, on the other hand by pure geographical distance, which is an extremely precise measure as the location of the headquarters is exactly known. Industrial relatedness is measured by the positioning of both primary business activities

in an industrial classification system as well as by means of a newly developed co-occurrence measure, which denotes how often two activities co-occur in the same company.

Compelling evidence shows that about 71% of the bidders are home biased and about 76% are industrial relatedness biased. On average, the selected target is 28 km closer than the average target on the national M&A market. If the potential target is strongly industrially related instead of only weakly related or unrelated, M&A formation is 89 times more likely. These results are seen as strong evidence for the existence of a home bias and industrial relatedness bias in domestic M&As. One question remaining is to what extent different underlying causes are responsible for this home bias. Although the effects of information advantages, familiarity and strategic reasons cannot be disentangled it is possible to control for localization economies. Doing that, the estimates are lower, but in general remain significant, which means that localization economies are not the only cause of the observed home bias.

An exploration of home biases on different spatial scales shows interesting results. The home bias for deals within the same province is about 10 km, for deals within the same COROP region about 5 km and for deals within the same municipality about 1 km. When controlling for localization effects the home bias on a municipality level turns insignificant. That means that for deals within the same municipality localization economies are the main contributor to the home bias. Comparing the relative home biases, home bias is largest for deals within the same province. Industrial relatedness bias is about the same on all spatial scales. A logistic regression model shows that geographical proximity and industrial relatedness reinforce each other. The total impact of the present variables on M&A formation is, however, rather small. If M&A formation occurs randomly the average likelihood of deal announcement between two particular companies is 0.12%, whereas the modeled average likelihood is 0.55%. This means that M&As are rare events which arise from history and coincidence and depend on many additional factors.

While it remains about impossible to actually predict deal announcements between two particular companies, this research shows systematic and robust evidences on the role of geographical proximity and industrial relatedness in domestic M&As. One stunning finding is that home bias even inheres in bidders that selected a company from the same municipality, COROP region or province. By providing these evidences this thesis makes an important contribution to the M&A literature as well as to the proximity literature in economic geography. The thesis also illustrates that geographical proximity remains to be a significant in economic decision making processes.

Keywords: **M&As, Geographical proximity, Industrial relatedness, Dyadic formation, The Netherlands**

Contents

Tables and Figures

Appendix Tables and Figures

Abbreviations

Divers	
M&A	Mergers & Acquisitions
NACE	Industry standard classification (Nomenclature statistique des activités économiques dans la Communauté européenne)
Methodology	
ALL	Assumption that bidder can select between any potential target
SG	Assumption that bidder can only select between potential targets of its actual target's industry
D	Deals
ND	Non-deals
b	Bidder
rt	Real target
pt	Potential target
HB	Home Bias
IRB	Industrial Relatedness Bias
Variables	
SPCM	Categorial variable indicating geographical proximity
P	Dummy variable indicating co-location in the same province (but not COROP region) of the bidder's and target's headquarter
C	Dummy variable indicating co-location in the same COROP region (but not municipality) of the bidder's and target's headquarter
M	Dummy variable indicating co-location in the same municipality of the bidder's and target's headquarter
USDG	Categorial variable indicating industrial relatedness
S	Dummy variable indicating same NACE section (but not NACE division) of the bidder's and target's activity
D	Dummy variable indicating same NACE division (but not NACE group) of the bidder's and target's activity
G	Dummy variable indicating same NACE group of the bidder's and target's activity
UU	Dummy variable indicating that bidder and target are unlisted companies
SiSi	Dummy variable indicating that bidder and target are single-locational companies
MuMu	Dummy variable indicating that bidder and target are muli-locational companies
UnUn	Dummy variable indicating that bidder and target are undiversified companies
DiDi	Dummy variable indicating that bidder and target are diversified companies
Statistics	
df	Degrees of freedom
B	Regression coefficient
CI	Confidence interval
SE	Standard error
SD	Standard deviation
OR	Odds ratio
RR	Relative risk
N	Number of observations
***, **, *	The asterisks denote statistical significance at the 1%, 5% and 10% level, respectively. Although many results are significant at the 0.1% level, this is not separately indicated.

1 Introduction

The introduction illustrates the aim of this work and the ideas behind. It is shown how this research builds on different strands of literature and what the contributions of this research are. Finally, an outline is provided that guides through the paper.

1

Geographical proximity is an important facilitator of economic processes between two or more actors. Economic geography and other disciplines have shown that economic actors are often affected by a so-called home bias. Home bias describes the propensity of individuals or companies to act within the own local environment. In other words: geographical proximity is preferred over geographical distance. This is not only documented for international activities, for which primarily the institutional setting might be responsible, but also for domestic events. The existence of a domestic home bias has been identified for various dyadic and *k*adic activities, such as investments (e.g. Coval and Moskowitz, 1999; Chen et al., 2010; Cumming and Dai, 2010), trade (e.g. Wolf, 2000; Hillberry, 2003), co-inventorship (e.g. Ter Wal, 2010), or other forms of collaborations (e.g. Katz, 2005; Broekel and Boschma, 2009). A few attempts have been undertaken to show that a domestic home bias also exists in M&A formation. Eun and Mukherjee (2006), Grote and Umber (2006) and Chakrabarti and Mitchell (2008) found evidence for large U.S. deals.

M&A formation is the selection of a target company by a bidding company, or vice versa, which consequently leads to the announcement of a deal. An M&A deal describes an economic transaction in which one company is bought by another or in which two companies set up a new company jointly. In the formation phase domestic home bias can have four mayor causes: strategic decisions, familiarity, localization economies and information advantages. It is not difficult to imagine that companies might tend to acquire local competitors; to select a target from the region in which the manager has used to live; to select a target from the same region because the bidder is simply located within the local cluster; or because a bidder has more and better information on local targets, what facilitates the decision making process through target identification and valuation advantages.

Although there is evidence on the existence of home bias we lack sophisticated, systematic research. The four known studies have basically four shortcomings. First, the studies neglected the choice set of the individual bidder. Grote and Umber (2006), for example, solely estimated an average and absolute home bias, which was 555 km in his case of U.S. deals, not taking into account the distance to all potential targets. A standardization of the absolute estimates would provide a more meaningful estimation of the home bias. Second, the studies investigated an average home bias. They left us in the dark about the individual home bias of every single and the proportion of bidders that were home biased. Both shortcomings are addressed by developing an individual and relative home bias measure that indicates a bidder's home bias on a scale from -100 to 100. Third, all four studies investigated actually implemented deals. Home bias, however, exists in the selection of targets and therefore it would be more meaningful to investigate deal announcements. In this thesis deals do not only include completed deals but also announced deals that were not completed.

Fourth, the studies did not control for other factors that ease M&A formation. Boschma (2005) argues that geographical proximity can be substituted or reinforced by other forms of proximity, such as cognitive, institutional, organizational or social proximity. Therefore, it is important to control for other relational factors. As M&As are often seen as an instrument to realize synergies industrial relatedness seems to be another crucial factor next to geographical proximity. Consequently, this thesis explicitly addresses the role of industrial relatedness. Similarly as home bias refers to company locations, industrial relatedness bias refers to the activities or industries bidder and target are active in. Industrial relatedness has been studied within the voluminous corporate diversification literature (e.g. Montgomery, 1994; Servaes, 1996), which usually distinguishes between related and unrelated diversification strategies. These strategies can be embarked internally or externally, by M&As. Among scientists and practitioners opinions differ whether related or unrelated diversification increases firm value (Martin and Sayrak, 2003). Also M&A researchers have addressed this question (e.g. Limmack and Mcgregor, 1995; Flanagan, 1996; Homberg et al., 2009) and have often associated related deals with the realization of synergies. Although it is understandable that managers, shareholders and researchers are interested in the effect of industrial relatedness on value creation it is still surprising that there is hardly any research that explicitly addresses the propensity to acquire industrially related targets. To me only the study of (Hussinger, 2010) is known that explicitly addresses the role of industrial relatedness on target selection in M&As. In order to investigate the industrial relatedness bias of bidders a new industrial relatedness measure is developed, based on the Tanimoto coefficient. This is a similarity measure often used in ecological research.

So far, the role of geographical proximity and industrial relatedness has hardly been addressed in literature on M&A formation although there is an abundant body of empirical M&A literature. This literature has been created, starting about 35 years ago, by many different fields, such as management and finance, strategy, sociology, economics, accounting and others (Cartwright, 2006). Despite a formidable amount of articles the mayor research foci of empirical works are focused on the motives to acquire (e.g. Seth et al., 2000; Berkovitch and Narayanan, 1993; Luypaert and Huyghebaert, 2007), on post-acquisition performance (e.g. Lubatkin, 1983; Barney, 1988; Morck et al., 1990; Chatterjee et al., 1992; Matsusaka, 1993; Ahuja and Katila, 2001; King et al., 2004; Capron and Shen, 2007) and on target characteristics (e.g. Hasbrouck, 1985; Powell, 1997, 2004; Tsagkanos et al., 2008; Brar et al., 2009; Cai and Tian, 2009). Research on M&A formation from a dyadic perspective, i.e. by addressing the relation between the two involved companies, is hardly developed. One strand of research shows which companies are more likely to become acquisition targets and a very few papers show which companies are likely to acquire. Both views, however, neglect the characteristics of the other company, i.e. the target likelihood studies do not consider the characteristics of the acquiring company and the acquirer likelihood studies disregard the targets.

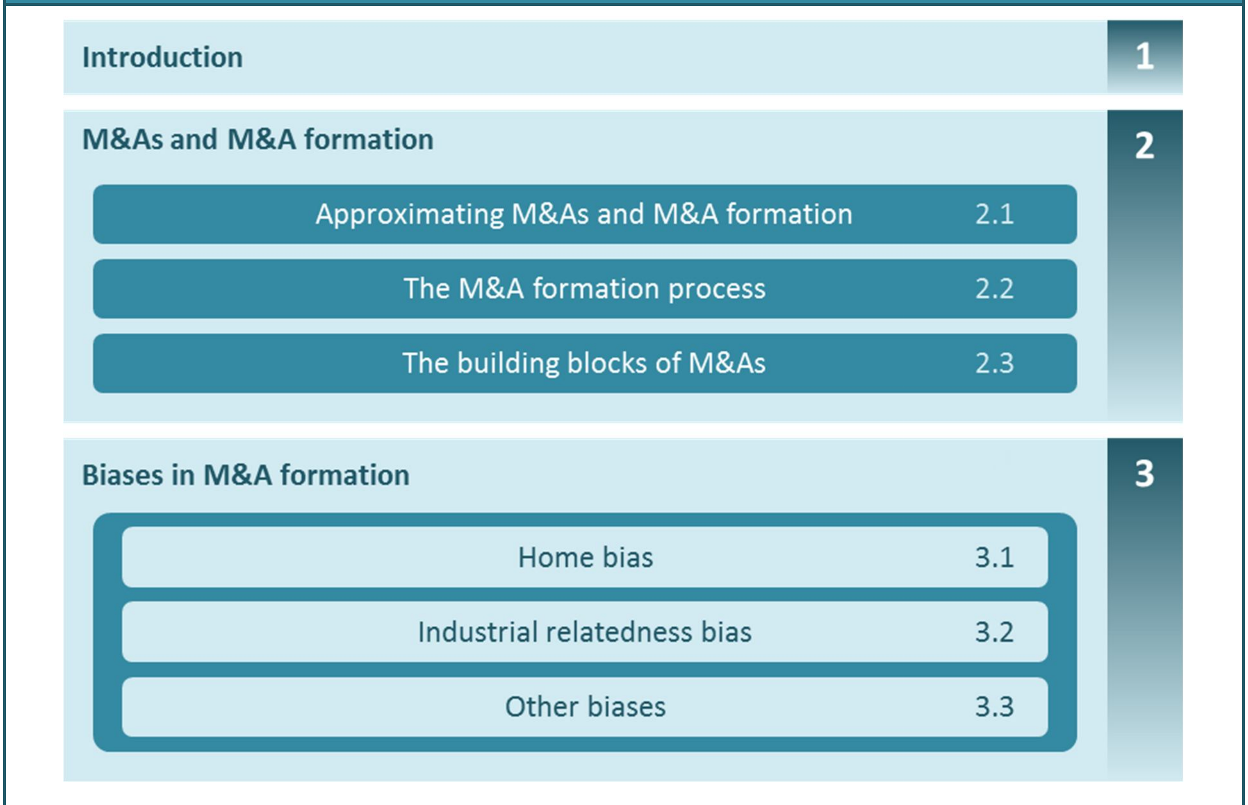
While the formation of M&As has hardly been researched from a dyadic perspective other phenomena are subject of dyadic research. An everyday example is partnering between humans. Sociologists have kept themselves busy by explaining the selection of dating or marital partners (e.g. McFarland, 1975). Also the formation of wars and alliances between states is subject of dyadic research within international relations (King and Zeng, 2001a). In this paper M&As are seen as purely dyadic phenomena because usually an M&A occurs between solely two companies. Only in very exceptional cases more than two companies form a merger. It must be clear, however, that this thesis only reveals biases when it comes to M&A formation. Dyadic research cannot determine which two particular entities will match. Given about $36*10^{18}$ possible dyads between humans it is unthinkable that any model would accurately predict the relatively small amount of friendships. The chance of friendship between two particular persons is extremely small. The same is true for M&A formation. Whether two particular companies engage in an M&A or not depends on history and coincidence and on a complex interaction of much more factors than only geographical proximity and industrial relatedness. Thus, M&A formation implies that M&As are extremely rare events.

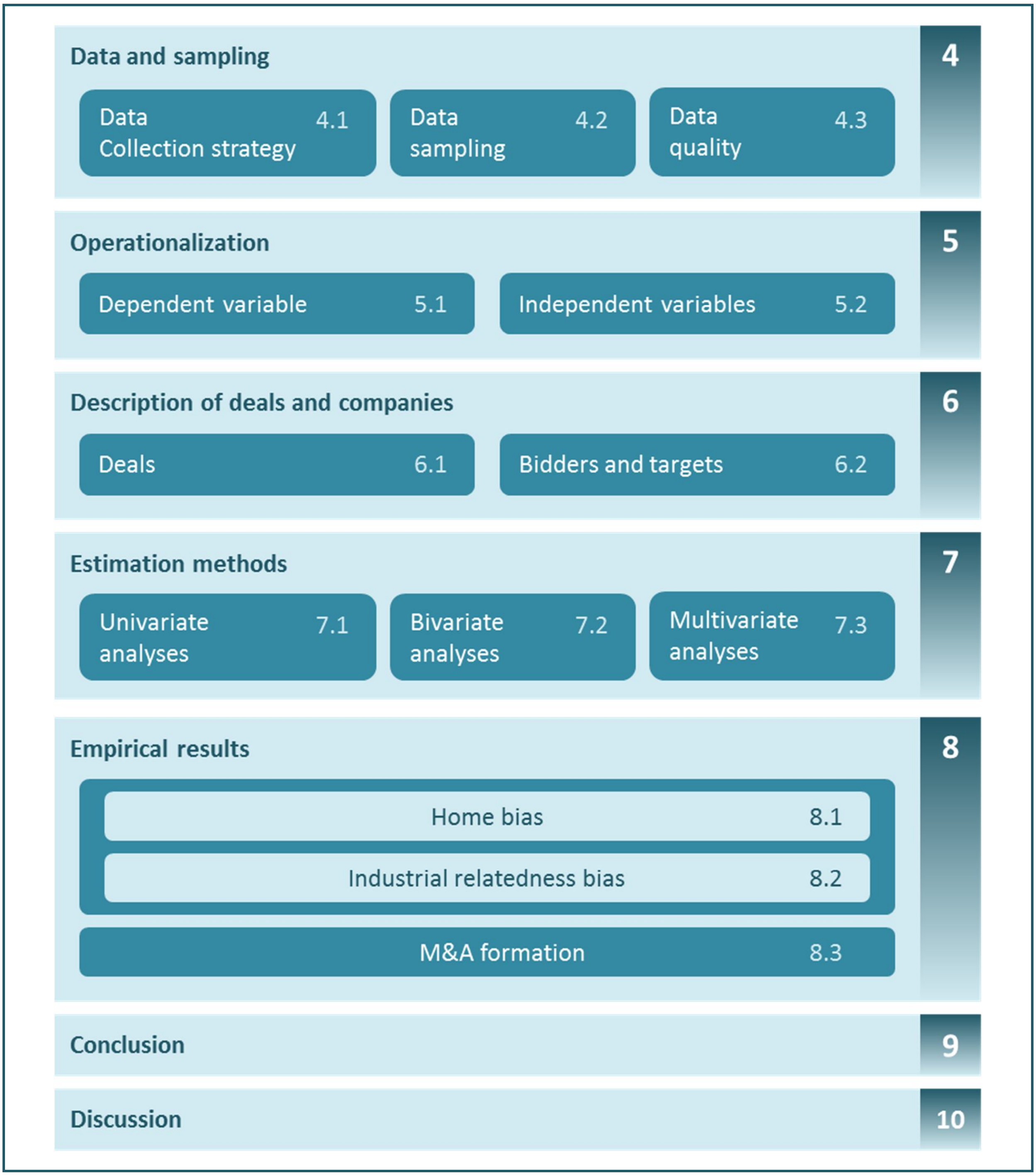
Given the theoretical background of the literature on the role of geographical proximity, industrial relatedness and the literature on M&As, the aim of this thesis is twofold. One the one hand, I aim to systematically measure the extent of home bias in domestic M&As addressing the mentioned shortcomings of the previous studies and to systematically measure the extent of industrial relatedness bias in domestic M&As, which is so far about completely unknown. On the other hand this thesis explores two further issues. First, I want to test how geographical proximity and industrial relatedness interact and in how far these dimensions contribute to the explanation of M&A formation. Second, I want to test whether home bias also exists on different sub-national scales.

The second part of the aim is exploratory in character. It is not known to what degree geographical proximity and industrial relatedness actually contribute to M&A formation. It is neither known whether both dimensions serve as substitutes or whether they reinforce each other. This is explored by means of a logistic regression model, which is adapted to the specificities of rare events. This model also includes some further control variables. Besides those explorations the thesis presents completely new findings: The existence of regional home biases. So far, home bias was investigated on an international and domestic scale, but not for M&A deals that take place within small regions in the Netherlands. The detection of home biases for bidders that selected a target from the same province, COROP region and even municipality is enabled by a very detailed locational data of the company headquarter locations. This data is based on the postcode of the companies and on average one postcode contains only 17 addresses.

After this introduction the remainder of this work is organized as following. The first theoretical part, chapter 2, introduces the phenomenon of M&As and M&A formation. It is shown how M&A literature explains M&As from a focal, i.e. unrelational, perspective, which motives of the bidder to acquire can be identified and which the typical characteristics of bidders and targets are. Based on that, chapter 3 presents a synthesis and elucidates which dyadic factors might explain M&A formation. The role of geographical proximity and industrial relatedness is extensively discussed, while some other factors that facilitate M&A formation are briefly addressed thereafter. Chapter 4 documents how the data is collected and sampled. It also contains a discussion of the data quality. The operationalization of the dependent variable (M&A formation) and the independent variables (geographical proximity, industrial relatedness and control variables) is specified in chapter 5. Chapter 6 provides some descriptives of the deals in the sample and of the involved firms. This provides a first impression of the role of geographical proximity and industrial relatedness. Chapter 7 explains in detail the methods used for the statistical tests. These are bivariate methods, the construction of individual bias measures, which are analyzed by univariate techniques, and logistic regression models. Thereafter chapter 8 provides the results, evidence on domestic and regional home bias and industrial relatedness is presented as well as the M&A formation model and its assessment. Chapter 9 provides the discussion of the results. These are discussed in a wider context in the concluding chapter 10. The appendix contains some additional research and provides all empirical results that are placed in the report.

Figure 1 Outline of the thesis.




Data and sampling
4
Data Collection strategy 4.1
Data sampling 4.2
Data quality 4.3
Operationalization
5
Dependent variable 5.1
Independent variables 5.2
Description of deals and companies
6
Deals 6.1
Bidders and targets 6.2
Estimation methods
7
Univariate analyses 7.1
Bivariate analyses 7.2
Multivariate analyses 7.3
Empirical results
8
Home bias 8.1
Industrial relatedness bias 8.2
M&A formation 8.3
Conclusion
9
Discussion
10

2 Focal views on M&A formation

This chapter starts by defining the business transactions this work is about (2.1) and the processes that lead to deal announcements and M&A formation from the perspective of the bidder (2.2). Section 2.3 illustrates the characteristics of a typical target and a typical bidder. This enables the relational synthesis made in chapter 3.

2

2.1 Approximating M&As and deal announcements

2.1.1 Definition of mergers and acquisitions

M&A is usually a buzzword for three things. First, it describes the sector dealing with mergers and acquisitions, consisting of advisory firms or investment banks. Second, it is the heading for different forms of business transactions, which is the relevant use of the term in this paper. Third, it terms the legal and economic discipline dealing with M&A transactions. M&As as business transactions are largely about integrating two independent companies into one organization. This form of corporate restructuring can be achieved by two different means: by means of an acquisition or by means of a merger. An acquisition is the purchase of one company by another; a merger refers to setting up a new establishment by two previously separate companies. Whether one company takes over another or whether two companies merge is above all mostly a decision depending on legal issues (Maynard, 2009). In economics, finance and law both kinds of transactions are commonly investigated under the label M&As.

In an acquisition the acquirer buys at least 50% of the assets and liabilities or stock of the target (Maynard, 2009). If a bidder acquirers less than 50%, but more than 5%, we would talk about a "partial acquisition" (Kang and Kim, 2008). The target's shares can be bought by the acquirer's stock, by cash or by both. Thus, an acquisition is basically an economic transaction in which ownership is changing. Literature can be confusing because for this kind of transaction there are several different names in different strands of literature or in different contexts. In fact, many articles talk about "takeovers" (e.g. Palepu, 1986; Berkovitch and Narayanan, 1993; Cai and Tian, 2009). In the UK, however, takeovers only refer to deals in which a public company is being acquired. In the U.S. acquisitions are often called "statutory mergers" (Maynard, 2009), although a merger is something different, and in the Netherlands we talk about overnames.

In a merger acquirer and target are integrated by establishing a new legal and economic entity while the two old companies legally disappear. Therefore a merger is sometimes also called "merger by establishment". Another term is "statutory consolidation" and in the Netherlands mergers are called fusies. Despite the fact that mergers occur very rarely in reality, we often encounter the term "merger" in media and scientific literature. This is because the terms "acquisition", "purchase" or "takeover" carry a negative connotation and the term "merger" offers a euphemistical replacement. Newspapers told us, for example, that the deal between the Chrysler Corporation and the Daimler-Benz AG was a merger. Little effort is required to find out that it was actually a takeover of the Chrysler Corporation by the Daimler-Benz AG (History, 1998). In a great deal of scientific M&A literature the authors even use the term of mergers although both acquisitions and mergers are meant (e.g. in Blackburn and Lang, 1990; Matsushima, 2001; Colcera, 2007; Schenk, 2008b). In this work no distinction is made between acquisitions and mergers, as it is common practice in economic M&A research. This means that also in mergers there is an acquirer (after the deal), bidder (before the deal) and target role, which makes conceptualizations and textual explanations much easier than constantly referring to both kinds of transactions.

There are two forms of M&As that are not considered in this work. These are deals in which the bidder was a non-company and cross-border deals. In general, every legal person may buy a company or a part of it. In fact, companies are often acquired by their own managers. These deals are called management buyouts (MBOs). If the bidder consists of private persons which have not been involved in the management of the target the purchase of the company is called management buy-in (MBI). If a company is bought by private persons at the stock market, thus if a company offers its stock for the first time on the stock market, we talk about an initial public offering (IPO). In all those cases where private persons are involved the bidder does not have an industrial profile (i.e. product portfolios, technological capabilities, etc.) and in the case of MBOs not even a different location. Therefore these cases are not useful objects for the purpose of this research. Only legally and economic independent companies are considered. Being registered at the Chamber of Commerce and therefore being a legally independent entity is, however, not sufficient because pure divisions or branches depend on the economic decisions made in their headquarters. Subsidiaries, however, have some economic autonomy and can therefore be involved in an M&A.

Cross-border deals are not considered either in this work for two reasons. First, the effect of geographical proximity would be biased by different institutions in different countries. Cross-border deals often follow a different logic and are based on other motives than domestic deals. In fact, cross-border M&As are often seen as the main form of foreign direct investment and as an instrument to get access to new markets while domestic M&As could be seen as an event of restructuring. Also political, legal or cultural factors might play a role in target selection. Second, the empirical investigation requires a closed system. Considering all potential dyads of companies that could have merged all over the world is simply not possible.

2.1.2 M&A formation

The phenomenon explained in this work is M&A formation. M&A formation refers to announced deals taking into account the relation between the two involved companies. The relation stems from the unique characteristics of the bidder and target. The interrogative form of M&A formation would be "Which two particular companies announced to engage in an M&A?". In other words, M&A formation is the deal announcement in consideration of the two involved companies and their relations to each other (**Figure 2**). Deal announcement refers to the pure fact that an M&A deal is announced, ignoring whether the deal was finally completed or not. This wide definition is made because I am interested in the behavioural bias of companies, i.e. which M&A partner is preferred, and not in whether a particular company dyad de facto completed an M&A. Reasons for non-completion might be that the deal becomes forbidden by an antitrust agency, because financing partners might bail out, due to a competing offer, because one of the firms was misevaluated, etc. (Weitzel and McCarthy, 2009). In the ten largest M&A markets the number of non-completion is about 14%, and about 20% in the Netherlands (Zephyr).

Essential is that M&As can be seen from different perspectives. These are the views from the bidder, the target or from external entities, such as advisory companies or the government. About the complete M&A literature implies that bidders are the dominant and deciding entities and therefore works build up on their perspective; the proactive role of companies that want to sell their business is widely neglected. While this view is missing in M&A literature the seller's side is addressed in the so-called divestiture research. This is a very small stream within corporate restructuring research. However, the divestiture of one firm does not necessarily become the acquisition of another one, because also spin-offs, carve-outs (sale of a minority share of the child company) and split-ups (split of a company in two or more new companies) are included. These processes do not involve an external buyer. If at all, only sell-offs can be regarded as mirror images of an M&A. However, whereas M&A literature inheres a bidder dominated view the divesting company selling off is regarded as having full discretion over the transaction process (Brauer, 2006). Thus, M&A literature and divestiture research on sell-offs have different views on the company that finally decides about a deal announcement.

Although M&As can be described from different perspectives choosing one viewpoint facilitates conceptualization and a better understanding. Therefore I take, as most M&A literature, a focal bidder perspective and define M&A formation as deal announcement from the perspective of the bidder considering the relation to the target. In other words: M&A formation is conceptually equal to target selection. Whereas deal announcement solely refers to the announcement of a deal between two companies, M&A formation refers to the "partnering" between two particular companies (see **Figure 2**).

Figure 2 The concept of M&A formation.

2.2 The M&A formation process

The M&A formation process can be initiated by both potential bidder and potential target. That means, M&A formation depends on the intention of the potential bidder to acquire and the willingness or efforts of the potential target to sell. While the role of the target is hardly understood (Graeber and Eisenhardt, 2004) there is a bidder dominated view in M&A research. This implies that the bidder is seen as the decisive entity. In this thesis this view is followed because the theory can much more easily be communicated and conceptualization becomes simpler as well.

2.2.1 From the search process to M&A formation

The nature of the bidding process, i.e. the action that is closely associated with deal announcement and M&A formation, largely depends on the ownership structure of the target. If a company is listed at the stock market an acquisition can be made by simply purchasing the target's stock on the open market. Often, bidders gradually and quietly buy stocks of a defenseless victim without asking the target's management for approval (creeping tender offer). Such hostile takeover practices were especially popular in the third merger wave from 1965-1975 (Luypaert and Huyghebaert, 2007). As these kinds of hostile takeover strategies are only possible if the target is a listed firm it becomes clear that the number of such deals is relatively small. Normally, deals are friendly, i.e. the bidder informs and negotiates with the target's board of directors, which finally does or does not agree on the takeover. If no agreement is reached the prospective bidder can try to persuade a simple majority of the shareholders to replace the management which will approve the takeover. In other words, the bidder tries to initiate a proxy fight.

M&A formation is always the result of preceding decision making processes. It is not necessary to understand these processes in detail, but a minimum insight is required in order to understand in how far geographical proximity and industrial relatedness can affect M&A formation. For this purpose I outline, from the perspective of the bidder, two possibilities that lead to the M&A formation between two particular companies: a more or less thorough and conscious search process of either the bidder or the target and pure opportunity (see **Figure 3**).

The search process model assumes that companies first make the decision to acquire another business and consequently search for an appropriate candidate. That means that the intent to acquire is internal to the company. The search process can be done proactively by the managers of a company or by external parties, such as M&A advisory companies, which search and propose potential targets. Often, the searching company has certain criteria potential targets have to fulfill in order to be identified as promising targets. These can relate to the geographic location, the industry or technology provided, the ownership structure, certain financial characteristics, the company's age or experience, its corporate culture, its reputation and much more.

However, not all companies exert such a strategic search. The other model describes cases of opportunities, which implies that one of two companies that are already in a kind of relationship takes the opportunity to initiate a deal. This might be applicable if two firms have already been in a buyer-supplier or a trading relationship before the deal announcement. This can be compared to marriages, for which it would be a good guess that the bride and groom have been in a liaison before, at least in Western cultures. Dependent on certain target characteristics companies might seize the chance to bid on the target, analogically to the guess that it would be a better opportunity to marry one's partner than somebody unknown. If two firms do businesses together or if individual persons of two companies know each other it might be suspected that an M&A could be beneficial for one or both of them. This can lead to M&A formation, but it is not known how many deals are outcomes of search processes and how many deals took place due to opportunity. The notion that opportunities trigger M&As can be summarized by Say's law "supply creates its own demand".

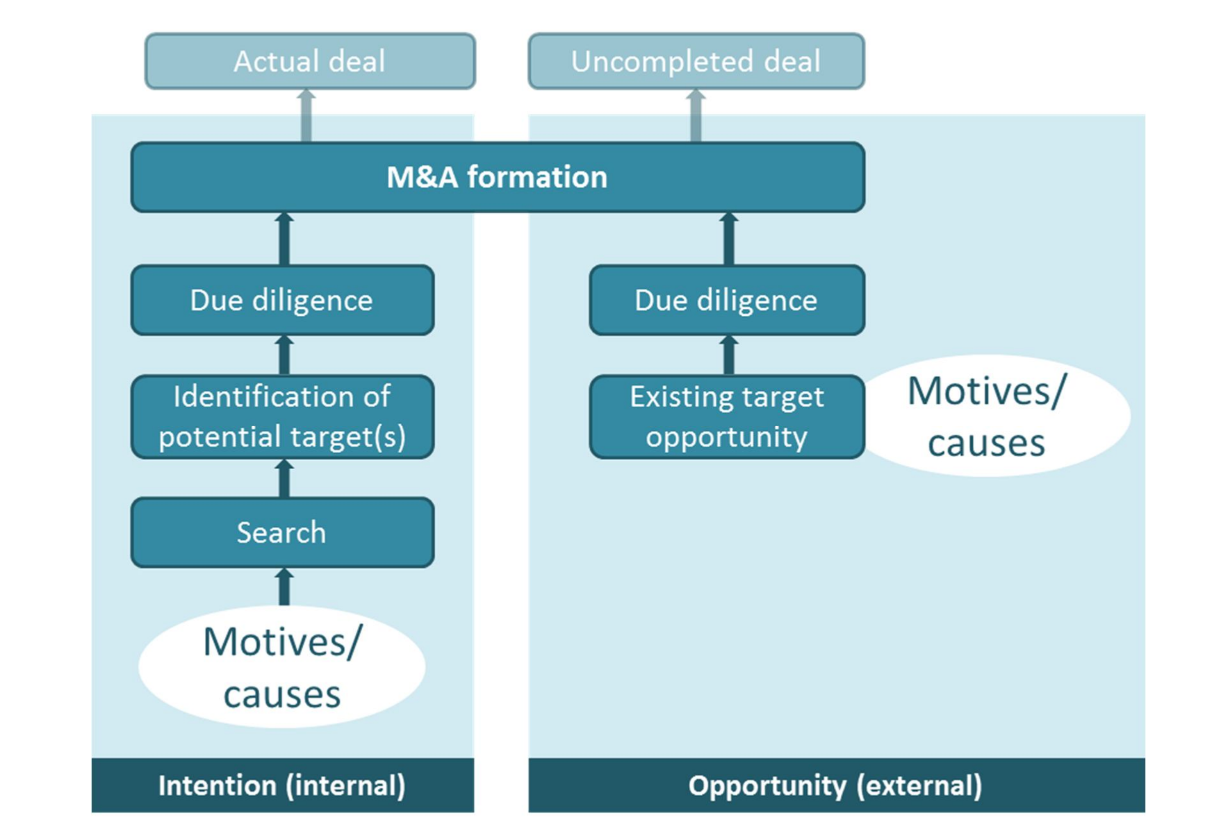
Once one or more promising potential targets have been identified, either resulting from a search process or by taking an opportunity, a letter of intent is prepared for the potential target, which states the intent to proceed to the next step, which is the due diligence process. Due diligence is a thorough assessment of the target's business in order to either assign the potential target to the bin or the leave it on the desk. It is possible that this phase is skipped in cases where the prospective bidder already knows the prospective target very well or where risk and deal value is very low. However, the larger the deal and the less known about the prospective target the more important is this phase for the success of the deal. Due diligence is almost always done in collaboration with external advisors and the target itself. It is often protected by a non-disclosure agreement in order to assure confidentiality. Usually it is the most time-consuming and expensive stage in the whole process (Paulson, 2001). If the due diligence process has revealed clarity about whether to go on with a particular company, price and method of deal financing the deal structure is formulated and negotiations follow. The target board normally gets a deadline for rejecting or accepting the premium bid. Sometimes targets just do not react (Chapple and Treepongkaruna, 2006), but if both firms confirm that the deal is going ahead the deal is getting announced. It must be clear that deal announcement is not equal to bids. There can be many bids that do not lead to deal announcements and, as already said, there can be deal announcements that do not lead to an actual deal.

The processes preceding a deal announcement, or an actual deal, can be very complex: Many entities can be involved and the processes can take very long. Preparing a bid often takes months (Paulson, 2001). Search process, negotiations, due diligence and all other facets that lead to a deal are unique for every deal. These facets rarely follow a linear sequence but resemble an odyssey or at least an adventure. The sober quantitative researcher, however, puts all those happenings into a black box. It normally remains obfuscated who has had the largest impact on the decision to bid on a particular target and why a particular target was chosen. For this research it is sufficient to know which final decision has been made: announcing an M&A with a particular company or not.

2.2.2 Motives and causes

If the bidder is a financial investor, mostly a private equity or venture capital firm, deals can be seen as pure investments. In those so-called institutional buy-outs (IBO) the bidders are solely financial sponsors, not integrating their businesses and not even being involved in their target's management. Given that M&As are considered as strategic and not as financial deals, and as domestic and not as cross-border deals, scientific literature basically discusses two types of motivation to acquire another company: firm based strategic motives and agency based managerial motives. In fact, those motives imply that the company first makes a conscious decision to acquire and consequently starts a search process (internal intention). If M&As take place due to external triggers, namely by the existence of an opportunity, as argued in the previous section, motives are different and can be related to both, economic and personal factors.

Figure 3 The M&A formation process from the perspective of the bidder.



2.2.2.1 Strategic motives

There is a multitude of strategies that might result in an acquisition, or at least in the intention to acquire. Some essential motives are (a) growth, (b) access to strategic assets, (c) product differentiation and (d) synergy effects by business integration. In general, deals that are motivated by such strategic motives are associated with abnormal returns and wealth creation (Berkovitch and Narayanan, 1993) and the neoclassical thinking that M&As take place as a result of profit maximizing behaviour. Common to all four causes is the notion that M&As are an instrument of fast growth. Consequently, the causes of M&As can be seen as an underpart of firm growth theories.

M&As enable firms getting quick access to strategic assets, such as patents, skilled labour, licenses, brands or management skills, or valuable or new technology (Graebner, 2004). For example, TexIRBook publishers had started to acquire small software companies because they wanted to supplement their materials with multimedia software (Maynard, 2009). The theoretical foundation of this motive is the resource-based view of the firm. According to this view, companies possess asymmetrical abilities and M&As are a mean of gaining abilities and resources that the bidder does not possess. These abilities and resources include operational capabilities, processes, technology, management expertise and other strategic capabilities. If two firms have the absorptive capacity to understand and to know how to use each other's resources combining their baskets of resources would be a value-creating process.

M&As always imply that the two involved firms, although this can happen to very different degrees, integrate their businesses. This integration is seen as a source of synergy effects, often expressed as 1+1=3. The existence of these effects is based on traditional cost efficiency theory, which assumes that economies of scale exist. We can distinguish between operational, market, marketing, financial and managerial synergies. Sharing operations is most likely possible if the two firms are engaged in the same sector (Brakman et al., 2006). One resource of a company is its seller's market. By combining two corporations market power can be increased, which reduces competition and therefore assures a competitive advantage. Such price related (or collusive) synergies can be especially realized when the two firms operate in the same or in a similar market (Brakman et al., 2006). Financial synergies refer to the costs of capital. Increasing firm size through a merger or acquisition can lead to easier access to cheaper financing and/or create an internal market where capital can be allocated more efficiently (Brakman et al., 2006; Luypaert and Huyghebaert, 2007). If the bidder management believes that the management of a certain company has valuable managerial expertise that is complementary to its own expertise this could lead to a transaction as well (Matsusaka, 1993; Brakman et al., 2006). Also Galbraith and Kazanjian (1986) argue that similarities in organizational structure, management processes, beliefs and values do matter a lot (Pehrsson, 2006).

Another source of value-creation is suspected to exist in product differentiation. Some firms use M&As as a method to differentiate their product lines in the hope to spread risks. Industrial organization theory argues that product differentiation is a determinant of a firm's market position and source of competitive advantage (Nachum and Wymbs, 2005). Also technological diversification is related to superior performance (Miller, 2006). Other studies, however, show that specialized firms are doing better than diversified firms (Fan and Lang, 2000). Often, the question whether diversified or specialized companies are doing better is not only a theoretical, but also a conceptual and methodological problem.

2.2.2.2 Managerial causes

All those motives suggest that M&As facilitate the earnings of abnormal returns. However, a successful, value-creating deal is more the exception than the rule. This fact is paraphrased as the so-called merger paradox (Brouthers et al., 1998; Schenk, 2008c; Bogan and Just, 2009). There are two main reasons that are responsible for the failure of M&As: overpaying and bad integration. While the latter might be related to wrong target selection or failed post-acquisition actions the first is often related to the interest and behavior of managers. The reasons why managers want to acquire other companies, and overpay or fail to successfully integrate the businesses, are pure self-interest and incompetence. M&As motivated by managerial instead of strategic reasons are suspected to be value destructing and often undertaken by managers who do not own the company they manage.

The so-called theory of empire building suggests that managers strive to increase their power, prestige and remuneration. This is in line with the popular view that the behavior of managers is often associated with greed. It appears that acquisitions are seen as the coronation of a manager's career. The one who cannot show an M&A transaction within his career did not deliver his masterpiece. Especially for managers of large listed companies M&As can be a strategy of assuring their employment because running a larger firm and getting more complex management tasks lead to increasing salaries. Here, we have to deal with a typical principal agent problem. The managers (agents) of the bidding companies are primarily concerned with increasing their own welfare instead of increasing the value of the companies they manage (what would be in the interest of the shareholders, the principals). In consequence, such personal ambitions lead to systematic, and often knowing, overpaying for targets at the expense of the shareholders (Morck et al., 1990).

Assuming that managers behave in the interest of the company M&As can still fail if the managers are incompetent or infected by hubris (Berkovitch and Narayanan, 1993). This incompetence can be revealed in the target search (misselection), due diligence (misevaluation) or integration process (misintegration). In these cases managers overpay unintentionally. Morck et al. (1990) show that that the worst acquisitions are made by well-performing firms because their managers are most likely to be infected by hubris. Incompetence can also be related to the mimic and herding hypothesis, which says that companies imitate their peers to reduce the risk of falling behind (Schenk, 2008a). Acquisitions are often simply the "thing to do" and this also explains why they occur in waves (Brouthers et al., 1998; Matsushima, 2001; Luypaert and Huyghebaert, 2007). Both the hubris and herding hypotheses are widely acknowledged and can hold for listed as well as unlisted companies.

2.2.2.3 Opportunity

In line with the argument that M&As can take place just due to opportunity one can imagine that the management of a certain company is very bad in the opinion of the bidder, for example because they under-utilize their firm's assets resulting in low profitability. Also such cases can lead to a transaction, after which the target management is replaced. This phenomenon is known as the inefficient managers hypothesis (Matsusaka, 1993; Hasbrouck, 1985; Palepu, 1986). Weitzel and McCarthy (2009) show that poorly performing firms are often victims of hostile acquisitions. One can imagine many other factors that entice companies to acquire a certain business. Reasons for that are mainly the characteristics of bidders, such as the valuation and price of the company, its business portfolio, its possessed knowledge, etc. Such characteristics are discussed in the following section.

2.3 The building blocks of M&A deals

In the previous two sections causes and processes that lead to the intention to acquire, have been discussed. This is the basis for M&A formation. The conceptually easiest and most graspable tactic to investigate M&A formation is to assume that all companies that want and are able to sell their business, or acquire another one, are present on the M&A market. The companies that want to and are able to sell their business are called potential targets and the companies that want and are able to acquire are named potential bidders (see **Table 1**).

The companies that are more or less active on this M&A market are not random companies. Here, I argue that bidders and targets are different from non-bidders and non-targets. The rationale behind is that often the target, but also the acquirer, has to fulfill certain criteria which are independent from the other firm's characteristics in order to become selected by the other entity. Additionally, not only the expectations of the other party but also the own motivations to acquire another business or sell the own business might be linked to certain firm characteristics. Some researchers kept themselves busy with identifying idiosyncratic bidder and target characteristics. While research on bidders is very limited M&A researchers have been successful in developing target prediction models. This literature has risen over the last 30 years and is mostly financial in nature. Generally, we can distinguish between internal characteristics, such as size or age and external characteristics, such as the characteristics of the sector the company is active in, or of the region in which the company is located.

Due to the fact that an acquisition requires a considerable amount of cash or stock it is not surprising that bidders are significantly larger than non-bidders (Luypaert and Huyghebaert, 2007). This is closely related to the fact that relatively more listed than unlisted firms are active in acquisitions. Listed firms can raise new equity very easily through open-market stock issues and they can pay by stock (Luypaert and Huyghebaert, 2007). In contrast, private firms often cannot raise enough money at one moment and might opt for internal growth investments instead. Another rationale is the agency problem. In large, listed firms managers are not sufficiently controlled by the owners and can therefore more easily acquire another company, also against the interest of the shareholders. Although M&As can be financed by loans (so-called leveraged buy-outs) firms with high portions of bank loans are less likely to acquire, while debt ratio is positively related (Luypaert and Huyghebaert, 2007). Another characteristic is experience. Firms that have exerted some M&A deals in the past are more likely to acquire another company. The fact that experienced firms are more successful (Haleblian and Finkelstein, 1999) speaks for their significant appearance on the M&A market. Finally, Cefis et al. (2009) found that firms with a greater product line scope are more likely to be acquirers. This is because these firms have greater integration capabilities (Mitchell and Shaver, 2003).

Referring to external characteristics a stylized fact is that M&As concentrate in industries in which a regime shift of technological or exposure to deregulation can be observed (Rossi and Volpin, 2003; Gorton et al., 2005). Schoenberg and Reeves (1999) refer to, among others, the deregulation of the insurance and air transportation industry in the EU during the 90s. External factors are not only a company's industry but also its location. Zademach (2005) shows that in Germany M&As accumulate in large cities, which reinforce their role as centers of power and control by this kind of economic transactions. Not only economic agglomerations but also political power is seen as a driver of M&A formation.

While bidders tend to be large, one would intuitively say that targets are smaller than non-targets, if measured by number of employees, market capitalization, market share or sales. This is, indeed, what the studies of Palepu (1986), Gorton et al. (2005), Shen and Reuer (2005) and Brar et al. (2009) demonstrate. Small companies are much easier to buy because they cost less, involve less transaction costs and require less integration efforts. Consequently, some companies that wanted to avoid becoming a target were trying to grow by means of acquisitions themselves and thus avoided being eaten by others (Gorton et al., 2005). Also multilocational targets, i.e. firms with many subsidiaries, are less likely to become targets as it is more difficult to monitor and assess their complex structure (Böckerman and Lehto, 2006).

Additional to the size the legal form of the target is a distinguishing factor as well. On the one hand, in hostile M&As, it would not surprise that bidders prefer unlisted targets because they tend to involve lower transaction costs and they are easier to evaluate (Shen and Reuer, 2005). In fact, companies use initial public offerings (IPOs) to increase their chance of being acquired. In friendly deals, and in the Netherlands almost all deals are friendly, the legal form of the target should not matter much. On the other hand, listed firms provide more information and estimating the right value for unlisted firms is due to a lag of information. Indeed, Song and Walking (1993) found that targets have lower managerial ownership than non-targets.

Table 1 Definition of company roles in M&As.

Potential bidder	=	Company that is able and willing to acquire another company at time *t=0*. I.e. company is on the M&A market at time *t=0*.
Potential target	=	Company that is able and willing to sell its business at time *t=0*. I.e. company is on the M&A market at time *t=0*.
Bidder	=	Company that announced the acquisition of another company at time t=0. Note: (a) Companies that solely bid on a target but did not announce a deal are not considered in this thesis. (b) If the same company announced the acquisition of two or more targets at different points in time this company occurs as two or more different bidders (see multiple bidder).
Acquirer	=	Company that acquired another company at time *t=0*.
Target	=	Company that announced the sale of its business at time *t=0*. OR: Company that got acquired by an acquirer at time *t=0*.
Non-bidder	=	Company that is not able or willing to acquire another company at time *t=0*. I.e. company is not on the M&A market at time *t=0*.
Non-target	=	Company that is not able or willing to sell its business at time *t=0*. I.e. company is not on the M&A market at time *t=0*.
Multiple bidder	=	Company that announced the acquisition of at least two potential targets at *t*=0 and *t*=1.
Multiple acquirer	=	Company that acquired at least two potential targets at *t*=0 and *t*=1.
Deal	=	Dyad of bidder and target that announced a deal. Note: Deals denote the announcement of deals, not the actual deals. For the purpose of simplicity, however, I refer to deals. The same holds for non-deals.
Non-deal	=	Dyad of potential bidder and potential target that did not announce a deal.

A mayor factor seems to be the financial situation of the targets. Firms are more likely takeover targets when their market value is low compared to their book values (Powell, 2004). Referring to the managerial discipline theory it makes also sense that firms with inefficient management, measured by profit margins, return on investments, return on capital and Tobin's Q, are more likely targets (Hasbrouck, 1985; Palepu, 1986; Morck et al., 1990; Powell, 1997; Tsagkanos et al., 2008). Brar et al. (2009), however, could not find much evidence for that. Also a low market to book ratio, a low price - earnings ratio, the possession of accumulated tax losses or high balance sheet liquidity (Schoenberg and Reeves, 1999) and high current financial liquidity (Hasbrouck, 1985) have been associated with target likelihood.

Although the relative amount of targets might differ between different industries the target literature mainly focuses on internal characteristics. Some geographical work shows that a target is more likely to be acquired if it is located in an agglomeration in terms of population and GDP (Rodrígues-Pose and Zademach, 2003). Also Cai and Tian (2009) show that targets in urban areas are 8.3% more likely to receive a bid. He explains that by better information provided by urban firms, less entrenched management teams and higher synergies.

3 Biases in M&A formation

Integrating the theory of M&A formation and the description of bidders and targets as illustrated in the previous chapter this chapter elaborates on biases when it comes to M&A formation. The role of geographical proximity (3.1), industrial relatedness (3.2) but also other factors (3.3) is explicitly discussed.

3

It is almost needless to mention one necessary condition of any M&A formation: the bidder can only select between targets that are known. Thus, the saying "What the eye does not see, the heart does not grieve over" can be translated into "A company can never bid on a target that the bidder's eyes do not see". Given this condition there are many relational factors that facilitate M&A formation, but the underlying reasons are often the same: information advantages, trust and strategic considerations. The idea behind information advantages is that the better and the more information available the more likely M&A formation. This is because, in general, companies are risk-averse and avoid acquiring a firm about which not much information is available. Note, that information advantages are sometimes termed information asymmetries in the literature. In its actual sense information asymmetries refer to superior information available to one party over another, whereas information advantages do not only refer to the quality but also to the quantity of information available. The quality of information is also related to trust. The significance of trust has been proven for joint ventures (e.g. Zutshi and Tan, 2009) and other inter-organizational relations (e.g. Bachmann, 2001). Therefore it should also be crucial for M&As, at least for friendly acquisitions.

Strategy, however, might be one of the most important factors that determine M&A formation. Depending on the bidder's motive to acquire only special types of potential firms are taken into account. This can be related to the target's industry, location, management style, etc. If a firm wants to diversify the target can belong to nearly any industry, if the bidder wants to exploit synergy effects it might look for a company with complementary products, similar production techniques, similar markets etc. If a firm wants to strengthen its local market power probably a competitor located nearby is acquired. Also prospects on a good integration should be crucial if firms seek for benefits. Firms might be more easily being integrated if there is a cultural fit among the managers, but also among the employees. Keeping costs low is also seen as a strategic factor. If bidders are interested in value-creating deals targets should be selected that involve the least costs. Costs can be related to the information seeking process before the deals but also to integration costs after the deal.

3.1 Home bias in M&A formation

Home bias is the bias to act in the own local environment, based on geographical proximity. Geographical proximity per se is just a death physical measurement, but for the formation of company dyads it is not meaningless, although not a prerequisite at all. Its significance is based on various underlying mechanisms. In the M&A formation process model (see **Figure 3**) we assume that bidders either make the decision to actively search for promising potential targets (among known or unknown companies) or that an opportunity can be taken (bid on a known company). Irrespective of whether the potential targets were known or not during the decision to acquire, the home bias of bidders can be attributed to (a) information advantages, (b) familiarity, (c) strategic reasons and (d) localization effects.

Information advantages refer to hard (codified) and soft (tacit) information which are both more likely to be transmitted between geographically proximate entities. Hard information refers, for example, to the potential target's name, activities, financial information, address or financial data – information as used in this research. Hard information can easily be saved and transmitted and therefore transmitted on long distance. However, it can still be argued that even hard information is more likely to be transferred in geographical proximity. Think of lists with potential targets available at local consulting companies. Soft (tacit) information is harder to transmit on long distances because it requires face-to-face contacts. It might refer to feelings or valuable confidential information. This kind of information is even more associated with geographical proximity than hard information because it is easier to meet if both actors are in geographical proximity. In innovation literature geographical proximity is often associated with local knowledge spillovers. While knowledge mostly refers to very specific types of information, geographical proximity in the case of M&As is associated with different types of information.

If companies are explicitly seeking among unknown companies especially hard (codified) information is crucial to identify possible potential targets and to make a selection. If a bidder is looking for a potential target among companies that he knows, soft (tacit) information play a more crucial role. Also for short-listing and the due diligence process soft information is essential, which is eased if the managers can visit and observe each other's facilities and operations more easily. Even financial assessments are easier to make in geographical proximity because they often involve subjective judgments (Böckerman and Lehto, 2006). Having superior information a deal is more likely. If a deal takes place due to opportunity the bidder bids on one of the companies it is tied to. As most business relations take place within geographical proximity this effect reflects in the target selection decision. Summing up, as on each stage of the M&A formation process information is required geographical proximity can affect the final decision making on each stage.

While information advantages should have an effect on average, its meaning depends on the characteristics of the bidder and target. Grote and Umber (2006) show that especially small and relatively opaque targets receive bids from local companies. If soft information is more decisive than hard information this can be explained by the fact that small companies are not required to publish financial information and in general provide much less codified information. Also on the bidder side size determines how crucial information advantages are. For large companies it should be easier to overcome distance. Large firms have more resources during the search and identification phase and the due diligence process, which allows to identify and assess distant targets in a similar way as local targets. Bidders that are experienced with M&As might also be more likely to overcome distance (Chakrabarti and Mitchell, 2008).

Especially for small bidders another factor should be significant: familiarity. Every person has a personal relation to certain locations and persons are prone to act in their own and known environment before they engage in more distant and foreign environments. This is a form of cognitive bias. In a widely appreciated paper Huberman (2001) documents the geographic distribution of the shareholders of seven regional Bell Operating Companies in the U.S. and comes to the conclusion that "people simply prefer to invest in the familiar". This familiarity effect may hold in particular for small, owner-controlled firms. It might even be reinforced by the action and home bias of the vendor, the actor, mostly a company, that helps identifying and selecting targets, because vendors are like all companies more likely to know local firms that are on sale.

While the effects of information advantages and familiarity imply to be unconscious factors, bidders sometimes choose a proximate target strategically, i.e. explicitly involve a spatial and locational element in the target selection decision. Laulajainen (1988) describes in detail the spatial considerations of The May Department Stores Company Inc. and shows that next to management fit, price and other factors also spatial considerations can be a crucial factor in target selection. Reasons might be easing price competition and the possibility to share common assets after the acquisition (Böckerman and Lehto, 2006). This motivation for seeking local targets might especially hold for targets that operate in the same market. The effect of sharing common assets holds for deals which are motivated by synergy effects and for deals in industries with immobile assets. Rodrígues-Pose and Zademach (2006) found, for example, that nearly 50% of the deals within the energy sector were intra-regional, while it was only 25-30% for automotive, heavy manufacturing, IC/TC and chemicals. For small offices, that can easily relocate, sharing common assets is certainly not a great deal. Keeping transaction costs as low as possible can also be seen as a strategy. In general, implementation costs in post-acquisition integration are lower for local deals. When looking at post-acquisition performance Kang and Kim (2008) show that targets returns are higher in local deals, i.e. for targets it would be better to look for a local buyer.

As bidders strategically opt for local targets they can likewise strategically opt for distant targets. Reasons are to penetrate new geographical markets or to enter regions with lower production costs. Also increasing the access to suppliers and buyers, or the location of subsidiaries and competitors might be pivotal in the target selection process.

Economic activity within geographical proximity is often associated with localization economies. This is the effect of industrial clusters on inter-firm activities, which can be related to the notion of opportunity developed in the previous chapter. Imagine a greenhouse company in the so-called Westland in South Holland intents to acquire another greenhouse company. As the Westland hosts more greenhouse companies than any other region the chance to find a suited target is much higher within geographical proximity. Eun and Mukherjee (2006) tests, however, show that the clustering of M&A is only weakly explained by localization economies. Some own research (see "**Additional Research 1**" in the **Appendix**) shows that especially on municipality level bidders are 3.6 to 5.8 times more likely to bid on a target from the same municipality than suggested by the geographical distribution of potential targets. These numbers even include the home bias that stems from other causes and are estimated under the assumptions that bidders can only select between potential targets of the same industry as their real target.

Although bidders can strategically or unconsciously opt for distant targets, the discussed unconscious (information advantages and familiarity), conscious (information advantages and strategy) and external (localization effects) factors should lead to an overall home bias in domestic M&As. Therefore the first hypothesis is

Hypothesis A1

On average, the headquarters of bidders and targets involved in an M&A deal announcement are geographically closer to each other than expected if M&A formation would occur randomly.

This hypothesis is falsified if the average geographical distance between bidders and targets in deals is not significantly lower than in non-deals. As there is some evidence that companies which bid on domestic targets are home biased (Chakrabarti and Mitchell, 2008; Eun and Mukherjee, 2006; Grote and Umber, 2006) it would be rather surprising if this hypothesis could be falsified. While it is difficult to disentangle the conscious and unconscious factors it is possible to exclude the localization effects. As in this case there would be only three underlying causes left, home bias is expected to be weaker, but still being present.

If an average home bias exists it is likely that there are more home biased than distance biased bidders. This is, however, not self-evident. If about the half or even less of the bidders inhere an extremely strong home bias there could still be an average home bias although not most bidders are home biased. So far, we only know that there is an average home bias, but the actual proportions of home and distance biased bidders are unknown. In order to test whether the majority of bidders is home biased or not the second hypothesis is:

Hypothesis A2

The proportion of home biased bidders is larger than the proportion of distance biased bidders.

While previous studies provide evidence on the home bias of companies on a national level there is no evidence yet that bidders are home biased if they select a target from their own region. The regions that are used in that research are provinces, COROP regions and municipalities. I conjecture that bidders are also home biased if they had to select between potential targets of their own region. This might be due to the same reasons as discussed for the domestic level, but as there is no research done yet on such small spatial levels it is difficult to set up a well-supported hypothesis. Instead I formulate a research question:

Research question A3

Does home bias exist on different spatial scales and how strong is home bias on different spatial scales?

3.2 Industrial relatedness bias in M&A formation

Industrial relatedness refers to the degree to which two firms are active in the same industry. This is associated with shared technological experiences and knowledge bases (Knoben and Oerlemans, 2006) and with similar products and markets. M&A literature talks about horizontal mergers if bidder and target are industrially related and about conglomerates if both companies are completely unrelated. While home bias is believed to mainly stem from unconscious causes, industrial relatedness bias is mainly the result of strategic decision, although information advantages might play a little role, too.

Both quantity and quality of information can be associated with industrial relatedness, just as geographical proximity. If both firms are active within the same industry their managers are more likely to know each other and to exchange information, which affects the identification phase (Chatterjee et al., 1992). During the phase of due diligence bidders have an advantage when assessing industrially related targets because the value of the target is more easily to determine. This notion is based on the concept of absorptive capacity, which is the firms' capability to process and value external knowledge. Capron and Shen (2007) argue that targets in the same product market are especially attractive targets.

More often, however, selecting a target from the same or similar industry is a conscious strategic choice. Industrially related deals are related to a better exploitation of operational and market synergies. It is easier to integrate knowledge and combine operations, duplicative functions can be reduced (Capron, 1999) and economies of scales can be realized. In unrelated deals realizing synergies is more difficult and more integration efforts are required, leading to less benefits and higher costs. In fact, investors and strategy researchers value related acquisitions more than unrelated ones (Lubatkin, 1983; Matsusaka, 1993; Flanagan, 1996; Fan and Lang, 2000). The question remains, however, which degree of industrial relatedness is most beneficial.

While deals can be either industrially related or unrelated, industrial relatedness exists on various degrees. If bidder and target are active in exactly the same industry, e.g. in manufacturing of motorcycles, we talk about strongly related deals. If bidder and target are active in related industries we talk about diversifying deals. This is, for example, if a manufacturer of motorcycles acquires a manufacturer of bicycles. As bicycles and motorcycles are transportation vehicles the two industries are related and the motorcycle company would follow a diversification strategy. The main characteristic of diversified companies is that they are adhered to a single primary industry base. For example, the German *TUI AG* is active on the markets of airlines, travel agencies, tour operators, hotels, and cruises. However, all business activities are related to the leisure time industry and have therefore a certain degree of industrial relatedness among each other. In this paper industrially related deals are always horizontal deals. Vertical deals, i.e. deals between supplier and buyer are not seen as industrially related deals.

Although related deals, being strongly related or diversifying, are associated with value creation, we observe that also completely unrelated deals take place. This is often the case if the bidder is a conglomerate or a financial investor. Conglomerations are companies with completely unrelated business activities. For example, the Japanese *Yamaha Corporation* is active, among others, on the markets for musical instruments, electronics, motor cycles and sports equipment. If the *Yamaha Corporation* bought a manufacturer of toothpaste this activity would not be related to any of the businesses of its subsidiaries and therefore being an unrelated deal. However, also these kinds of transactions can be related to value creation due to risk reduction. If the bidder is a pure financial sponsor the relation between the activity of the bidder, mostly a private equity company, and the activity of the target is entirely meaningless. Thus, about all financial deals are unrelated deals.

As diversification strategies are in general more risky than coherence strategies and because they can be motivated by managerial motives especially large and listed companies are expected to exert unrelated deals while small companies are generally more risk-averse. So far, not much is known about the actual industrial relatedness bias of bidders. To my knowledge there is only the paper of Hussinger (2010) that shows the propensity of companies to acquire targets with a related technology portfolio. Although bidders can opt for completely unrelated deals there are good reasons why, on average, bidders should prefer industrially related targets. Consequently the third hypothesis in this paper is:

Hypothesis B1

On average, the business activities of bidders and targets involved in an M&A deal announcement are more related to each other than expected if M&A formation would occur randomly.

Also here it is interesting to look at the proportion of industrially related bidders. We know from past developments that strategic trends, which are often based on belief or herding behaviour, have been changing over time (Bouwens, 2005). For example, while the third US merger wave (1965-1975) was dominated by conglomerates, the first (1895-1904) and second wave (1920-1930) were dominated by horizontal mergers. The actual extent of this bias was however never estimated. Also about the 6th merger wave, which is the time frame of the data used in that work, not much is known. Nonetheless, I expect that most bidders are industrial relatedness biased. Therefore following hypothesis is tested:

Hypothesis B2

The proportion of industrial relatedness biased bidders is larger than the proportion of industrial unrelatedness biased bidders.

Although financial deals are included it would be surprising if those two hypotheses were falsified. Technically, hypotheses B1 and B2 work in a quite similar way as hypotheses A1 and A2 and answer the first part of this thesis' aim. The second part of the aim is to explore the interaction between the two dimensions and to estimate their contribution to the explanation of M&A formation. The proximity concept known in economic geography implies that different relational variables can either substitute or reinforce each other when it comes to dyadic or *k*adic formations (e.g. Boschma, 2005 for the impact on innovation). In the case of M&As it is not known how geographical proximity and industrial relatedness interact. I conjecture that geographical proximity affects M&A formation independently from industrial relatedness because the underlying causes are only partly overlapping. This means that both bidders that bid on related and bidders that bid on unrelated targets are home biased. This interaction is explored by addressing the following research question:

Research question B3:

How do industrial relatedness and geographical proximity interact in predicting M&A formation?

Another question is in how far both dimensions contribute to the explanation of M&A formation. It was argued that M&A formations are rare events and the actual prediction of which two particular companies would try to engage in an M&A deal is about impossible. That means the actual explanatory power of geographical proximity and industrial relatedness on M&A formation is expected to be very small. It is also interesting to compare the explanatory power of both dimensions. As the realization of synergy effects is central in the motives to acquire my guess is that industrial relatedness has by far a larger impact on M&A formation than geographical proximity. In order to address these two issues the last research question is:

Research question B4

What is the single impact of industrial relatedness and geographical proximity on M&A formation and what is their contribution overall?

Given this four hypotheses we can estimate in how far bidders are prone to bid on industrially targets on average and individually. It also guides some tests concerning the interaction of the dimensions and their overall contribution to M&A formation. I expect that the overall contribution of the investigated dimension is fairly small, because there are many others, especially non-relational, factors that affect M&A formation. While non-relational factors root in history and contextual circumstances some other relational factors are identified.

3.3 Other biases

One can imagine that there are several other biases in M&A formation. **Figure 4** lists some relational factors that facilitate M&A formation and mention on which bidder and target characteristics they can be constructed. Except for business relationships, such as buyer-supplier relations, all factors can be constructed based on single characteristics of the bidder and target. As geographical proximity and industrial relatedness these factors might only facilitate but never determine M&A formation. M&A formation is often a product of coincidence and contingency and not a phenomenon that automatically occurs when certain conditions are fulfilled. It is furthermore not known how important the contribution is of each. The factors that are speculated to be very influential are placed at the top of the table while factors with probably weak effects are placed at the bottom.

As most deals take place within the same country I think that institutional similarity is one of the most important factors that facilitate M&A formation. The highest degree of institutional similarity is if bidder and target are located within the same country. In domestic M&As the firms share the same laws, language, a highly similar culture, similar ethnic values, attitudes, style of work and organization of companies. These factors can be related to information advantages, superior integration possibilities and a generally easier deal process. Cross-border transactions should consequently being more likely between companies located in states that are institutionally very similar, for example between Germany and Austria. The results of Cartwright and Price's (2003) study show that culture influences partner selection decisions. For example, companies of the highly individualistic cultures of the USA, U.K. and Northern Europe are more likely to acquire a firm of the same places than a firm from Southern Europe or Asia. Also here institutional similarity can be associated with post-acquisition performance, while it is disputed which degree of similarity is most beneficial (Morosini et al., 1998).

I argued that M&As can be an outcome of opportunities. If two companies are tied by a business partnership, such as a joint venture or buyer-supplier relationship, it might occur to one participant to buy the other company. Therefore, M&A formation likelihood should be higher between companies that are in a business relationship. In fact, buyer-supplier relationships are often associated with M&As in M&A literature and called vertical M&As. Also takeovers of the mother company by a partially owned subsidiary (downstream merger) are possible. M&A formation between companies that are already engaged in a business relationship is not only higher because they already know each other but also because the quality of information available to both parties is much higher than to a third party. It is not surprising that M&As between alliance partners are associated with a better post-acquisition performance (Porrini, 2004).

Usually M&A deals are not between companies of equal size. Deals should occur more often between large bidders and small targets, as small targets are normally cheaper and less risky to buy (Weitzel and McCarthy, 2009). Therefore I argue that M&A formation is more likely if the bidder is larger than the target. Related to that idea is age difference. As often startups are acquired I argue that the bidder is normally older and more experienced than the target.

Not only direct ties between potential bidders and targets can affect deal likelihood, also belonging to the same superordinate or overarching organization can be associated with M&As. Indeed, so-called sister company mergers are not an unknown phenomenon. One example of a sister company merger is the merger between Yamaha Music Central Europe GmbH and Yamaha Music Holding Europe GmbH into Yamaha Music Europe GmbH in 2008. The two companies served different geographical markets and were integrated because their owners, the Yamaha Corporation, wanted to think of Europe as one market (Yamaha Corporation, 2008). Other reasons that sister companies merge might be the similarity in their organizational structure that enables the exploitation of synergy effects and increases the chance of a successful integration.

Many M&A guidebooks preach to look for the so-called management fit between the two companies. If management styles are completely different the integration process is less promising, especially if the board of the target is not replaced. As many managers probably read such books and because deals between companies with similar management styles might lead to higher gains one might expect that the likelihood of an M&A deal is higher between firms with a high degree of similarity in management style. Management styles can refer to resource allocation pattern, such as low cost strategy, or investments in R&D, risk taking attitude, reward orientation, autonomy orientation. Yang and Lin (2005) measured the so-called strategic distance as the degree of dissimilarity between two firm's business strategies.

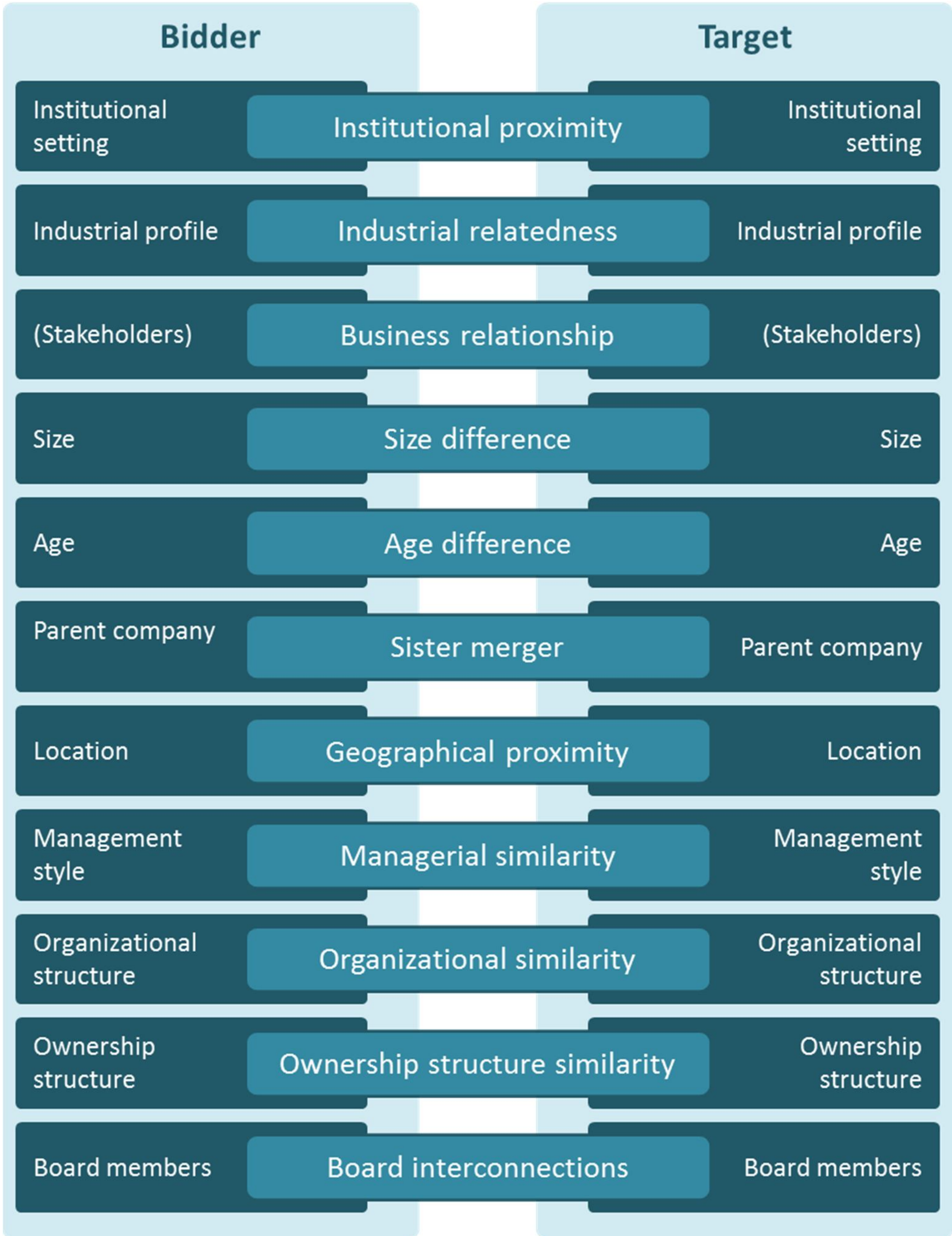
Also different organizational structures and cultures, e.g. a strictly hierarchical structure versus a very flat structure, are difficult to integrate. Therefore one could assume that managers look not only for management but also an overall organizational fit (Graebner and Eisenhardt, 2004). Due to higher success probability organizational fit could have an impact on M&A formation. Organizational fit could, for example, be indicated by equal legal forms. While most targets are unlisted companies one can wonder whether listed targets are more likely acquired by listed acquirers and unlisted targets by unlisted acquirer or not. As listed companies are normally large companies their acquisition requires some experience and resources. Consequently, it seems to be reasonable that the acquirer is more likely a listed company as well. While stock market listing is an aspect of ownership structure there might also be legal aspects that serve as a facilitator.

Other important sources of information that might facilitate M&A formation are social relations between companies. In the last decades research has shown increasing interest in the configuration of social relations within and between companies (Yeung, 2005). Thus, another facilitator might be board interconnections. Board interconnections occur if one board member of one company has served for the other prior to the acquisition, be it simultaneously (interlocking directorship) or in succession. A direct link is what Cai and Sevilir (2009) call a "first-degree connection", while a "second-degree connection" refers to a connection by serving for a third company. I argue that through board connections the acquiring company is not only tied to a potential target but also gets superior information, which allows better assessment and increases. Thus, the likelihood of acquiring the connected firm is higher due to information advantages in terms of quantity and quality. In fact, the existence of information flows between companies that share the same directors can hardly be denied (Mizruchi, 1996; Zajac and Westphal, 1996; Haunschild and Beckman, 1998). If board connections provide superior information we can also wonder whether they are strategically installed with companies that are pre-selected as a potential target or whether acquisitions are an outcome of a connection that came into being for other reasons.

Board connections do, however, not necessarily have to be positively related to M&A formation. Interlocking directors have fiduciary duty to both the bidder and target company. This might constrain their action in the interest of the bidder as well as of the target. In fact, Chapple and Treepongkaruna (2006) found a negative relationship. Own research (see "**Additional Research 2**" in the **Appendix**) has shown that out of 2113 deals only 9 board connections have been found that succeeded an interlock. Due to incomplete data this number is, however, not valid. In the U.S. Cai and Sevilir (2009) identified 65 connected deals out of 1664. It remains unclear which impact board connections have on M&A

formation. Next to board connections one can also imagine that companies are tied by relationships between managers or important employees. Such relationships might originate from shared past employment or educational affiliation, as in Ishii and Xuan (2010). Social relationships can be associated with trust and superior information and thus an increasing likelihood of M&A formation.

Figure 4 **Relational factors facilitating M&A formation.**



The relational variables in the bright fields can be constructed on the basis of the individual bidder and target characteristics indicated in the dark fields (except business relationship).

4 Data and sampling

This chapter starts by explaining which data is used and how the peer group of non-deals is constructed. Then, the sampling procedure is presented, followed by an assessment of the data quality of the ZEPHYR database.

4

4.1 Data collection strategy

This research is based on one basic idea: comparing firm dyads that did engage in an M&A with firm dyads that did not. The first represent the phenomenon that needs to be explained, the latter embody the peer group of dyads that arise if M&A formation occurred randomly. Finding a list of M&As and the involved companies is the least problematic issue in data collection. In fact, two M&A deal dataset are available: The Thomson ONE Banker database[1] and the ZEPHYR database by Bureau van Dijk (BvD)[2]. Usually, empirical papers on M&As use the first one, while the latter one is far from being exploited in scientific research. Both datasets contain information on M&A deals all over the world, but there are crucial differences in scope, information on firm activities and location as shown in **Table 3**. There are three main reasons why I finally opt for ZEPHYR. First, I want to investigate the target selection behavior of a representative set of companies, not only of large and listed ones. Second, Thomson ONE Banker does not provide information on the location of the headquarters within a country, which is indispensable for investigating domestic M&As. Third, in order to construct a co-occurrence based industrial relatedness measure information on all activities occurring in one firm is required. Also the latter is given in ZEPHYR and by taking it from the same database we avoid biases that might occur due to different methods of allocating activity codes to firms.

Table 2 Differences between the databases Thomson ONE Banker and ZEPHYR.

Data content	Thomson ONE Banker	ZEPHYR
Temporal scope	2000 – 2010 for global deals 1985 – 2010 for U.S. deals	2000 – 2010 for global deals
Deal volumes	Mainly large deals	Deals of every size
Place of bidder's and target's headquarter	Country	Country City Postcode (for some countries)
Information on activities of bidder and target	Classification codes for primary activity (SIC)	Classification codes for all activities (SIC, NACE, NAICS)

ZEPHYR provides a list of deals, but the challenge remains to get a peer group of non-deals. In fact, constructing an appropriate list of firms that did not merge requires some wise considerations. Theoretically, every firm could acquire any other firm. Assuming that there are about 60 million companies worldwide, there were 3.6 septillion possible firm dyads. However, most of these firms have never had the chance to consider engaging in an M&A because they have not been on the same M&A market. Consequently, in this work, non-deals can only involve companies that have been active on the M&A market. If companies are included that never considered engaging in an M&A apples and peaches would be compared (King and Zeng, 2001a). Considering only companies that were on the M&A market reduces the number of possible dyads drastically. This number can furthermore being halved by assigning each company either the role of a bidder or the role of a target. As bidders and targets significantly differ from non-bidders and non-targets and also from each other (see 2.3) constructing dyads with distinguishable members is theoretically and empirically meaningful (Kenny et al., 2006).

[1] Thanks to U. Weitzel, Utrecht University, Faculty of Law, Economics and Governance, for providing the data.
[2] Available at Utrecht University Library.

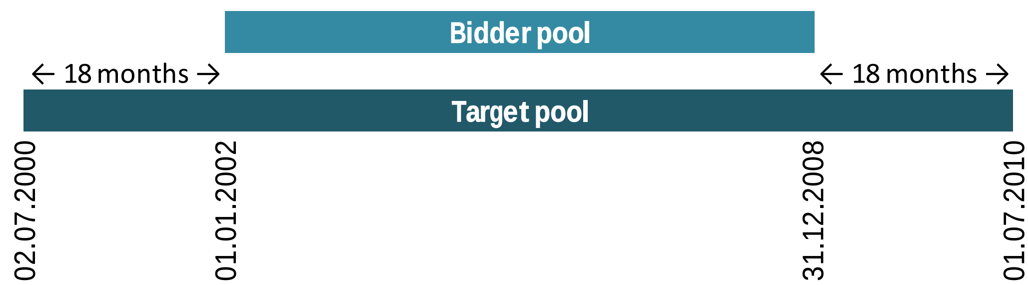
Given the condition that all companies must have been active on the M&A market and either have to take a bidder or target role the approach to identify potential bidders and potential targets is rather straightforward: Every company that bid on another company at a certain time and place is considered as a potential bidder and every company that received a bid at a certain time and place is a potential target. Matching all potential bidders and all potential targets a set of observations is obtained, that set includes non-deals but also the actual deals. Considering solely companies that have been active on the M&A market assures that every potential bidder and every potential target is equally likely to become involved in an M&A deal announcement.

Given this first condition the spatial and temporal dimensions need to be specified. In order to eliminate the effect of institutions, which might have a considerable impact on cross-border transactions, as already explained, only domestic deals are considered. This facilitates the creation of a meaningful set of non-deals because most potential bidders and targets are active on the national M&A market instead of the international market. Considering non-deals between companies that are active on different markets would create many irrelevant observations resulting in an unmanageable amount of observations.

Next to the definition of a spatially closed system temporal conditions are required. Imagine that bidder B_A announces a deal with target T_A at time *t* (day of deal announcement). B_A and T_A form a real deal. The alternatives of B_A are, among others, potential targets PT_B and PT_C. PT_B was involved in a deal with bidder B_B at time *t-x* and PT_C was involved in a deal with bidder B_C at time *t+y*. In order to construct a set of all potential targets for every single bidder (target pool) time span *x* and time span *y* need to be defined. This time span indicates how long a target has been available on the M&A market. It is assumed that a target has been available on the M&A market prior to the day it became involved in a deal. Normally, the decision to bid on a particular potential target starts about 12 months before the deal announcement (Grote and Umber, 2006). This means, during that time the potential target was on the M&A market. Announcing a deal does, however, not lead to an immediate disappearance of the target from the M&A market. As long as it is not acquired it can receive more bids and announce more deals.

Given that targets are active on the M&A market before and after a deal is announced the sizes of time *x* (before deal announcement) and time *y* (after deal announcement) need to be determined. Grote and Umber (2006) use a time window of 18 months before and after the date of deal announcement. I test whether the size of the time span has a significant impact on the outcomes. Consequently, I run the final logistic regression model under assumption x=549 days, y=549 days and x=366 days, y=366 days and x=183 days, y=183 days. As expected this has no significant effect on the outcomes. Furthermore, x=549 days, y=0 days and x=0 days, y=549 days is tested (**Appendix Table 6**). Also here the results are not time-sensitive. As a higher number of non-deals allow more accurate subsampling, resulting in a sufficient amount of observations, I opt for taking the time window of +/- 18 months. In conjunction with the required time lapse we have to select the targets of all deals that were announced between 02.07.2000 and 01.07.2010 (**Figure 5**). These targets can be allocated to all bidders that bid on a target in the years 2002-2008.

Figure 5 **The temporal dimension in the construction of non-deals.**



Construction of non-deals: the time windows for the bidder pool (all potential bidders) and the target pool (all potential targets).

4.2 Data sampling

The data collection strategy determines the sampling procedure. This includes the selection of a country and the selection of appropriate and suitable deals. Research would especially being relevant for countries with high M&A activity. **Table 3** shows the ten largest M&A markets of the world with the amount of domestic and cross-border M&As between 2000 and 2010. The proportion of domestic deals per country varies between 64% and 93%. On a global scale M&A deals are, as most economic processes, spatially concentrated. 69% of all domestic deals take place within these ten countries and 88% of all bidders are from one of these countries.

As it makes sense to investigate any of these countries the actual choice of the site solely depends on the available data. In order to calculate the distance between firms locational information is needed. For most companies there is, however, no information on the postcodes of their headquarters, but the highest coverage is for the Dutch deals. As the postcodes need to be linked to geographical coordinates additional data about the geographical location of postcodes is required, which is only available for the Netherlands[3]. All in all, there is not much choice left for selecting the geographical location. Unsurprisingly, the M&A behavior of Dutch companies is strongly internationally oriented. Only 67% of the deals take place within the own country. This means that many deals undertaken by Dutch bidders are beyond the scope of this work. However, we can take comfort from the fact that for no other country the completeness of data regarding both company location and activity is as good as for the Netherlands.

Table 3 The ten largest M&A markets in the world.

Country	No. of domestic M&As	No. of bidders	Domestic M&As (%)
United States	52,772	63,979	82
United Kingdom	19,158	25,500	75
Russian Federation	9,097	9,831	93
Japan	7,645	8,627	89
France	6,809	10,139	67
Canada	6,500	10,120	64
Germany	6,270	9,744	64
Finland	6,114	7,281	84
The Netherlands	5,936	8,901	67
China	5,831	6,370	92
All countries (N=177)	**181,567**	**242,051**	**75**

The 10 largest domestic M&A markets in the world 2000-2010. The numbers are based on all announced M&As as defined by BvD. The actual numbers of deals is slightly higher because for 13,021 deals out of the total population of 255,072 deals the bidder country was not given in the data although queried (see data quality, 4.3).

ZEPHYR does not only contain M&As as defined in this paper but also other kinds of M&As or corporate restructuring, respectively. **Table 4** gives an overview including the amount of domestic deals in the Netherlands within 2000 and 2010. The only relevant transactions are acquisitions and mergers. However, a close inspection of these cases shows that by far not all of these observations are M&As as defined in this paper.

Having chosen a country as well as the relevant deal types a tailored sampling procedure is needed that not only results in a list of deals, but also enables the construction of non-deals. Both deals and non-deals are based on transactions taking place between 02.07.2000 and 01.07.2010. Therefore this time period is taken as a start of the sampling procedure. When constructing non-deals all potential targets are allocated to each bidder that was engaged in a deal within a 36 month range around the announcement date of the actual deal. This means that only the behavior of bidders between the

[3] Thanks to J.J. Harts, Utrecht University, Faculty of Geosciences, for providing the data.

middle of 2001 and the middle of 2009 can be investigated. In order to investigate complete years the time period 2002-2008 is studied. For the whole period between July 2000 and July 2010 ZEPHYR shows 4540 deals (**Table 5**). A detailed investigation of those deals shows that many of those deals are not M&As as defined in this work. I consider M&As as mergers or acquisitions of 50% or more of the target's stakes. However, not all deals which are coded as mergers or acquisitions by BvD are M&As that fulfill this criterion. Therefore I exclude all stake increases (i.e. the acquirer already acquired a minority or majority stake before 2000), restructurings (coded by BvD as stake increase from 100% to 100%), acquisitions of minority stakes (<50%), acquisitions of unknown stakes and mergers with unknown stakes. These are 342 deals in total.

Table 4 M&A deal types in Dutch domestic M&As.

Deal type	Deals (#)	Deals (%)
Acquisition	5646	54.5
Acquisition minority stake	1607	15.5
Acquisition increased	555	5.4
Merger	410	4.0
Demerger	290	2.8
IBO	282	2.7
MBO	368	3.6
MBI	30	0.3
Share buyback	904	8.7
Joint Venture	267	2.6
IPO	2	0.0
Total	**10361**	**100.0**

The numbers of deals include all announcements of Dutch domestic deals between 2000 and 2010.

Furthermore, I consider M&As as a phenomenon between companies, but it appears that there are many bidders and targets that cannot be clearly identified as legally independent companies. Therefore I exclude all deals in which the bidder's name is unknown (empty field), if the bidder is a private person or family (e.g. "*Mr Roger Houben*" or "*Ven Bekhoven family*"), if it is the government or municipality (e.g. "*Netherlands government*" or "*Leiden municipality*"), if it is a joint venture (e.g. "*Avenu Autogroep BV*" or "*Pon Automotive BV's holding joint venture*"), if the bidder is not in the Netherlands coded as a Dutch company by ZEPHYR, and if the name is unspecified (e.g. "*MBO Team - Netherlands*"). Furthermore, all deals are excluded in which the target's name is unspecified, too (e.g. "*Three regional taxi operators*"), if the target is a governmental institution or not located in the Netherlands, furthermore if the target only consists of certain assets, activities, branches or divisions of a company (e.g. "*Roantex Kantoorinstallaties BV's certain assets*", "*Moes Holding BV's construction materials activities*", "*JSR Multi Design BV's Novell division*" or "*Bruna BV's 10 shops*"). These are 1031 deals. The remaining number is 3167, which is believed to be about the total population of M&A deals in the Netherlands as defined in this work, although a black number is certainly remaining.

It was not possible to clearly identify private equity companies on the basis of their name or NACE code although it appears that many private equity firms (e.g. *NPM Capital*, *HAL Holding NV*, *Gilde Investment Management Benelux BV*) are in NACE class **6499**. As this class does not exclusively contain private equity companies and as I do not know which of the 134 bidders in NACE class **6499** are private equity companies all kinds of financial businesses remain in the sample.

Unfortunately the total population cannot be used due to missing data. For the analysis the geographical location of all companies is required. However, for 198 bidders and 403 targets a postcode is not available or only the postcode of the post office box is given. For these firms the postal address needs to be searched on the web. With the help of the company name I found 183 bidder and 304 target postcodes. For 15 bidders and 99 targets the internet did not show any information on the address, the company name was not distinct or the company name did not even appear on the internet. Therefore 99 cases have to be excluded (For 15 deals both bidder and target location are unknown.).

Explorations of the deals show that in many deals bidder and target headquarter have exactly the same postcode. As every postcode is allocated to on average 17 addresses only, this is fairly suspicious. A match of these suspicious cases with data from the REACH database[4] shows that in most of these cases the bidder shares exactly the same address as the target. I assume that postcode in these cases does not reflect the location before the deal but the location after the deal. Therefore I exclude all cases with identical addresses and all cases for which no address information was available.

Table 5 Sampling criteria and procedure.

Criteria		No. of deals
1	All domestic Dutch mergers and acquisitions that are announced or completed between 02.07.2000 and 01.07.2010	4540
2	Deals that are no M&As as defined in this paper	-342
		=4198
3	Bidder ≠ company (e.g. deals which are management buy-ins but not coded as such)	-365
4	Target ≠ company (e.g. spin-offs or carve-outs)	-683
Population (02.07.2000 - 01.07.2010)		**=3167**
5	Bidder headquarter cannot be localized	-15
6	Target headquarter cannot be localized	-99
		=3068
7	Bidder and target headquarter have the same street name and house number	-372
8	Bidder and target headquarter share the same postcode but no information about the street name and house number is available	-68
		=2628
9	Bidder NACE code on a 3 digit level is not available	-55
10	Target NACE code on a 3 digit level is not available	-15
	Deals from 02.07.2000 - 01.07.2010 which provide all potential targets (target pool)	**=2558**
	Deals from 01.01.2002 - 31.12.2008 which provide all potential bidders (bidder pool)	**=1855**

Next to the geographical location also the activity of both bidder and target needs to be known. The database provides primary and all secondary US SIC, UK SIC, NAICS and NACE codes with about the same coverage. For the analyses I use NACE Rev. 2 codes on a three digit-level. This choice is made because the NACE system reflects the industries of the European economy and is therefore regarded as the most appropriate choice of classification. If using NACE codes on a four digit-level (NACE class), which would be more detailed, many cases would be lost because for 180 bidders and 147 targets NACE class(es) are not available. As the three digit-level (NACE group) is still very detailed it is seen as sufficient to use NACE groups. However, for 55 bidders and 15 targets there is still no NACE class, but only NACE division, given. An own assignment of these cases to NACE classes with the help of online information would lead to biases as the allocation procedure by BvD is not known. Therefore 70 cases in which either the bidder or target NACE group(s) (3 digits) are not available are excluded from the sample.

4.3 Data quality

During the exploration of the data several inconsistencies were noticed (see **Table 6**). Another irregularity is that the data in the ZEPHYR database does not only show a merger wave, but also a merger tsunami at the end of 2010 (**Figure 6**). This huge increase of M&A deals cannot be reliable. Neither global data nor the Thomson ONE Banker database suggests such an enormous eruption of M&A deals. It appears that 167 out of 359 deals in the third quarter and 371 out of 954 deals in the

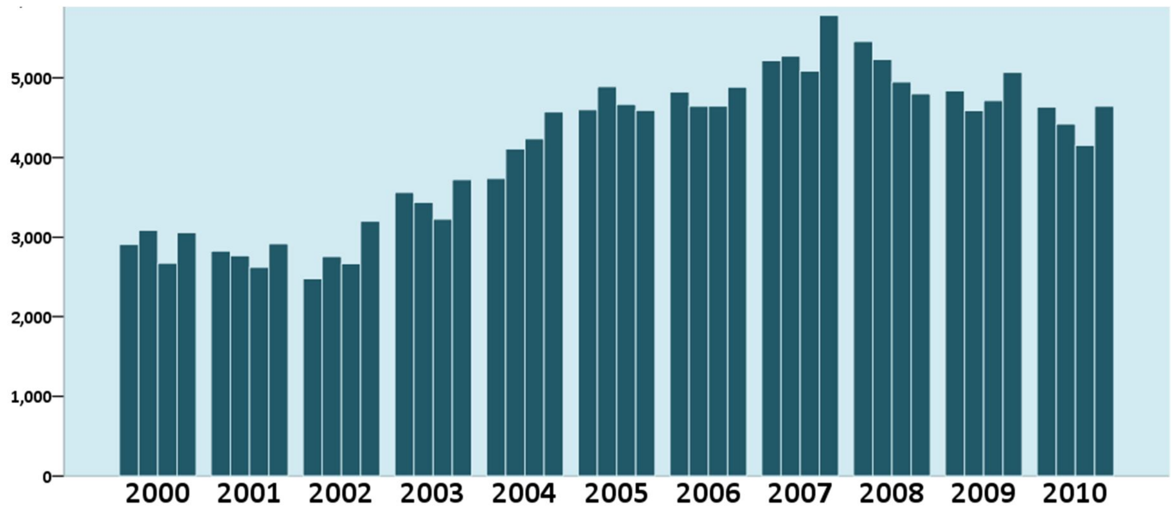
[4] Available at Rotterdam University Library.

fourth quarter of 2010 are actually not M&As but restructurings, although they are coded as M&A deals by BvD. But even without considering those deals the remaining number of deals is unrealistically high. An explanation is owed. Fortunately, this outbreak is just out of the temporal scope of this research. Despite those drawbacks the deals in ZEPHYR are believed to be more complete and representative than the deals in the Thomson database.

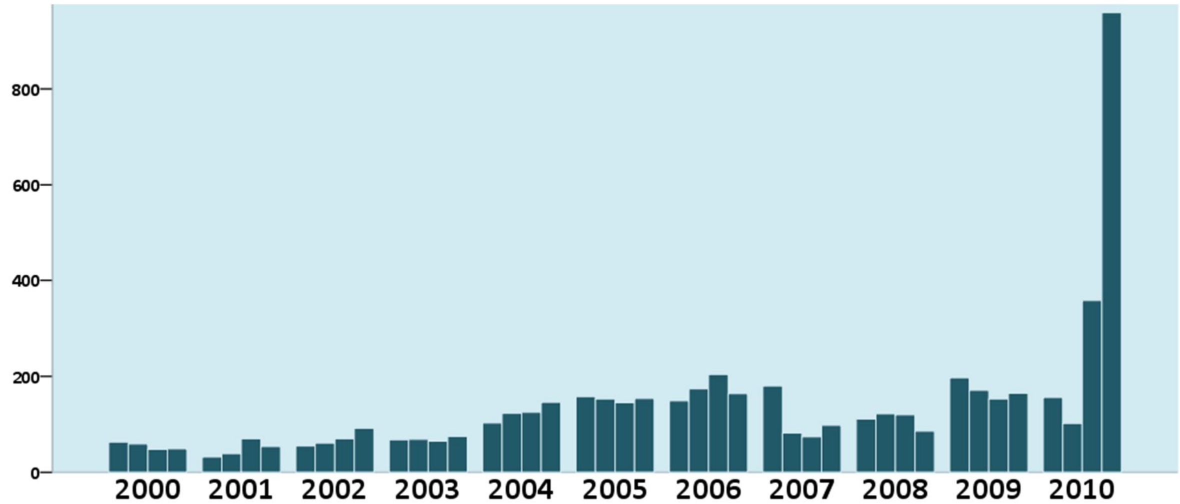
Table 6 Data irregularities in Dutch domestic deals in the ZEPHYR database.

Bidder country
In some cases there is a mismatch between query and actual data. If selecting Dutch domestic deals some foreign firms, mostly bidders, or firms of an unknown country, do still appear. These cases are excluded.
Deal type
Many deals that are coded as mergers or acquisitions are actually just restructurings, as the variable "deal subtypes" suggests. Furthermore, the meaning of some deal coding is unclear (e.g. "Merger 100%" vs. "Merger"). These cases are excluded.
Bidder and target names
In several cases the bidder or target is very unspecific or completely unknown. As many bidder and target names suggest M&As are defined quite broadly. This requires a removal of irrelevant cases by hand (see sampling procedure).
Parent firm of target
The variable which refers to the target parent company often contains the name of the bidder. As a bidder cannot buy its own subsidiary I assume that this information reflects the post-acquisition instead of pre-acquisition situation. Furthermore, it is unclear how complete and reliable this variable is. Therefore this variable can unfortunately not being used in any analysis.
Postcodes of targets and bidders
In many deals the postcode of the bidder and target coincides. Also here I assume that the target postcode reflects the location after instead of before the deal. The suspicious cases are excluded.
Uncomplete values
Compared to other countries there are not that much missing values for Dutch deals. Nevertheless, the coverage of postcode information is rather low, also activity codes are incomplete. This requires an extensive and time-consuming revision of the data.

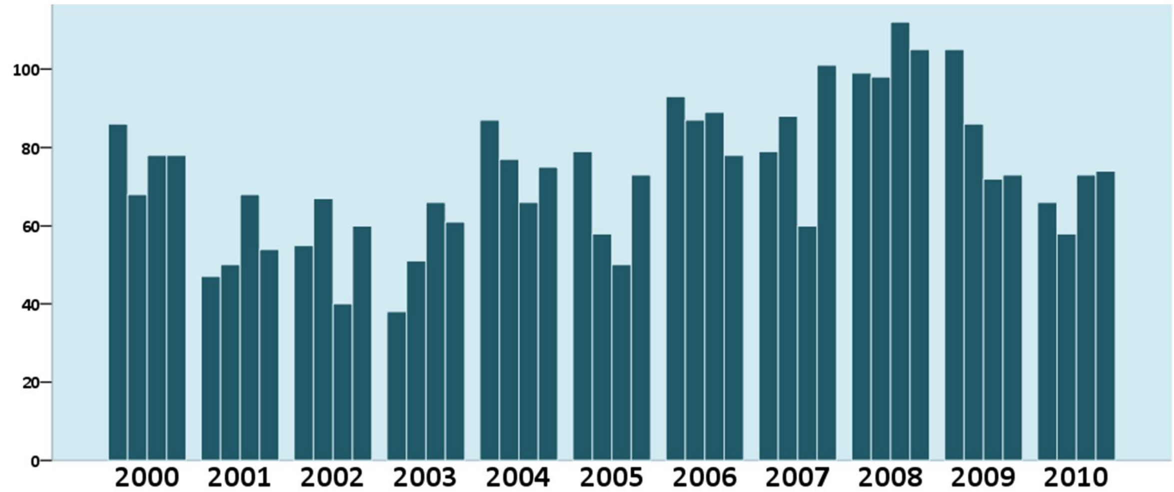
Figure 6 **Domestic M&As 2000-2010 worldwide.**



Domestic M&As worldwide 2000-2010 per quarter, N=181,169.
The graph reflects all Domestic deals coded as M&As by BvD.



Domestic M&As in the Netherlands 2000-2010 per quarter, N=5,901.
The graph reflects all deals coded as Dutch domestic M&As by BvD.



Domestic M&As in the Netherlands 2000-2010 per quarter, N=3,228.
The graph reflects all Dutch domestic deals taken from Thomson ONE Banker.

5 Operationalization

This chapter elaborates on the dependent variable (deal vs. non-deal) and explains which subsamples are of it are tested. Furthermore, the measures of the independent variables, geographical proximity, industrial relatedness and some control variables, are explained.

5

5.1 Dependent variable

The dependent variable simply indicates whether a dyad belongs to the group of deals (1) or non-deals (0). As already mentioned at several places of this thesis deals and non-deals denote deal announcements and not only actual deals. Due to simplicity it is referred to deals and non-deals although announcements and not the actual completed deals are meant. In medical research the 1's would represent the treatment group and the 0's the control group. In this work the 1's denote the actual deal announcements and the 0's the deals announcements that would occur if M&A formation happens randomly. We already know that there are 1,855 deals. Now, we have to construct the group of non-deals following the criteria we developed in the previous chapter. In this way we obtain 1,607,608 non-deals and the total number of observations is 1,607,008 + 1,855 = 1,608,863 deals. As the exact location and industry of every firm is known some explorational tests can be exerted, which refer to the predefinition of the target's industry by the bidder, which is to control for localization economies, and to the regional scale of deals, which shows whether home bias exists on different spatial scales.

For tests on the role of geographical proximity I introduce an assumption that allows to control for localization effects, i.e. for the fact that the selected industry is differently distributed than all other companies. So far, it is implied that bidders are open to select a company of any industry. This is subsequently called assumption ALL (**All** potential targets can be selected). However, many firms might decide on the industry of the target firm before the target identification process, because this is closely related to the bidder's motives (see theoretical discussion in 2.2). In order to consider that as an alternative option I make the assumption that a bidder can only choose between partners that are from the same industry as the partner which is actually selected. Applying this criterion, subsequently called SG (Only potential targets of the **s**ame NACE **g**roup can be selected), we obtain 33,212 industry-matched non-deals. Note, that under this assumption also the group of deals became smaller because for 89 bidders there is no alternative target of the same industry as the real target available, i.e. the bidder would not have had any choice to select any other target. For the test of industrial relatedness and for all multivariate tests this assumption is not applicable.

Furthermore, all tests are done for four spatial levels. So far, it is implied that bidders would select partners from all over the Netherlands but one could also assume that a bidder requires a target in the same region, be it due to better integration, because of gaining access to a certain geographic market or other reasons. Therefore, I create subsamples that contain only deals between companies that are located within the same region. These regions are provinces, COROP regions and municipalities. The deals in overarching regions always include the deals on lower regional scales. The numbers of deals and non-deals considering the alternative assumptions are given in **Table 7**.

Table 7 Number of deals and non-deals on different spatial scales (ALL and SG).

Spatial level of deals	Deals	Non-deals (ALL)	Non-deals (SG)
Netherlands	1,855	1,607,008	33,212
Intra-province	654	210,995	4,988
Intra-COROP region	430	87,008	2,377
Intra-municipality	233	25,301	896

5.2 Independent variables

5.2.1 Two measures of geographical proximity

The basis of geographical proximity is the location of the headquarter location of the target and bidder. Many firms keep several subsidiaries or plants at different locations. It might be possible that the location of such subsidiaries might have an influence on the decision to select a particular firm that is located nearby such a subsidiary. However, we do not have data on the location of subsidiaries. Therefore I use as most other studies only the headquarter locations to determine geographical proximity between two firms.

5.2.1.1 Intra- and inter-regional deals

A straightforward way of measuring geographical proximity is to indicate whether the headquarters of two firms are located in the same region (intra-regional deal) or not (inter-regional deal). I use three non-overlapping types of regions for the Netherlands: provinces, COROP regions and municipalities. There are 418 different municipalities, 40 different COROP regions and 12 different provinces. COROP regions always consist of several municipalities and are always within the same province; provinces always consist of several COROP regions. COROP regions were developed by the Coordination Commission Regional Research Programme (**Co**ördinatie Commissie **R**egionaal **O**nderzoeks**P**rogramma) in 1971 and have hardly been changed since then, while municipalities had been changing from time to time, e.g. by mergers.

This hierarchical classification of regions can be used to construct a measure of geographical proximity. If two headquarters are within the same municipality geographical proximity is high, if they are in different provinces geographical proximity is low. Based on this logic a variable is constructed with 4 categories (see **Table 8**). The categories are mutually exclusive in order to avoid biases by subordinate spatial levels. If a superordinate category contains deals within smaller spatial units deal likelihood might be determined by the fact that the firms are located within the same smaller spatial unit rather than in the larger one. Thus, the variable controls for the effect that deal likelihood is high because of subordinate regions. In the analyses the variable SPCM (**S**tate**P**rovince**C**OROP**M**unicipality) is divided into dummy variables S, P, C and M.

Table 8 Geographical proximity in categories.

Value	Spatial scale of the deal	Geographical proximity
S	Inter-province \ Intra-province \ Intra-COROP \ Intra-municipality	Low
P	Intra-province \ Intra-COROP \ Intra-municipality	Moderately high
C	Intra-COROP \ Intra-municipality	Moderately high
M	Intra-municipality	High

5.2.1.2 Distance

Imagine that two firms are both located near the border of two adjacent provinces. The actual geographical distance would be very low (i.e. high geographical proximity) but S, P, C and M would all be 0 (i.e. low geographical proximity). Thus, SPCM depends on the structure of the different regions and using the actual distance would be a more accurate way of measuring geographical proximity. High distance indicates low geographical proximity and a low distance high geographical proximity. Distance can be calculated by the Great Circle Distance Formula (Pearson, 2011):

$$d_{i..j} = r\left(2\arcsin\left(\sqrt{\left(\left(\sin\left(\frac{radians(lat_i)-radians(lat_j)}{2}\right)2\right)+\cos(radians(lat_i))\cos\left(radians(lat_j)\right)\left(\sin\left(\frac{radians(long_i)-radians(long_j)}{2}\right)2\right)\right)}\right)\right),$$

with $r = 6378\ km$ as the radius of the earth and lat_i and lat_j as latitude coordinates and $long_i$ and $long_j$ as longitude coordinates of the headquarters *i* and *j*.

In fact, the available data allows a very accurate calculation of the distance between bidders and targets. For most Dutch companies the 6-digit postcode (four numbers and two characters) of the headquarter is known (for a few cases the 4-digit postcode has to be used). Every postcode out of the total amount of approximately 440.000 different postcodes is allocated to on average 17 addresses. This means that the postcodes represent the actual location very well. A dataset from the University Utrecht provides RD coordinates (**R**ijks**d**riehoekcoordinaten stelsel) for most postcodes from 1993. As some new postcodes have been added I have to search their coordinates manually. Via Google Maps I find the longitude and latitude coordinates for the missing cases. The given RD coordinates are transformed into longitude and latitude coordinates (WSG 84) with the help of ArcGIS.

5.2.2 Two measures of industrial relatedness

Industrial relatedness refers to horizontal and conglomerate mergers. A high degree of industrial relatedness means that the involved firms possess an overlapping basket of industrial resources (horizontal merger). These resources refer mainly to products and production technologies or required human capital. If the degree of industrial relatedness is low the involved firms have not many industrial resources in common (conglomerate merger). In conglomerate mergers the two firms might be similar in different ways, however, for example by using the same distribution channels. Vertical relatedness (buyer or supplier relationship) as in Fan and Lang (2000) is not covered by industrial relatedness.

The measures of industrial relatedness are based on the primary industries or activities in which the firms are active. These are usually classified within different systems. Such systems are the four-digit **US** **S**tandard **I**ndustrial **C**lassification (US SIC) system for the US, which is the most often used system, or the UK SIC for business establishments in the United Kingdom, the six-digit **N**orth **A**merican **I**ndustry **C**lassification **S**ystem (NAICS) which has largely replaced the US SIC system in Canada, Mexico and the United States or **N**omenclature statistique des **a**ctivités économiques dans la **C**ommunauté **e**uropéenne (NACE), which is an European industry standard classification consisting of a four-digit code developed by the EU on the base of the **I**nternational **S**tandard **I**ndustrial **C**lassification of All Economic Activities system (ISIC Rev. 3) by the UN. NACE Rev. 2 has been applied since 01.01.2008. The ZEPHYR database contains codes of several systems with about the same coverage. I regard the NACE Rev. 2 system as the most appropriate one.

As firms can be present at several locations, firms can also be active in several industries. This means that some firms have several industry codes. Morck et al. (1990), for example, compared the SIC among their codes of the involved firm's top three activities. Although information on all industries is available and some related connections might remain undetected I only use the primary industry code in order to keep the analyses as simple as possible. This leads to a bias, but this is believed to be small as most Dutch firms engaged in M&As are only active in one industry.

5.2.2.1 Intra- and inter-industry deals

The standard method of measuring industrial relatedness is to compare activity codes within the hierarchy of classification systems. NACE codes can be assigned to sections, division, groups and classes (**Table 9**). The classification level of classes is mostly quite detailed. For many firms in the dataset only the group is available. Therefore I do not make use of classes. Following the same logic as behind SPCM a variable USDG (**U**nrelated**S**ection**D**ivision**G**roup) with mutually exclusive categories is created (**Table 9**). Also this variable is divided into dummy variables, which are U, S, D and G. Whenever the primary activity of a bidder and target is in the same section, division or group, the deal can be regarded as a horizontal merger and as conglomerate if the deal is between companies of different sections (inter-section).

Table 9 The structure of NACE Rev. 2.

Level	N	Notation and range	Example
Section	21	[A … U]	"Information and Communication"
Division	88	[01 … 99]	"Publishing activities"
Group	272	[01.1 … 99.0]	"Publishing of books, periodicals and other publishing activities"
Class	615	[01.11 … 99.00]	"Publishing of newspapers"

Table 10 Industrial relatedness in categories.

Value	Spatial scale of the deal	Industrial relatedness
U	Inter-section \ Intra-section \ Intra- division \ Intra-group	Low (conglomerate)
S	Intra-section \ Intra- division \ Intra-group	Moderately high (horizontal)
D	Intra-division \ Intra-group	Moderately high (horizontal)
G	Intra-group	High (horizontal)

5.2.2.2 Tanimoto coefficient

As SPCM depends on the structure of regions, USDG depends on the structure of the NACE system. This can lead to unobserved connections between firms and to observed connections which are actually not related. On the one hand, two firms which are both active in the same section or division do not necessarily need to be related. On the other hand, two firms which do not share the same section or division can still be related. In order to determine the industrial relatedness in a more accurate way, not based on the hierarchy of the NACE system, I construct a co-occurrence measure. The idea is that the more often two activities occur within one company the more related they are. The underlying rationale is that if two activities often occur within one company they probably refer to similar products, production technologies or required human capital and other similarities that are associated with synergy effects.

Similarity measures can also measure the co-occurrence of two identical activities. If a similarity measure of two identical activities *i* and *i* is lower than the combination of this activity with another one *j* it means that activity *i* belongs to *j* and does, for any reason, not occur alone. For example, in the ZEPHYR database (bidders worldwide) it can be seen that only 136 firms are exclusively active in software publishing, but 24,557 firms are active in software publishing and computer programming. We can interpret this as a dependence of software publishing and computer programming. The same would hold for musical instruments. Whereas piano and violin occur very often together, two pianos can be seldom heard.

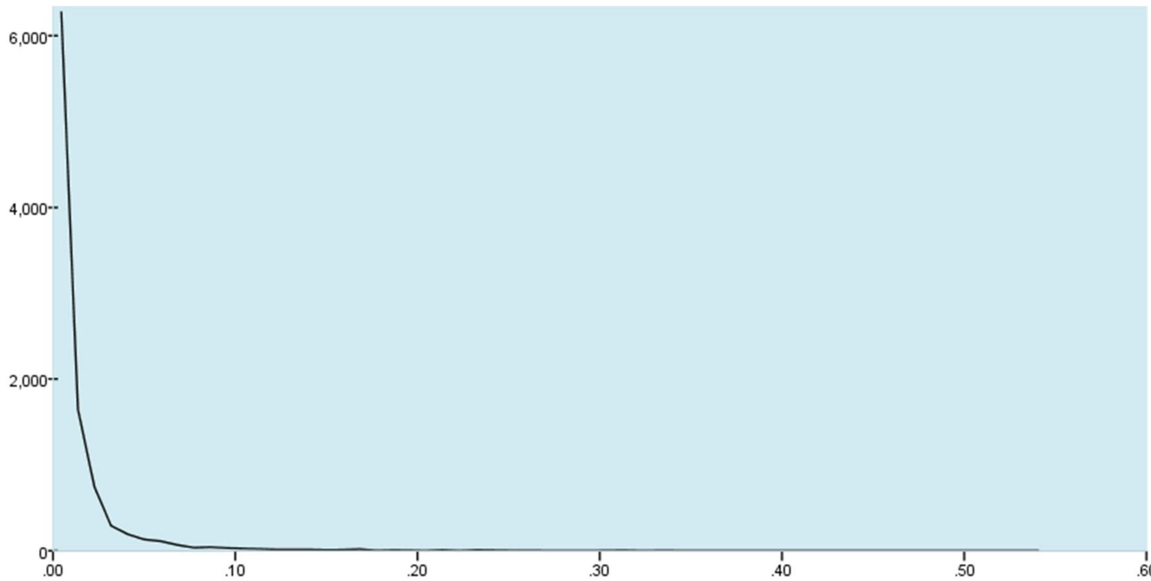
There are many measures that measure the similarity between objects based on co-occurrence. The Jaccard index, Tanimoto coefficient, Cosine similarity, Sørensen's quotient of similarity or Mountford's index of similarity are some examples of so-called similarity measures. Similarity measures can be used whenever the similarity between objects is studied and are often used in the field of ecology and biogeography (Real and Vargas, 1996; Weitzel and McCarthy, 2009) or citation analysis (Hamers et al., 1989). In literature that deals with the relatedness of firms they seem to be hardly used. The study of Broekel and Boschma (2009) uses Cosine similarity, which measures the similarity between two vectors by measuring the cosine of the angle between them (Garcia, 2006), for testing the proximity paradox in the Dutch aviation industry. Bottazzi and Pirino (2010) use Monte Carlo-p scores to show the diversification patterns of firms. In this work I use a powerful but simple measure, which is the Tanimoto coefficient (Tanimoto), also known as the Extended Jaccard Index. The Tanimoto coefficient for the combination of activities *ij* within company *a* is

$$Tanimoto_{ij} = \frac{\sum_{a=1}^{n} ij_a}{\sum_{a=1}^{n} i_a + \sum_{a=1}^{n} j_a - \sum_{a=1}^{n} ij_a}.$$

The Tanimoto coefficient is always within the range [0, 1]. It is 0 when activities *i* and *j* never co-occur in one company and 1 if activity *i* always co-occurs with activity *j*. As activity I use the NACE group code of the firm's primary activity. As there are 272 different NACE groups each bidder selects a target between one of those 272 codes. As the bidder has also one of the 272 activities there are 272*272=73,984 possible combinations of NACE groups. For each combination the coefficient needs to be calculated.

The best way to calculate reliable coefficients for all 73,984 combinations is to use as much data as possible. I decide to retrieve data from the same dataset the sample is taken from because this avoids potential biases as we do not know how the NACE groups are allocated to firms. I selected all bidder companies with one, two, three and four activities from all over the world. Selecting companies with more than four activities would not only lead to a strong bias towards these companies, it is also believed that two activities within a company with many activities are less likely to be related. Companies with only one activity are also selected because also the calculation of industrial relatedness between two companies that share the same NACE code is required. In total, we have 271,187 firms with one activity + 91,982 firms with two activities + 44,684 firms with three activities + 19,068 firms with four activities = 426,921 firms and correspondingly 271,187 + 2*91,982 + 3*44,684 + 6*19,068 = 611,633 cases. Out of the 73,984 possible combinations 13,624 can be found in the real world data, i.e. Tanimoto>0. Only very few of them have a high Tanimoto coefficient; most coefficients are very close to zero (**Figure 7**).

Figure 7 Frequency of Tanimoto coefficients.



Frequency (y-axis) of Tanimoto coefficients (x-axis) for the 13,624 related activity combinations.

It is important to state that this measure is believed to be independent from the dependent variable. One might suspect that two activities co-occur as a result of an acquisition, which would result in an endogeneity problem. This would mean that a circular cause effect bias would influence the results: We would explain the selection of a related target by a variable that is constructed on the basis of bidders of the same industry which has acquired similar targets. However, an analysis of multiple bidders shows that the composition of secondary NACE codes do not change after M&A deals. This is because acquired companies are normally just new subsidiaries of the acquirer without transferring their activity codes to the acquirer. REACH allows to follow some variables over time from 2003-2007 and it appears that also after the deal most targets remain operating as economically independent entities although control and ownership has changed and the target has stopped to exist as a legal entity.

5.2.3 Control variables

Geographical proximity and industrial relatedness are the only dimensions discussed in the theoretical framework we can measure with the data available. However, some more relational variables can be constructed in order to control for unexpected effects. Those variables build upon the legal form, number of locations and number of activities of each firm. It might be that deals are more likely between two firms that are either both listed or both unlisted. In the M&A literature the first type is called public takeover, but this kind of deals does not occur in our data of deals. Therefore there is no need do control for this. The latter type is called public takeover and indicated by the dummy UU (**U**nlisted **U**nlisted) =1. Furthermore, I control for deals in which both firms are either multi-locational or single-locational. If both firms have at least one subsidiary MuMu (**Mu**lti-locational **Mu**lti-locational) =1 and if both firms have no subsidiaries SiSi (**Si**ngle-locational **Si**ngle-locational) =1. A third pair of binary variables indicates whether both firms are diversified or undiversified, respectively single- or multi-businesses. DiDi (**Di**versified**Di**versified) =1 if both firms have at least two different NACE groups and UnUn (**Un**diversified**Un**diversified) =1 if only one NACE group is assigned to both firms. Next to the variables that indicate the same characteristic for both firms, also other combinations are possible (see **Table 11**).

6 Description of deals and companies

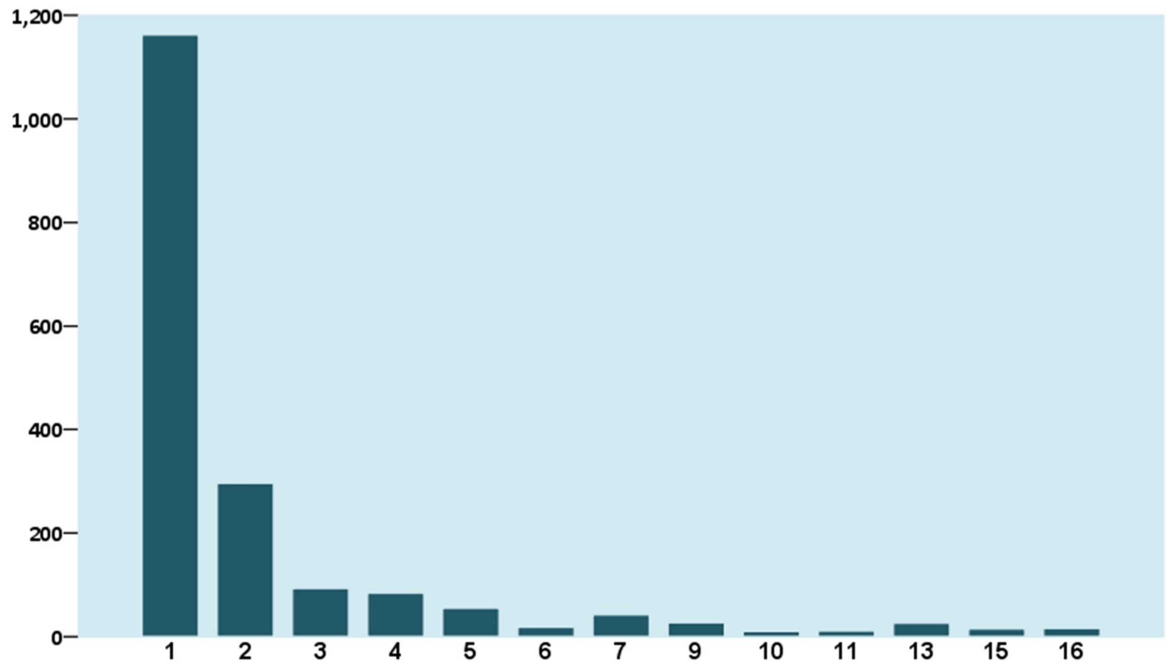
In this chapter the features of deals and non-deals are presented (6.1) as well as some characteristics of the bidders and targets (6.2).

6

6.1 Deals

In total 1,855 deals in the years 2002 to 2008 are identified. Those deals were undertaken by 1391 different companies. 29 companies announced a deal with 5 or more targets, 52 companies with 3 or 4 targets and 148 companies with 2 targets. We have 1,162 bidders that announced a deal with only one target during the period between 2002 and 2008 and 693 bidders with two or more targets (see **Figure 8**). To avoid confusion, bidders do not denote different companies but a company per deal (see **Table 1**). That means if the same company announced a deal with two targets this company occurs as two different bidders in the dataset. As the temporal scope of the dataset is limited it is not known how in how much deal announcements a bidder was involved before 2000. Interpretation is furthermore difficult as the age of the bidders is not known.

Figure 8 **The amount of multiple bidders.**



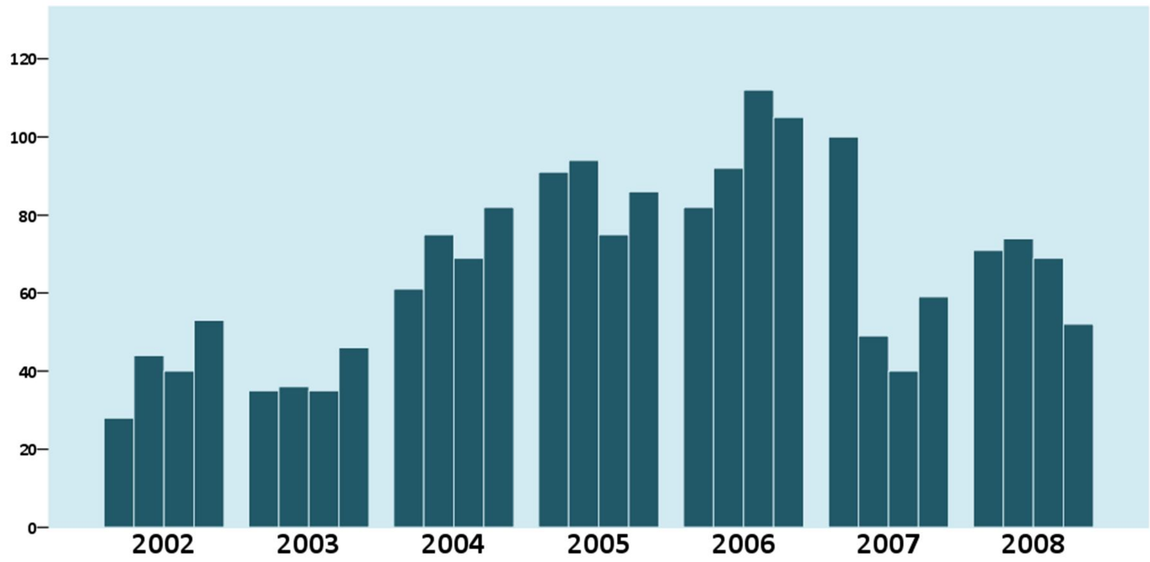
Frequency (#) of multiple bidders (y-axis) with their number of targets (x-axis), N=1,855.

Figure 9 shows all announced deals per year and quarter. The number of deals is varying strongly over time. As everywhere in this thesis deals refer to announced, not only actually completed, deals. The proportion of uncompleted deals is on average 24.1%, ranging from 19.4% in 2006 to 30.3% in 2002. 2002 was not only a year with many uncompleted deals, also the number of deal announcements was very low. M&A activity then took off until the economic crisis starting in 2007. This period of increased M&A activity equates to the sixth merger wave that has been globally identified by the so-called merger wave literature. In the U.S. the first wave is generally dated in the 1890s, the second wave between

1916 and 1929, the third wave at the end of the sixties, the fourth wave in the eighties and the fifth wave at the end of the nineties (Matsushima, 2001; Bouwens, 2005; Gorton et al., 2005).

Every wave has its unique characteristics and asks for different explanations. The fifth merger, dated from 1993-2000, is characterized by many cross-border transactions, huge transaction volumes, an increased involvement of the service sector, hostile takeovers and many leveraged buy-outs (Matsusaka, 1993). It ended with the burst of the Millenium Bubble. Only two years later the sixth wave started to rise. This recent wave is characterized by many financial deals and is likely to be caused, or at least facilitated, by the availability of low-interest financing and abundant liquidity. Consequently, many leveraged buy-outs took place. This implies that many large deals were paid by cash instead of stocks (Alexandridis et al., 2010; Lipton, 2006).

Figure 9 **Dutch domestic deals 2002-2008 per quarter.**

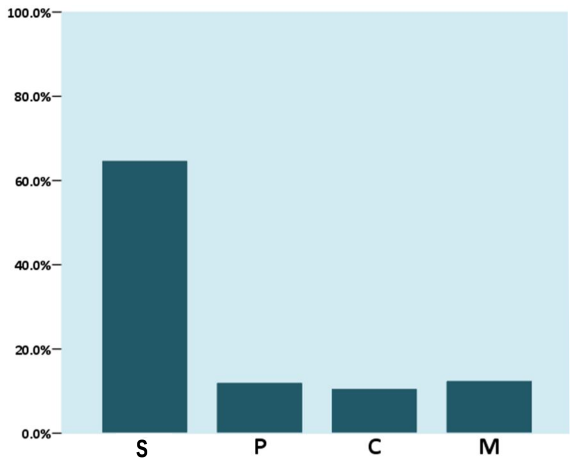


Absolute number of deals per quarter, N=1,855.

A comparison of the values of the geographical proximity and industrial relatedness variables between the group of deals and non-deals allows first insights into their significance. **Figure 10** shows the distribution of SPCM and Distance. It becomes clear that deals occur more often in geographical proximity than non-deals. While the distribution of deals follows a decaying function the distribution of non-deals is close to normal. Also for USDG and Tanimoto the differences in industrial relatedness seem to be significant (**Figure 11**).

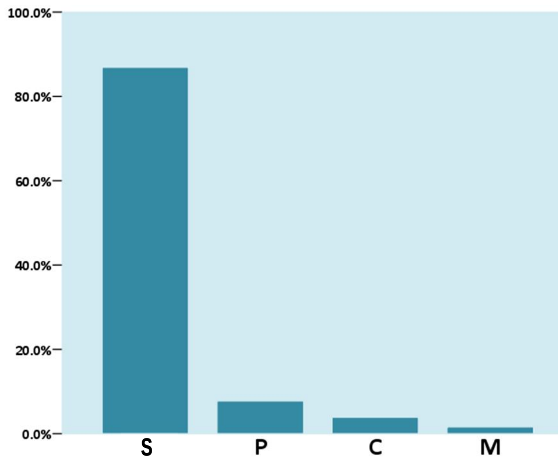
Table 11 shows the frequencies of the control variables. It is not surprising that most deals take place between two unlisted companies, which is to the simple fact that most companies are unlisted. Notably, there was no single deal between two listed companies, but five listed targets were engaged in a deal with an unlisted bidder and 173 unlisted targets with a listed bidder. When looking at the number of subsidiaries per company we see that about the half of all deals take place between a multi-locational bidder and a single-locational target. About 35% of the deals take place between single-locational companies and only in 75 cases between a single-locational bidder bid and a multi-locational target. While concerning listing and number of subsidiaries deals seem to be not much different from non-deals this is different concerning diversification. If the bidder is diversified the real target seems to be more likely to be diversified as well. Regarding deals about 60% of all cases take place between undiversified companies.

Figure 10 Geographical proximity in deals and non-deals measured by SPCM and Distance.



Deals

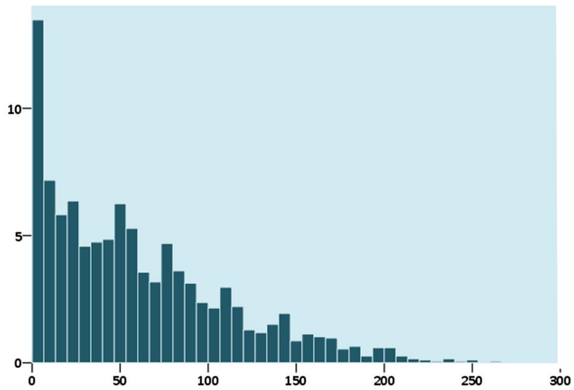
N=1,855.



Non-deals

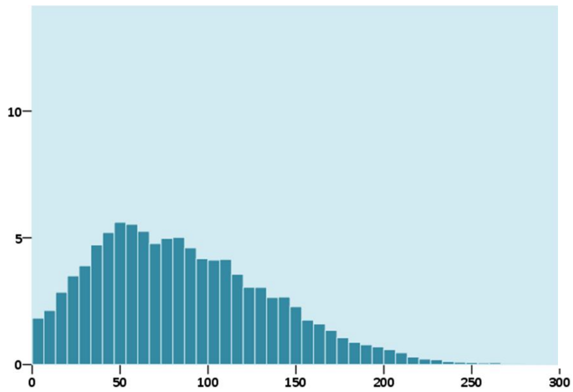
N=1,607,608.

Frequencies (%) of SPCM.



Deals

Max=308km, Mean=60.5km, SD=50.6km. N=1,855.

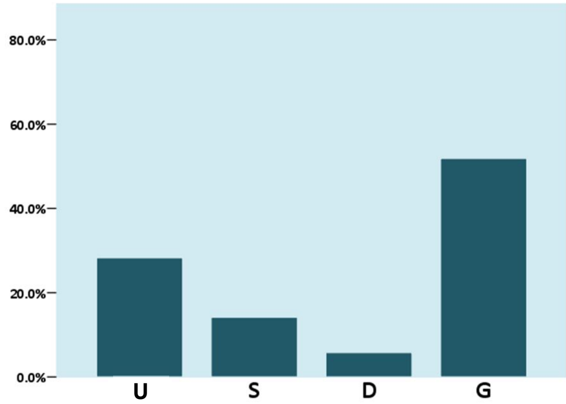


Non-deals

Max=282km, Mean=87.5km, SD=50.7km. N=1,607,608.

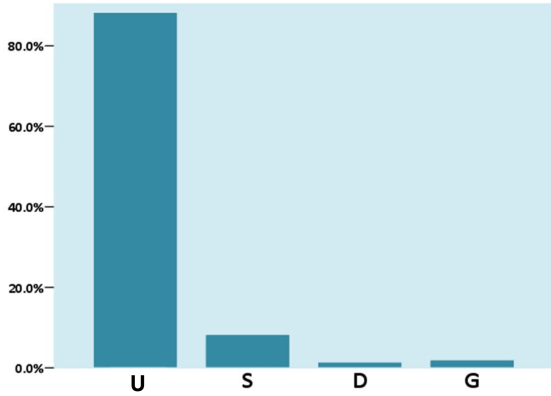
Histograms (%) and descriptives of Distance.

Figure 11 **Industrial relatedness in deals and non-deals measured by USDG and Tanimoto.**



Deals

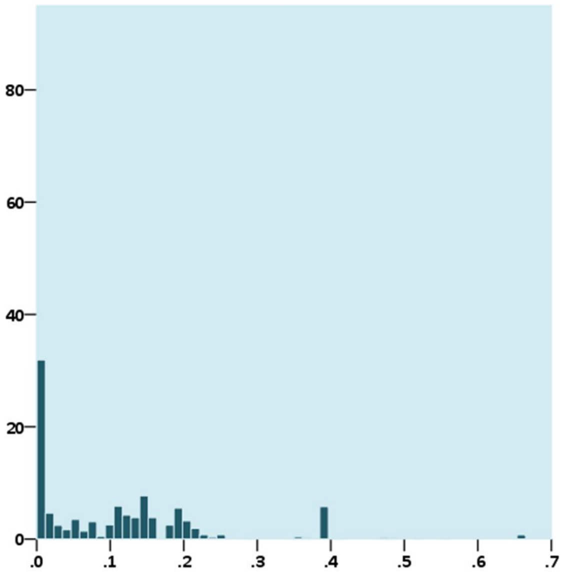
N=1,855.



Non-deals

N=1,607,608.

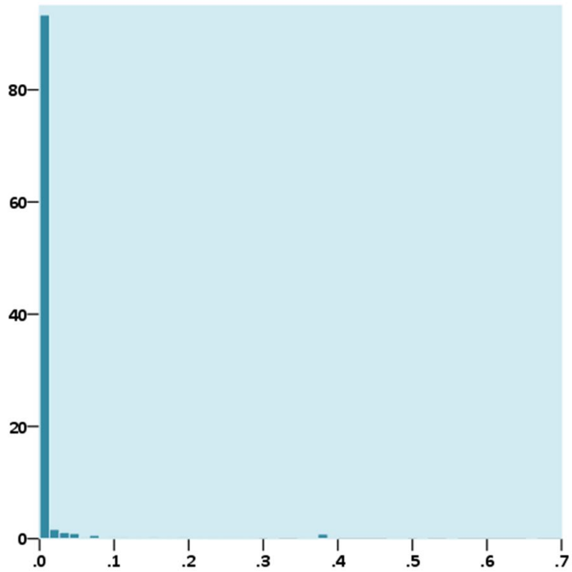
Frequencies (%) of USDG.



Deals

Max=0.658, Mean=0.108, SD=0.118.

N=1,855.



Non-deals

Max=0.658, Mean=0.009, SD=0.047km.

N=1,607,608.

Histograms (%) and descriptives of Tanimoto.

Table 11 The control variables and their frequencies in deals and non-deals.

Measurement	Bidder	Target	Variable	Deals	Non-deals
Listed/unlisted	Listed	Listed	LL	0.0%	0.0%
	Unlisted	Unlisted	UU	90.4%	90.6%
	Listed	Unlisted	LU	9.3%	9.1%
	Unlisted	Listed	UL	0.3%	0.2%
No. of subsidiaries	0	0	SiSi	34.6%	33.0%
	≥1	≥1	MuMu	10.0%	8.6%
	0	≥1	SiMu	4.0%	5.2%
	>1	0	MuSi	51.3%	53.1%
No. of activities	1	1	UnUn	58.4%	55.2%
	≥2	≥2	DiDi	13.3%	6.8%
	1	≥2	UnDi	11.5%	16.1%
	≥2	1	DiUn	16.7%	21.9%

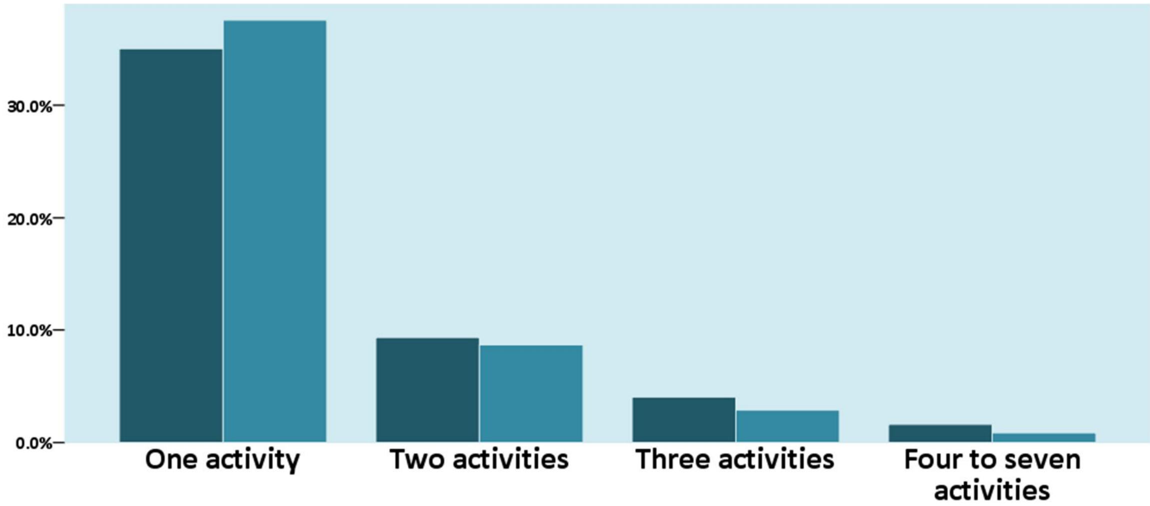
6.2 Bidder and target characteristics

The descriptions of the bidders and targets refer to all deals from 2002-2008 and allow a comparison between bidders and targets; a comparison with between bidders and non-bidders and targets and non-targets would be beyond the scope of this work. It stands out that bidders are more often financial companies than targets (**Table 12)**, which is in line with the argument that many deals are pure financial deals. Another distinguishing feature is that 9.3% of all bidders are listed while only 0.3% of the targets were traded on the stock market. Also this is in line with the theory. Furthermore, targets seem to have less subsidiaries and being less diversified than bidders (**Figure 12**). This is because bidders are generally larger than targets. **Figure 13** shows the geographic location of the headquarters of bidders and targets. It seems that targets are more often in peripheral locations than bidders.

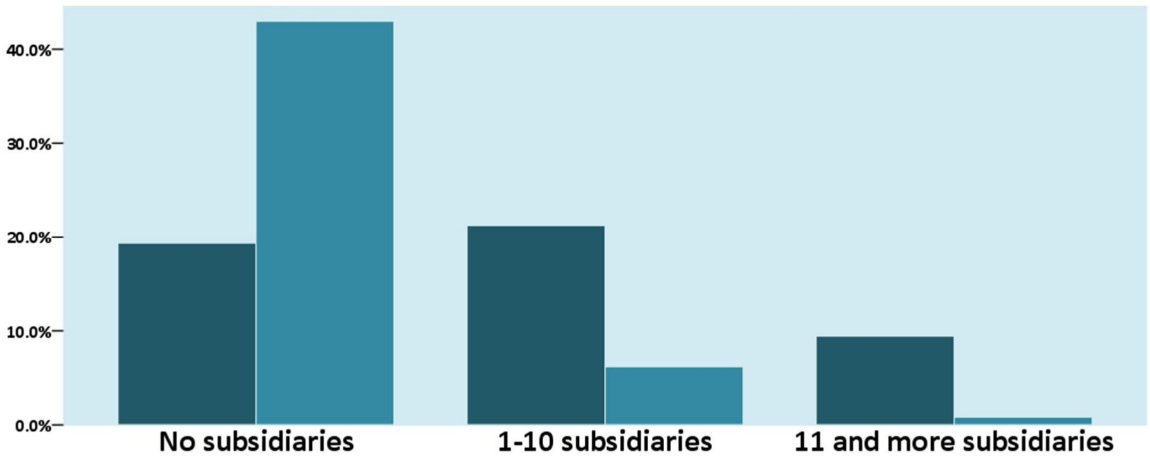
Table 12 The primary industries of bidders and targets.

NACE Section	Bidders		Targets	
	N	%	N	%
A: Agriculture, forestry and fishing	15	1%	19	1%
B: Mining and quarrying	2	0%	5	0%
C: Manufacturing	305	16%	297	16%
D: Electricity, gas, steam and air conditioning supply	16	1%	11	1%
E: Water supply; sewerage; waste l and remediation activities	19	1%	27	1%
F: Construction	119	6%	112	6%
G: Wholesale and retail trade; repair of motor vehicles and motorcycles	325	18%	366	20%
H: Transporting and storage	78	4%	81	4%
I: Accommodation and food service activities	47	3%	55	3%
J: Information and communication	323	17%	342	18%
K: Financial and insurance activities	212	11%	100	5%
L: Real estate activities	33	2%	30	2%
M: Professional, scientific and technical activities	132	7%	134	7%
N: Administrative and support service activities	126	7%	152	8%
O: Public administration and defence; compulsory social security	2	0%	1	0%
P: Education	21	1%	28	2%
Q: Human health and social work activities	39	2%	46	2%
R: Arts, entertainment and recreation	25	1%	32	2%
S: Other services activities	16	1%	17	1%

Figure 12 **Diversification and multi-locationality of bidders and targets.**

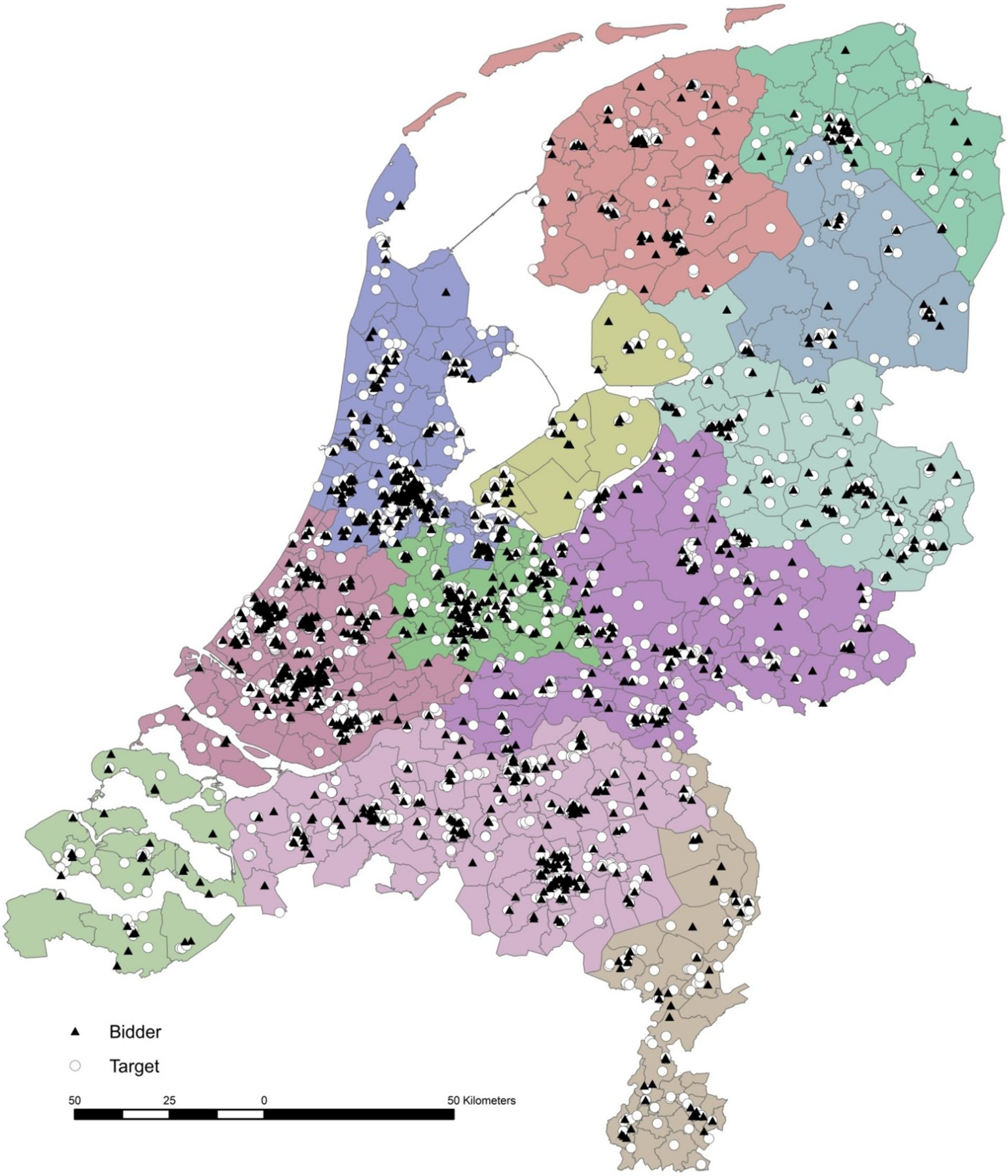


Frequency of undiversified (one activity) and diversified (two and more activities) bidders (dark) and targets (bright), N=1,855.



Frequency of single-locational (no subsidiaries) and multi-locational (1 and more subsidiaries) bidders (dark) and targets (bright), N=1,855.

Figure 13 The headquarter locations of bidders and targets.



Bidder and target headquarter locations. Colouring indicates different provinces and lines indicate municipality borders as in 2010. Coordinate system: RD (double stereographic), Datum: Amersfoort.

7 Estimation methods

7

Given the quantitative research strategy this chapter elaborates on the methods that are applied in order to test the hypotheses. It is shown how deals can be compared to non-deals in a bivariate setting (7.1), by means of logistic regression (7.2) and how the individual bias measures are constructed and tested (7.3).

The empirical tests are done based upon different assumptions (**Table 13**). As already explained in 5.1 non-deals are constructed under assumption ALL (every potential target can be selected) and under assumption SG (only potential targets that have the same NACE group as the real target can be selected). Tests under assumption SG control for localization economies (see 3.1). The effect of geographical proximity is therefore expected to be weaker when estimated under assumption SG. Furthermore, bivariate, univariate and multivariate analyses are applied. Whereas the bivariate tests do not control for any other effects the others do so. The univariate tests are based on the individual bias measures that control for the amount of available potential targets and their minimum and maximum value of geographical proximity and industrial relatedness. Therefore the univariate tests are regarded to be more valid than the bivariate tests. The multivariate tests (logistic regression) control for some other covariates (see 5.2.3) but not for the characteristics of every bidder's choice set.

Table 13 Overview of different methods applied to control for unwanted effects.

Controlling for localization economies	Controlling for other effects
No (Assumption ALL)	*No (Group comparison by bivariate analysis)*
No (Assumption ALL)	*Yes (Univariate analysis of individual bias measure)*
No (Assumption ALL)	*Yes (Multivariate analysis by logistic regression model)*
Yes (Assumption SG)	*No (Group comparison by bivariate analysis)*
Yes (Assumption SG)	*Yes (Univariate analysis of individual bias measure)*

The logistic regression model is only applied on the national scale. For the tests that only include deals and non-deals in which both companies are co-located in the same province, COROP region and municipality only the bivariate and univariate analyses are applied. The univariate analyses are always more valid than the simple bivariate techniques. The different methods are elaborated in detail in the following three sections.

7.1 Bivariate analysis

Hypotheses A1 and B1 can straightforwardly being tested by bivariate analyses techniques. The group of deals and the group of non-deal are compared to each other. If the independent variable is categorical (SPCM and USDG) simply contingency tables are drawn and relationship tested by Pearson's Chi Square Test. Consequently two different effect sizes can be calculated, which are odds ratios and relative risk. Odds ratios are calculated as

$$OR = \frac{D_{intra}}{D_{inter}} \Big/ \frac{ND_{intra}}{ND_{inter}}$$

with *D* as deals and *ND* as non-deals. *Inter* and *intra* refer to either whether the deal was intra- or inter-regional or whether the deal was intra- or inter-industrial. However, odds ratios are difficult to communicate to non-specialists. A more natural interpretable effect size is relative risk, sometimes called prevalence ratio. Relative risk is also used in the multivariate analyses, which allows a comparison of the effect sizes. For the bivariate analyses odds ratios are calculated as

$$RR = \frac{D_{intra}}{D} \Big/ \frac{ND_{intra}}{ND}.$$

If the independent variable is continuous (Distance and Tanimoto) other techniques are required. One could simply calculate the means or medians in each category and compare them. The significance of the difference, however, can only be estimated by statistical tests. For mean differences the parametric *t*-test needs to be exerted. However, the results of several graphical and numerical analysis methods (Park, 2008) suggested that the distribution of every sample is non-normal. Therefore the tests rely on non-parametric tests. The Independent-Samples Mann-Whitney U Test indicates whether two distributions are significantly different and the Independent-Samples Median Test indicates whether two medians significantly differ. Thus, I use those tests and consequently calculate the effect size and its 95% confidence interval by Hodges-Lehman Median Difference.

7.2 Multivariate analysis

The obvious disadvantage of bivariate analysis is that we cannot see how geographical proximity and industrial relatedness interact and we cannot include the control variables. The most popular statistical procedure for multivariate analysis and a binary dependent variable is logistic regression and seems therefore to be the natural choice. Thus, hypotheses A1 and B2 are tested by logistic regression as well. The logistic regression model predicts the probability that a deal is real over the probability of being a non-deal. The model assumes that

$$Prob(y = deal) = \frac{1}{1 + e^{-(\alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n)}}$$

with α as the intercept, β_n as the regression coefficient of the corresponding variable X_n. The final model used in this research is:

$$Deal = \alpha + \beta_1 P + \beta_2 C + \beta_3 M + \beta_4 LnDist + \beta_5 S + \beta_6 D + \beta_7 G + \beta_8 Tanimoto + \beta_9 UU + \beta_{10} MuMu + \beta_{11} SiSi + \beta_{12} UnUn + \beta_{12} DiDi + \varepsilon$$

with ε as the error term. According to Field (2009) logistic regression requires to meet five assumptions: (a) linearity of the logit, (b) independence of errors, (c) no multicollinearity, (d) no complete separation and (e) independence of observations. The *linearity assumption* requires that there is a linear relationship between the continuous predictors and the logit of the dependent variable. This assumption can be tested by estimating the interaction effects between the continuous variables and their logarithmic term. It appears that LnDistance*LnLnDistance is significant on a 5% level and Tanimoto*LnTanimoto is significant on a 1% level. Therefore this assumption is violated and the estimates of the continuous variables need to be interpreted with caution.

The assumption of *independence of errors* is the same as for ordinary regression and means that cases of the data should not be related. This is not the case because every dyad is unique. Although not an assumption, predictors should not be too highly correlated because *multicollinearity* leads to inflated standard errors. **Table 14** shows the correlation coefficients between all predictors. Distance and LnDistance are perfectly correlated. Therefore only one of these variables can be included into the model. Also G and Tanimoto are highly correlated because G almost exclusively covers an own range of very high Tanimoto coefficients (**Figure 13**). As the inclusion of Tanimoto does hardly affect the standard errors or odds ratios of the other predictors I always keep it in the model. The correlations between the regional dummies P, C, M and Distance are rather weak (**Figure 15**).

Another problem to avoid is *complete separation*. This is if the outcome can be perfectly predicted by one or more predictors. Complete separation occurs if there are many low and high values in a variable, which is not the case in our data. *Independence of observations* occurs if (a) the same dyad occurs several times, (b) the dyads share unexplained heterogeneity and (c) the dyads have monadic similarity (Poast, 2010). In fact, there is monadic similarity as 230 out of 1391 different bidding companies made a bid to more than one of the targets within the investigated time period (see 6.1). This leads to incorrect standard errors. **Appendix Table 9** shows the estimation of the final model with exclusively the 1,162 bidders that engaged in a deal with only one target. It shows clearly that the standard errors and confidence intervals are larger than in the model that includes all cases. This means that the confidence intervals of the model that includes all bidders are a bit overestimated.

Table 14 The correlations between the covariates used in the logistic regression model.

	(Distance-mean)²	LnDistance	P	C	M	Tanimoto	S	D	G	UU	MuMu	SiSi	UnUn	DiDi
DISTANCE	-0.30 ***	1.00 ***	-0.31 ***	-0.28 ***	-0.21 ***	-0.04 ***	-0.01 ***	0.00 ***	-0.03 ***	-0.02	-0.02	0.04 ***	0.03	0.00
(Distance-mean)²		-0.30 ***	0.05 ***	0.14 ***	0.15 ***	0.00	0.00 *	0.00 **	0.00	-0.04 **	0.01	-0.05 ***	-0.02	0.01
LnDistance			-0.31 ***	-0.40 ***	0.48 ***	0.00	-0.1 ***	0.00 ***	-0.04 ***	-0.01	-0.02	0.07 ***	0.05 ***	0.00
P				0.07 ***	0.06 ***	0.01 ***	0.00	0.01	0.03 ***	0.03 ***	0.02 *	0.08 ***	0.00	0.02 *
C					0.04 ***	0.01 ***	0.00	0.02 ***	0.07 ***	0.01	0.00	0.00	0.01	0.00
M						0.02 ****	0.04 ***	0.03 ***	0.13 ***	0.00	0.01	0.00	0.02 *	0.00
Tanimoto							0.09 ***	0.03 ***	0.77 ***	0.07 ***	0.01	-0.01	0.01	-0.22 ***
S								0.05 ***	0.10 ***	0.02	0.01	0.00	0.03 ***	0.04 ***
D									0.05 ***	0.02 **	0.01	0.01	0.02	0.01
G										0.00	0.00	0.02 **	0.05 ***	0.08 ***
UU											-0.04 **	0.20 ***	0.10 ***	-0.10 ***
MuMu												-0.22 ***	-0.05 ***	0.02
SiSi													0.07 ***	-0.04 **
UnUn														-0.38 ***

If both variables are continuous (Distance, (Distance-mean)^2, LnDistance, Tanimoto) Kendall's tau τ is used as most data is not normally distributed. If both variables are categorical (P, C, M, S, D, G, UU, MuMu, SiSi, UnUn, DiDi) Phi φ is used. If one variable is continuous (e.g. Distance) and one categorical (e.g. P) the point-biserial correlation coefficient r_{pb} is used, which has the same value as Pearson's product-moment correlation coefficient. Significance levels are two-tailed. Due to limited computational power the coefficients cannot be calculated on the full sample. Therefore an endogenously stratified sample is used with all deals and a random selection of non-deals, the same sample as used for the tests. N = 11,755.

The effect of each variable is given by its odds ratio; the *b* values reflect the rate of change in the log odds as the predictor changes and are therefore not very useful for interpretation. In contrast, the odds ratio indicates the change in odds after one unit of change in the predictor. The odds are the ratio of the likelihood belonging to the groups of deals and the likelihood of belonging to the group of non-deals:

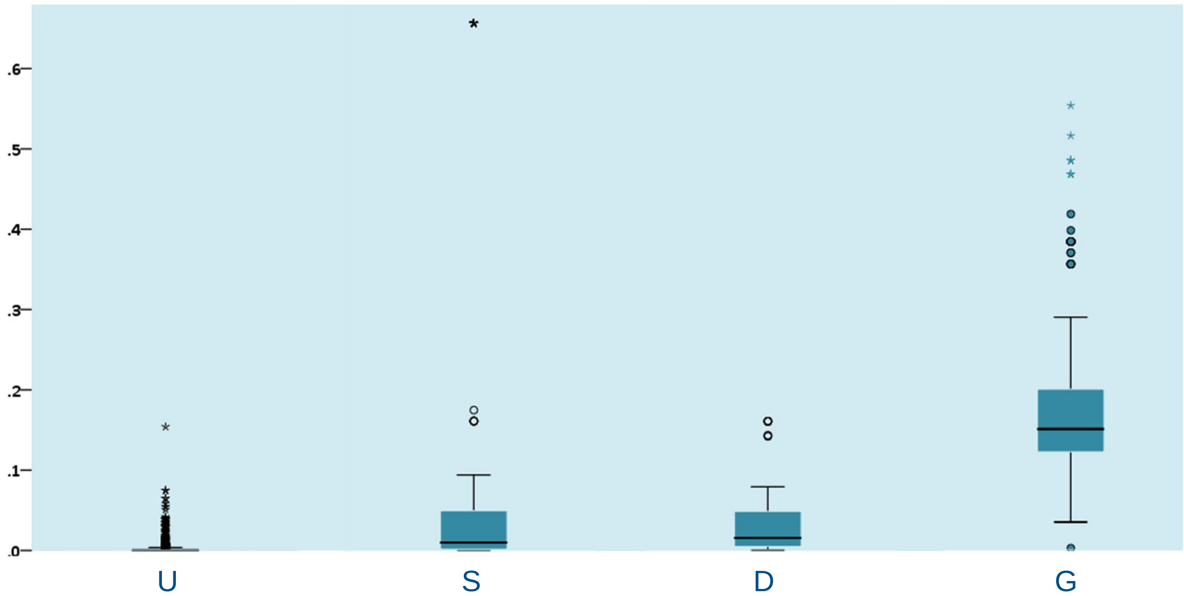
$$Odds = \frac{P(deal)}{P(non-deal)}.$$

Consequently, the odds ratio is

$$OR = \frac{odds\ after\ a\ unit\ change\ in\ the\ predictor}{original\ odds}.$$

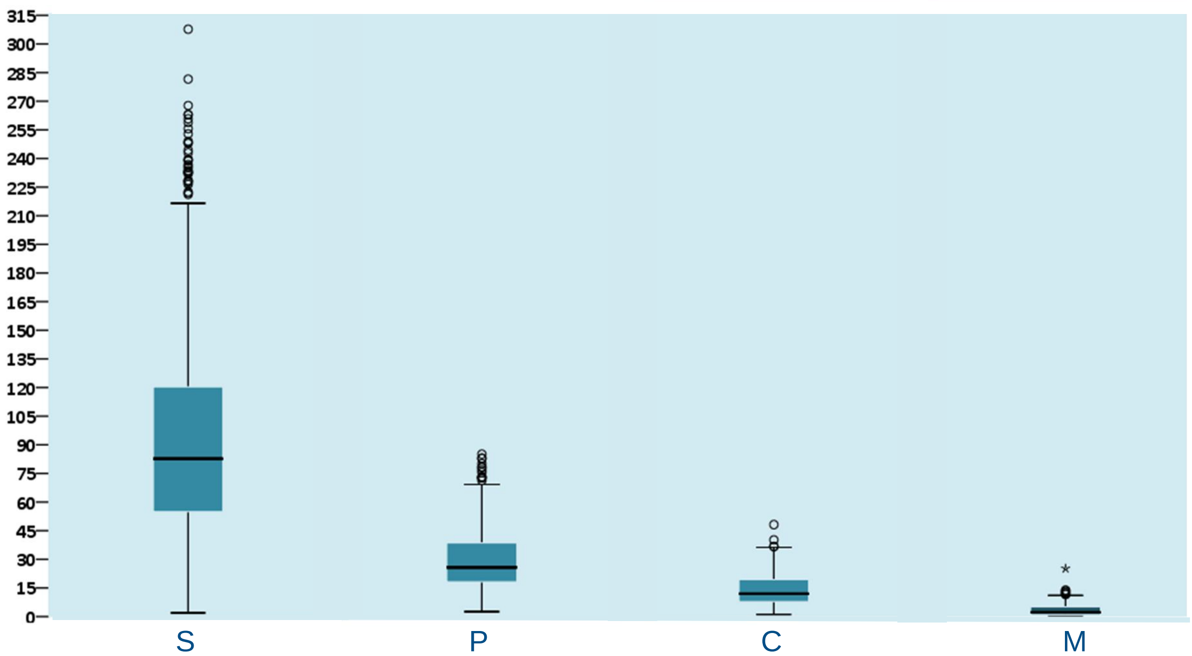
If the odds ratio is larger than 1 it means that the predictor has a positive effect on the outcome and if it is smaller than 1 it there is a negative effect on the outcome. When the upper or lower limit of the confidence interval passes 1 the predictor cannot be significant. Odds ratios are often misinterpreted as relative risk. Therefore it is useful to convert odds ratios into relative risk, which is possible with King and Zeng's (2001a) software ReLogit. However, in the models it appears that relative risk values differ extremely marginal from the odds ratios (compare **Appendix Table 8** and **Appendix Table 7).** Therefore I pass the manual correction of each odds ratio and just interpret the odds ratios as relative risk.

Figure 14 Tanimoto coefficients in inter- and intra-industry deals



Box-plot of U, S, D, G and Tanimoto. Due to limited computational power the figures reflect the data points of an endogenously stratified sample, N = 11,755.

Figure 15 Distances in inter- and intra-regional deals



Box-plot of S, P, C, M and Distance. Due to limited computational power the figures reflect the data points of an endogenously stratified sample, N = 11,755.

Next to the effect of each predictor we need to assess the absolute goodness of fit of the models, i.e. how accurately the outcome can be predicted by the predictors. This is a bit more difficult than it is for linear regression. I use two Chi-Square tests to assess the overall fit. The Omnibus Test indicates the overall significance of the model by comparing the χ^2 of the model without and the model with predictors. A significant value and high χ^2 indicates a good fit (UCLA: Academic Technology Services, 2011). In contrast, the Hosmer and Lemeshow Test needs to be unsignificant. If χ^2 is high and $p < 0.05$ it indicates a poor fit of the model (Bewick et al., 2005). If the test is significant it means that the proportion of observed deals is significantly different from those predicted by the model. In the models a significant difference is mostly observed together with overdispersion.

Overdispersion means that the discrepancies between observed variances are bigger than expected from the model and that means that there is a bad fit between model and data. Dispersion is measured by the dispersion parameter Φ, which is the ratio of the χ^2 statistic (Hosmer and Lemeshow test) to its degrees of freedoms. If $\Phi=1$ there is no over- or underdispersion. Φ should be between 1 and 2 because over- or underdispersion results in incorrect standard errors and therefore incorrect inferential statistics (Menard, 2002).

We cannot only test how well the model fits the data, we can also compare models among each other by means of pseudo R-squared measures. This is essential for assessing the contribution of each covariate. Note however, that these measures do not explain the proportion of variance as in OLS regression. Rather, they allow to compare models if the same data and same outcome variable is used. In general, R^2 in good logistic regression models are lower than in good OLS regression (Hosmer and Lemeshow, 2000). Therefore these values should not be underestimated. Hosmer and Lemeshow's R_L^2 is a very simple measure and compares the improvement from the null model (baseline), which consists only of the constant, to the fitted model. It is calculated as

$$R_L^2 = \frac{-2LL(model)}{-2LL(baseline)},$$

where -2LL is the model chi-square based on the log-likelihood (also called deviance). Log-likelihood is based on summing up the probabilities associated with the predicted and actual outcomes (Tabachnick and Fidell, 2007). R_L^2 is very high if there is only one predictor and in our case it is generally the higher the more observations are used. This is because with increasing sample size it becomes easier to predict 0's. Therefore, Cox and Snell's R_{CS}^2 is used, which takes the number of observations into account. It is calculated as

$$R_{CS}^2 = e^{\left[-\frac{2}{n}[LL(model)-LL(baseline)]\right]}.$$

As this measure never can achieve 1 it Nagelkerke adopted it by dividing it by the maximum possible value. This means, $R_N^2 = 1$ if the full model perfectly predicts the outcome. As this pseudo R-squared seems to be the best choice I always refer to R_N^2 in the discussions. It is calculated as

$$R_N^2 = \frac{1 - e^{\left[-\frac{2}{n}[LL(model)-LL(baseline)]\right]}}{1 - e^{\left[\frac{2(LL(baseline))}{n}\right]}}.$$

While the amount of 1's and 0's does not matter for the bivariate analyses, rare events (nearly two thousand 1's against more than one and a half million 0's) are a problem for logistic regression. In fact, logistic regression always underestimates the probability of rare events, which leads to errors in the same direction (King and Zeng, 2001a). Normally, probabilities in rare-event studies are very small for 1's and very large for 0's. Consequently biases in errors and coefficients can occur. I run the logistic regression including all predictors as in the final model and with all 1.6 million observations (**Appendix**

Table 7). The estimation of the full model shows that about all predictors are significant, which is not a surprise regarding the size of the sample. Furthermore, the goodness of overall fit is very bad and overdispersion is strong in the model estimated with the full sample. In order to cope with these inaccuracies an alternative is needed.

The alternative is to use a subsample, which is created by endogenous stratification, also called choice-based sampling in econometrics (e.g. Manski and Lerman, 1977) or case-control design in epidemiology (e.g. Breslow and Day, 1980). This simply means to split observations into a set of deals and a set of non-deals and then randomly, i.e. independently from all other variables, select from the 0's and select all available 1's. That is, the sample is stratified on an endogenous variable without losing consistency or efficiency compared to the full sample (King and Zeng, 2001b, 2001a). While it is clear to use all 1,855 deals the only question remains is how many non-deals should we draw from the tombola containing 1,607,008 non-deals? Following King and Zeng (2001a) an equal share (that is $\bar{y}$=0.5) is in many situations optimal or close to optimal. When the number of 0's exceeds the number of 1's each additional 0 contributes less to the explanatory variables' information content. Therefore one would normally collect not more than two to five times more 0's than 1's. It appears that Chakrabarti and Mitchell (2008) has randomly selected five potential targets for each deal (that is $\bar{y}$=0.2). In our case we can select as many 0's as we want as the data is available. I follow King and Zeng's (2001a) suggestion to proceed sequentially. I construct several samples ranging from $\bar{y}$=0.5 to $\bar{y}$=0.1, which appear to have different characteristics, depending on the number of 0's, although the trends are due to the random selection not linear (see **Table 15).**

Purpose of these estimations is to identify a subsample which fits the data well, i.e. which does not suffer from under- or overdispersion. Additionally, we have to watch out for large standard errors, an acknowledge for estimation uncertainties, and broad confidence intervals of the odds ratios. **Appendix Table 7** shows the estimations of the 50-50 split sample and the sample with 9,900 0's compared to the estimations with the full sample. It seems that the sample with N=11,755 is the best choice, as it shows a good fit, not too large standard errors and not too broad confidence intervals. One must, however, also state that, as the confidence intervals reveal, some odds ratios are significantly different from the ones estimated with the full sample. The effects of LnDistance, C, M and G are significantly smaller in this sample.

Table 15 Several endogenously stratified samples and their suitability for logistic regression.

1's	0's	$\bar{y}$	Dispersion Φ	R_N^2
1,855	1,855	0.500	0.40	0.564
1,855	8,000	0.232	1.16	0.513
1,855	8,500	0.218	1.13	0.513
1,855	9,000	0.206	1.30	0.512
1,855	9,500	0.195	1.13	0.509
1,855	9,800	0.189	0.83	0.502
1,855	9,900	0.187	1.05	0.514
1,855	9,950	0.186	0.69	0.502
1,855	10,000	0.186	0.93	0.509
1,855	11,000	0.169	0.75	0.490
1,855	14,000	0.133	0.87	0.474
1,855	1,607,008	0.001	4.16	0.231

When estimating a sample the values of deal likelihood need to be correct, which are much higher when based on the sample. An ex-post correction is virtually never done by applied researchers although the models are only valid when the finite sample is corrected (King and Zeng, 2001a). After the correction with ReLogit we receive, however, an average probability which is not significantly different from the average probability suggested by the full model estimation (**Table 16**). Therefore we can just use the likelihood values estimated on the basis of the full sample.

Table 16 Uncorrected and corrected deal likelihoods.

Data	Deal likelihood	95% Confidence interval Abs HB (SG)	
Full sample	0.00030	0.00027	0.00034
Sample (uncorrected)	0.03983	0.03462	0.04541
Sample (corrected)	0.00033	0.00027	0.00042

Deal likelihood when the values of all predictors are kept on the median value.

7.3 Univariate analysis

All hypotheses and research questions are tested by individual bias measures that indicate the home bias (HB_b) or industrial relatedness bias (IRB_b) of every single bidder b. The tests of *hypotheses A1* and *B1* are exerted by examining the distribution of these measures. The proportion of positively biased bidders (*hypotheses A2 and B2*) are tested by estimating the confidence interval of the bias measure means by bootstrapping. As the measure has positive and negative values the hypothesis is confirmed if the confidence interval is positive. It can also be tested by the non-parametric sign test, which leads to the same results. The following paragraphs show how the individual bias measures are constructed and explain some methodological issues.

First, I calculate the absolute home bias ($Abs\ HB_b$) of every bidder in km. This measure is the base of a relative measure with the range [-100…100]. The absolute home bias expresses the absolute distance which the bidder's real target is closer to the bidder than the average target. Basically, the distance between bidder *b* and its real target *rt* and the mean of the distance between *b* and all its potential targets *pt* is compared. The formula is

$$Abs\ HB_b = \frac{\sum_{b,ht=1}^{n} Distance_{b,pt}}{n} - \text{distance}_{b,rt}.$$

This measure does not take into account the different ranges of possible distances for every bidder yet. Therefore it is corrected by the minimum and maximum possible distance to every potential target. The relative home bias needs to be differently calculated for home biased and distance biased bidders. For home biased bidders we have to divide $Abs\ HB_b$ by the range between the distance to the closest target and the mean distance, and for distance biased bidders we have to divide $Abs\ HB_b$ by the range between the mean distance and the distance to the target which is furthest away. The values are multiplied with 100 and rounded. The formulas for the relative home bias $Rel\ HB_b$ are

$$Rel\ HB_b = \frac{Abs\ HB_b}{distance_{b,pt_{max}}} * 100 = \frac{Abs\ HB_b}{\frac{\sum_{b,ht=1}^{n} Distance_{b,pt}}{n} - distance_{b,pt_{min}}} * 100\ if\ Abs\ HB_b \geq 0,$$

$$Rel\ HB_b = \frac{Abs\ HB_b}{distance_{b,pt_{max}}} * 100 = \frac{Abs\ HB_b}{distance_{b,ht_{max}} - \frac{\sum_{b,ht=1}^{n} distance_{b,pt}}{n}} * 100\ if\ Abs\ HB_b < 0.$$

The above explained measure indicates the home bias for every bidder by comparing the distance to the actual target and the average distance to all potential targets. The number of potential targets is different for every bidder and sometimes very small, especially if assumption SG is applied (see **Table 17**). This might bias the results. Therefore I define a threshold of a minimum number of potential targets and test whether the number of potential targets has an influence on the home bias. I decide that the home bias can only be estimated for bidder that have at least 5 potential target under assumption ALL and at least 3 potential targets und assumption SG; the number of potential target is excluding the real target. If the threshold is higher too many deals would be lost, especially within municipalities and under assumption SG, which in turn would lead to unrepresentative results.

Having decided on a minimum threshold the number of potential targets is still different for every bidder. One can imagine that the differences of these numbers might have an impact on the home bias. I test this impact by means of a correlation analysis (see **Table 18**). I used the non-parametric Kendall's tau as correlation coefficient because the data is not normally distributed. Although there were some significant relations the effect of the number of non-deals on absolute or relative home bias is very weak. This means the home bias is hardly biased by the number of potential targets.

Table 17 The distribution of the number of potential targets per bidder.

Location of PT	Assumption ALL					Assumption SG				
	Min	Max	Mean	SD	N if PT>5	Min	Max	Mean	SD	N if PT>3
Netherlands	385	1,067	867.3	177.9	1,855	1	98	18.9	22.6	1,571
Same province	6	206	114.1	54.8	654	0	22	3.0	4.2	467
Same COROP	1	133	47.1	37.2	482	0	22	1.5	2.8	183
Same municip.	0	88	13.8	20.4	166	0	11	0.6	1.4	64

Table 18 The effect of the number of potential targets on home bias after thresholding.

Location of PT	Abs HB (ALL)	Rel HB (ALL)	Abs HB (SG)	Rel HB (SG)
Netherlands	0.016	0.026*	-0.022	-0.088***
Same province	0.069***	0.040	-0.068	-0.101**
Same COROP region	-0.079**	-0.098***	0.058	0.010
Same municipality	0.027	0.036	0.185	0.060

Kendall's tau τ correlation coefficients. Significance is 2-tailed.

Following the logic of the construction of the home bias measure I construct an absolute and relative industrial relatedness bias measure IRB_b (Tanimoto Bias). Here, oppositely to home bias, there is a bias if the value of Tanimoto is large in the deals and low in non-deals. Therefore, the sign of the bias measure needs to be turned:

$$Abs\ IRB_b = \left(\frac{\sum_{b,ht=1}^{n} Tanimoto_{b,pt}}{n} - Tanimoto_{b,rt}\right) * -1.$$

Also here a relative measure is used, which formulas are as follows:

$$Rel\ IRB_b = \frac{Abs\ IRB_b}{Tanimoto_{b,pt_{max}}} = \frac{Abs\ IRB_b}{\frac{\sum_{b,pt=1}^{n} Tanimoto_{b,pt}}{n} - Tanimoto_{b,pt_{min}}}\ if\ Abs\ IRB_b \geq 0,$$

$$Rel\ IRB_b = \frac{Abs\ IRB_b}{Tanimoto_{b,pt_{max}}} = \frac{Abs\ IRB_b}{Tanimoto_{b,pt_{max}} - \frac{\sum_{b,ht=1}^{n} Tanimoto_{b,pt}}{n}}\ if\ Abs\ IRB_b < 0.$$

I use the same threshold for the number of potential targets as before and I test by means of correlation analysis whether the different numbers of potential targets have an effect on relatedness. The non-parametric Kendall's tau reveals that there are no significant relations (see **Table 19**).

Table 19 The effect of the number of potential targets on ind. related. bias after thresholding.

Location of potential target	Abs. IRB	Rel. IRB
Netherlands	0.012	0.026
Same province	0.002	-0.062**
Same COROP region	-0.078**	-0.096***
Same municipality	-0.057	-0.177***

Kendall's tau τ correlation coefficients. Significance is 2-tailed.

8 Empirical results

This chapter presents the outcomes of the empirical tests for all spatial scales, with and without controlling for other effects. 8.1 presents evidence on the bidder's home bias, 8.2 on the bidder's industrial relatedness bias. 8.3 assesses the M&A formation model.

8

8.1 Home bias

8.1.1 Average home bias

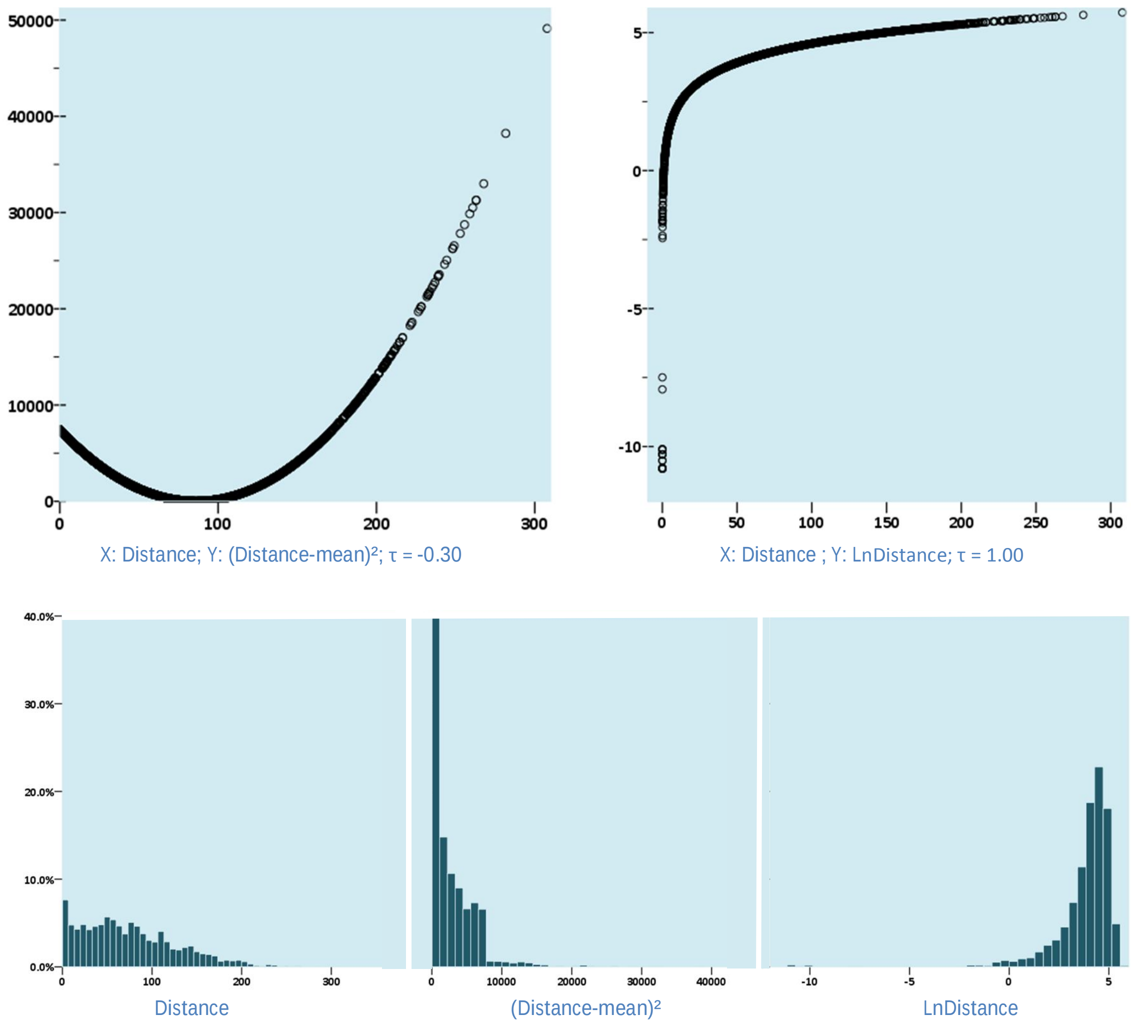
In order to test *hypothesis A1* the effect of SPCM on whether two particular companies announced a deal or not was tested. Unsurprisingly, Pearson's Chi-Square suggested a highly significant relation between SPCM and M&A formation. The bivariate estimates of relative risk, controlled for localization economies, show that deal chance was 4.6 times higher if both firms are located within the same municipality, 2.4 times higher within the same COROP region (but not the same municipality) and 1.5 times higher within the same province (but not the same COROP region). Without controlling for localization economies the numbers are higher. The contingency table and the estimations of χ^2, odds ratios and relative risks can be found for both assumptions in **Appendix Table 10**. Using logistic regression it was not only controlled for all other predictors, also confidence intervals could be estimated. **Table 20** shows also significant, but less optimistic, estimates. Here, deal likelihood was only 1.0 to 2.7 times higher for firms within the same municipality, 1.1 to 2 times higher for firms within the same COROP region (but not the same municipality) and 1.0 to 1.7 times higher for firms within the same province (but not the same COROP region). Thus, the strongest effect was found for intra-municipality dyads (M).

Hypotheses A1 can also be tested by estimating the differences in Distance between deals and non-deals. It was possible to estimate the mean differences (see **Appendix Table 11** and **Appendix Table 12**) but as the distributions were not normal their significance could not be tested. However, Independent-Samples Mann Whitney U Test suggested significantly different distributions and the Independent-Samples Median Test confirmed that also the medians were significantly different. The striking finding here is that this was not only true when all domestic deals were regarded, but also when deals within the same province, COROP region and even municipality were tested (*research question A3*). The average home bias was, measured by Independent-Samples Hodges Lehman Median Difference taking a 95% confidence level, 12.7 km to 26.9 km for bidders that selected any target within the Netherlands, 4.7 km to 6.8 km for bidders that selected a target from the same province, 2.0 km to 3.3 km for bidders that selected a target from the same COROP region and still 600 m to 1400 m for bidders that selected a target from the same municipality (see **Appendix Table 13**). The estimates were larger if the bidder could choose between targets of any industry (see **Appendix Table 14**).

On a national level, the effect of Distance on deal likelihood was also estimated in the logistic regression model, which controls for industrial relatedness and the control variables. Model A in **Appendix Table 15** reveals that Distance had a significant and negative effect on deal announcement. However, the overall fit of the model is worse (R_N^2=0.490) than of model B, in which only the regional dummies were included (R_N^2=0.503). No matter whether P, C, M or Distance were the only measure of geographical proximity, both models suffered from extreme overdispersion. Striking is that Distance is even significant in conjunction with the regional dummy variables (Model C). Additionally to the effect of regional co-location the chance of deal announcement decreased on average with 3% to 7% per additional kilometer. In terms of dispersion model C was superior (R_N^2=0.507) to the two previous ones, although in this model underdispersion instead of overdispersion was found.

Models D, E and F in **Appendix Table 16** address the problem of over- and underdispersion by using a quadratic and logarithmic term of Distance. The transformation of the variable can be seen in **Figure 16**. For the quadratic term first the distance mean was subtracted in order to avoid perfect correlation. For the logarithmic term I assigned -12 to 96 cases in which distance is 0 (same postcode, but different addresses). When including the quadratic term (Distance-mean)^2 the models improved from R_N^2=0.490 to R_N^2=0.502 but remains, although less, still overdispersed (model D). Overdispersion decreased when replacing Distance by the logarithmic term (they could not be estimated conjointly as τ = 1.00) and the model improved even more (model E). Adding the regional dummies overdispersion disappeard and R_N^2 increased to 0.514 (model F). Thus, LnDistance outperformed Distance and P, C, M even had an effect when estimated in conjunction with S, D, G. The logarithmic term predicts that deals are 1.4 times more likely when reducing the distance from 200km to 100km and they are also 1.4 times more likely when reducing distance from 2km to 1km. In fact, P, C, M and LnDistance did not measure the same effect because the joint correlation between SPCM and LnDistance is only r_{pb}=-0.47. **Table 20** shows all estimates of the final model. The final model shows that geographical proximity is significant in conjunction with S, D and G. This means that geographical proximity and industrial relatedness are not serving as substitutes (*research question B3*).

Figure 16 **The transformation of Distance for the logistic regression model.**



Figures of the transformation function and the frequency of the transformed values.

Table 20 Estimations of the final logistic regression model corrected with ReLogit.

Independent variables	Sample (N=11,755) Corrected with ReLogit				
	B	Robust SE	Relative Risk (95% CI)		
lnDistance	-0.50***	0.06	0.61	0.54	0.68
P	0.26**	0.12	1.31	1.03	1.66
C	0.38**	0.16	1.45	1.07	1.98
M	0.51**	0.25	1.67	1.02	2.68
Tanimoto	-0.49	0.72	1.00	1.00	1.00
S	1.66***	0.09	5.22	4.37	6.17
D	2.50***	0.14	12.10	9.10	15.97
G	4.54***	0.16	89.40	65.62	117.48
UU	-0.06	0.13	0.94	0.71	1.23
SiSi	-0.04	0.07	0.95	0.83	1.11
MuMu	-0.34***	0.12	0.71	0.56	0.89
UnUn	0.03	0.08	1.03	0.88	1.21
DiDi	-0.53***	0.13	0.59	0.44	0.76
Intercept	-5.21***	0.32			
Deals	1,855				
Non-deals	9,900				
Omnibus test, χ² (df)	4173.38*** (13)				
Hosmer and Lemeshow test, χ² (df)	8.38 (8)				
Dispersion Φ	1.05				
R^2_L	0.597				
R^2_{CS}	0.299				
R^2_N	0.514				

For the estimations of relative risks LnDistance and Tanimoto are hold at the mean and all others, i.e. all binary variables, at 0. The relative risk indicates deal likelihood if the independent variable is increased by 1.

8.1.2 Individual home bias

Using the measure of individual home biases, *hypothesis A1* but also *hypothesis A2*, stating that the fraction of home biased bidders is higher than the fraction of distance biased bidders, could be tested. Considering bidders that undertook deals within the Netherlands or the same province the results presented in **Table 21** are similar to the estimates of the Hodges-Lehman Median Difference: 17.9 to 22.4km and 4.1 to 7.2km, respectively. The estimate for inter-COROP region deals was a bit smaller (740m to 2.8km) and on a municipality scale there was no significant home bias any more (-400m to 1.2km), because the lower bound of the confidence interval was negative. This was not the case if bidders could select any target (see **Appendix Table 17**).

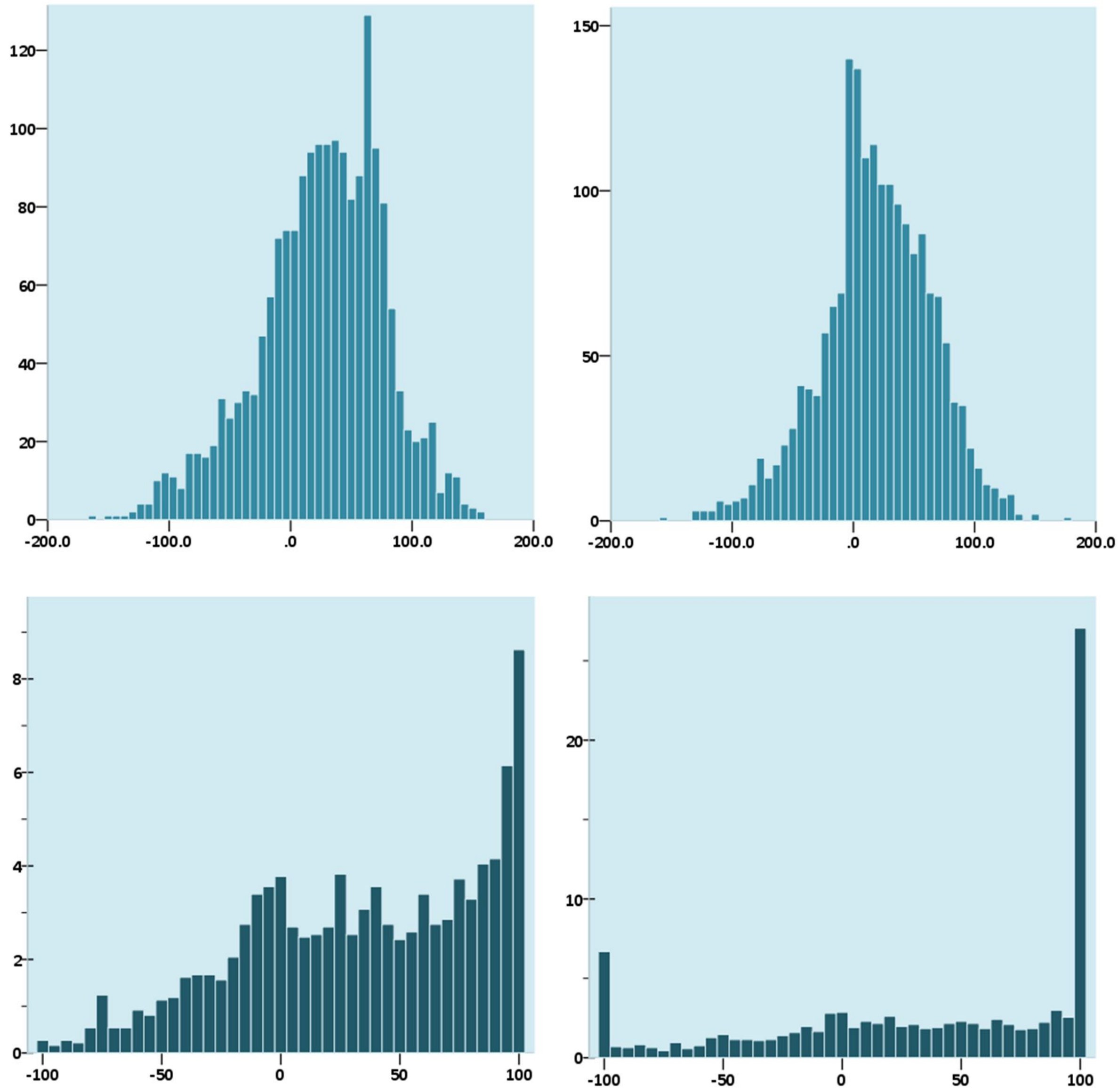
Table 21 The absolute home bias for deals on different regional scales (SG).

Absolute home bias (%)	Domestic			Same province			Same COROP			Same municipality		
Min			-153.6			-47.9			-16.4			-5.6
Max			175.3			51.0			17.4			3.9
Mean (95% CI)	20.28	17.87	22.39	5.67	4.12	7.23	1.87	0.74	2.84	0.40	-0.41	1.16
SD	47.21	45.48	48.95	13.50	11.93	14.95	5.97	5.02	6.77	2.54	1.92	3.02
Skewness	-0.33	-0.44	-0.20	-0.26	-0.71	0.31	0.36	-0.18	0.93	-0.70	-1.20	-0.15
Kurtosis	0.24	-0.01	0.47	2.01	0.92	2.92	0.87	-0.07	1.95	-0.25	-1.20	1.31
N			1,664			307			116			38

Only cases with at least 3 potential targets are included. Bootstrapping is based on 1,000 samples.

The distributions of the absolute relative home bias values are given in **Figure 17**. There is a striking difference when controlling for localization economies. Under this assumption many bidders selected the closest target available. 26.5% of all bidders had an extreme home bias of 99% or 100%, but a significant fraction also opted for the most distant target. If a bidder could select between all industries those distance biased bidders were missing. Also here I used subsamples for intra-regional deals and the distributions all looked fairly similar (see **Appendix Figure 3**).

Figure 17 **The absolute and relative home bias of bidders.**



Histograms (#) of absolute domestic home bias in km (top) and histograms (%) of relative domestic home bias (bottom) under assumption ALL (left) (N=1,855) and under assumption SG (right) (N=1,571).

Table 22 reveals that the home bias was strongest for deals within the same province. It shows furthermore the test of *hypothesis A2*, which was tested by the sign test and by the estimation of confidence intervals via bootstrapping. Both methods lead to the same results. While the proportion of home biased bidders is always significantly larger than 50% it is not when controlling for localization effects and considering bidders only that announced a deal with targets of the same municipality. **Table 23** and **Table 24** show all results concerning the role of geographical proximity.

Table 22 The relative home bias for deals on different regional scales (SG).

Relative home bias (%)	Domestic			Same province			Same COROP			Same municipality		
Mean (95% CI)	32.1	28.8	35.1	39.0	31.4	46.8	23.8	10.2	37.9	17.9	-9.6	45.5
SD	66.0	64.2	67.7	70.3	65.6	74.3	77.8	71.1	82.4	85.5	73.5	91.8
Skewness	-0.64	-0.71	-0.56	-0.84	-1.06	-0.65	-0.46	-0.80	-0.17	-0.33	-0.98	0.29
Kurtosis	-0.82	-0.95	-0.69	-0.70	-1.05	-0.18	-1.38	-1.62	-0.86	-1.69	-1.91	-0.68
Proportion of bidders with positive home bias in % (95% CI)	69	67	71	71	66	76	60	52	69	55	39	71
N	1,664			307			116			38		

Only cases with at least 3 potential targets are included. Bootstrapping is based on 1,000 samples.

Table 23 Evidence on Domestic home bias. Summary of findings.

Spatial scale	Contr. for localization effects	Controlled for other effects	Effect size	Variable	Estimate (95% CI)		
State	No (ALL)	No (Group comparison by bivariate analysis)	Relative risk	P	1.6***		
				C	2.8***		
				M	7.9***		
			Median difference	Distance	27.9***	25.5	30.3
		Yes (Univariate analysis of individual bias measure)	Mean	Abs. home	25.5***	23.0	27.9
				Rel. home bias	33.7***	31.4	36.2
		Yes (Multivariate analysis by logistic regression model)	Relative risk	P	2.2***	1.8	2.7
				C	3.7***	2.9	4.7
				M	10.4**	8.0	13.7
				LnDistance	0.6***	0.5	0.7
		Yes (Univariate analysis of individual bias measure)	Percentage of home biased bidders	Home bias	71***	70	74
	Yes (SG)	No (Group comparison by bivariate analysis)	Relative risk	P	1.5***		
				C	2.4***		
				M	4.6***		
			Median difference	Distance	19.9***	12.7	26.9
		Yes (Univariate analysis of individual bias measure)	Mean	Abs. home	20.3***	17.9	22.4
				Rel. home bias	32.1***	28.8	35.1
		Yes (Univariate analysis of individual bias measure)	Percentage of home biased bidders	Home bias	69***	67	71

Table 24 Evidence on Regional home bias. Summary of findings.

Effect size and variable	Regional scale	Assumption ALL			Assumption SG		
		Estimate	95% CI		Estimate	95% CI	
Median difference of Distance	Netherlands	27.9***	25.5	30.3	19.9***	12.7	26.9
	Intra-province	11.2***	8.6	13.9	5.7***	4.7	6.8
	Intra-COROP region	4.4***	3.5	5.3	2.7***	2.0	3.3
	Intra-municipality	1.7***	1.4	2.1	0.9***	0.6	1.4
Mean of Absolute home bias	Netherlands	25.5***	23.0	27.9	20.3***	17.9	22.4
	Intra-province	10.3***	8.9	11.5	5.7***	4.1	7.2
	Intra-COROP region	5.2***	4.5	6.0	1.9***	0.7	2.8
	Intra-municipality	1.0***	0.6	1.4	0.4	-0.4	1.2
Mean of Relative home bias	Netherlands	33.7***	31.4	36.2	32.1***	28.8	35.1
	Intra-province	44.8***	40.7	48.9	39.0***	31.4	46.8
	Intra-COROP region	43.7***	38.7	49.1	23.8***	10.2	37.9
	Intra-municipality	38.7***	29.2	47.3	17.9	-9.6	45.5
Percentage of home biased bidders	Netherlands	71***	70	74	69***	67	71
	Intra-province	78***	72	82	71***	66	76
	Intra-COROP region	77***	73	81	60**	51	69
	Intra-municipality	72***	65	79	55	40	71

8.2 Industrial relatedness bias

8.2.1 Average industrial relatedness bias

The existence of industrial relatedness bias (*hypothesis B1*) was tested in a very similar way as the existence of home bias. **Appendix Table 19** reveals that, other things being equal, the deal chance was 26.1 times higher if both firms were within the same NACE group, 4.1 times higher if they were within the same NACE division (but not the same NACE group) and 1.7 times higher if they were within the same NACE section (but not the same NACE division). While the effect sizes of P, C, M decreased in the logistic regression model the effect sizes of S, D, G increased when controlling for geographical proximity and the other control variables. Deal likelihood was 65.6 to 117.5 times higher when both firms were within the same NACE group, 9.1 to 16 times if they were within the same NACE division and 4.4 to 6.2 times if they were within the same NACE section (see **Appendix Table 16**).

Almost needless to say that also the means and medians of Tanimoto differed largely between the group of deals and the group of non-deals. While the median for non-deals was 0.0002 it was 0.0969 for deals, which is 485 times larger. Also here it was tested whether these results differ when only considering regional deals. Independent-Samples Hodges-Lehman Difference denoted a difference between 0.076 and 0.097, while there were no significant differences between the different spatial subsamples (**Appendix Table 21**). When estimating the effect of Tanimoto in logistic regression there was a similar problem as when estimating the effect of Distance. The model suffered from overdispersion and R_N^2 was low, here 0.391 (see **Appendix Table 20**). If S, D, G was used as the only measure of industrial relatedness R_N^2 became 0.513. S, D, G estimated in conjunction with Tanimoto turned to be insignificant. This corresponds with the biserial correlation r_{pb} between USDG and Tanimoto of 0.66. That means S, D, G counted for 43.6% of the variability in Tanimoto (see **Figure 14**).

Table 25 answers the first part of *research question B4*. While R_N^2=0.129 if only geographical proximity is taken as predictor R_N^2=0.465 if only industrial relatedness is taken as a predictor. That means that the impact of industrial relatedness is much stronger than the impact of geographical proximity. The final model (**Table 20**) shows that also two control variables are highly significant. These are MuMu and DiDi. Thus, M&A formation is less likely between two multi-locational and two diversified companies.

Table 25 The effect of industrial relatedness vs. the effect of geographical proximity.

Independent variables	Only geographical proximity					Only industrial relatedness				
	B	SE	Odds Ratio (95% CI)			B	SE	Odds Ratio (95% CI)		
P	0.21**	0.09	1.23	1.02	1.49					
C	0.38***	0.12	1.46	1.14	1.86					
M	0.53***	0.19	1.71	1.18	2.46					
lnDistance	-0.51***	0.04	0.60	0.55	0.65					
S						1.66***	0.08	5.23	4.46	6.21
D						2.56***	0.14	12.87	9.84	16.81
G						4.46***	0.09	86.73	72.24	104.11
UU	0.03	0.09	1.03	0.86	1.24	-0.00	0.12	1.00	0.80	1.26
SiSi	0.12**	0.06	1.13	1.00	1.27	0.06	0.07	1.06	0.92	1.22
MuMu	0.27***	0.09	1.31	1.10	1.57	0.29**	0.11	1.34	1.07	1.68
UnUn	0.30***	0.06	1.35	1.20	1.52	0.00	0.07	1.00	0.87	1.16
DiDi	1.00***	0.09	2.72	2.27	3.26	0.51***	0.12	1.66	1.31	2.09
Intercept	-0.11	0.21	0.90			-2.91	0.12	0.06		
Deals	1,855					1,855				
Non-deals	9,900					9,900				
Omnibus test, χ² (df)	919.51*** (9)					3711.05*** (8)				
Hosmer and Lemeshow test, χ² (df)	11.58 (8)					5.52 (6)				
Dispersion Φ	1.44					0.92				
R_L^2	0.913					0.640				
R_{CS}^2	0.075					0.271				
R_N^2	0.129					0.465				

8.2.2 Individual industrial relatedness bias

When estimating the industrial relatedness bias on the basis of the individual bias measure it was a bit larger than the estimates of the Hodges-Lehman Median Difference. The bias ranged from 0.088 to 0.112 and was also fairly similar on different spatial scales (**Table 27**). There is small trend, however, which indicates that industrial relatedness bias is larger on smaller regional scales. The positive values of skewness indicate a pile-up on the left of the distribution. The positive values of the kurtosis indicate a pointy and heavy-tailed distribution (see also **Figure 18**). The relative measure gives additional insight (**Table 27**). Between 70% and 84% of all bidders selected a related target, as tested by the sign test and the estimation of confidence intervals (*hypothesis B2*).

Table 26 The absolute industrial relatedness bias for deals on different regional scales.

Absolute IR bias (Tanimoto)	Domestic			Same province			Same COROP			Same municipality		
Min	-0.065			-0.062			-0.036			-0.031		
Max	0.614			0.615			0.608			0.558		
Mean (95% CI)	0.099	0.094	0.105	0.094	0.086	0.102	0.093	0.084	0.103	0.088	0.074	0.105
SD	0.112	0.106	0.117	0.102	0.093	0.112	0.101	0.090	0.114	0.100	0.081	0.120
Skewness	1.40	1.22	1.55	1.41	0.96	1.74	1.46	0.93	1.84	1.67	0.64	2.09
Kurtosis	2.77	1.92	3.50	3.47	1.42	5.19	3.92	1.23	5.79	4.80	0.05	6.92
N	1,855			654			428			166		

Only cases with at least 5 potential targets are included. Bootstrapping is based on 1,000 samples.

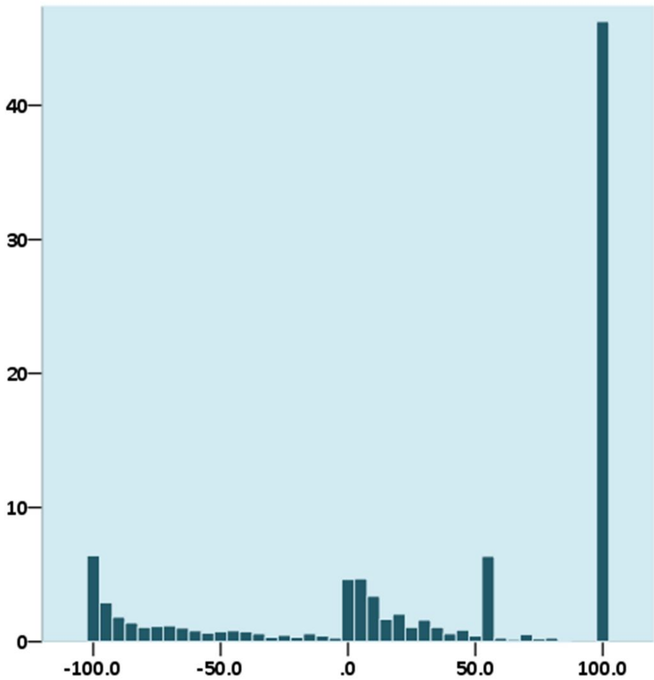
Table 27 The relative industrial relatedness bias for deals on different regional scales.

Relative IR bias (%)	Domestic			Same province			Same COROP			Same municipality		
Mean (95% CI)	36.4	32.9	39.8	44.2	38.0	50.2	46.7	40.3	54.0	47.4	35.8	58.6
SD	71.9	70.3	73.7	73.1	69.6	76.3	75.5	70.9	79.0	74.2	66.4	80.1
Skewness	-0.71	-0.79	-0.62	-0.93	-1.10	-0.77	-1.02	-1.25	-0.84	-1.03	-1.41	-0.72
Kurtosis	-0.93	-1.08	-0.78	-0.69	-0.99	-0.29	-0.63	-1.00	-0.07	-0.56	-1.21	0.49
No. of bidders with positive home bias in % (95% CI)	76	74	78	77	74	80	76	72	80	77	70	84
N	1,855			654			428			166		

Only cases with at least 5 potential targets are included. Bootstrapping is based on 1,000 samples.

Figure 18 shows that more than 46% of the bidders selected the most related target available. There were hardly any bidders with an industrial relatedness bias of 57% to 99%, in other words, strongly related targets were barely acquired. About 5% of the bidders had a bias of 56%. These were almost exclusively deals between firms of NACE group **620: Computer programming, consultancy and related activities**. As this industry occurs more often conjointly with **582: Software publishing** those intra-industry deals were only moderately related deals. All in all, deals were distinctly positively biased. A summary of all findings is given in **Table 28** and **Table 29**.

Figure 18 **Industrial relatedness bias for all bidders.**



Histogram (%) of domestic relatedness bias, N=1,855.

Table 28 Evidence on Domestic industrial relatedness bias. Summary of findings.

Spatial scale	Controlled for other effects	Effect size	Variable	Estimate (95% CI)		
State	No (Group comparison by bivariate analysis)	Relative risk	S	1.7***		
			D	4.1***		
			G	26.1***		
		Median difference	Tanimoto	0.080***	0.094	0.076
	Yes (Univariate analysis of individual bias measure)	Mean	Abs. IR bias	0.099***	0.095	0.093
			Rel. IR bias	36.4***	32.9	39.8
	Yes (Multivariate analysis by logistic regression model)	Relative risk	S	5.2***	4.4	6.2
			D	12.1***	9.1	16.0
			G	89.4***	65.6	117.5
			Tanimoto	0.6	0.5	0.7
	Yes (Univariate analysis of individual bias measure)	Percentage of industrial relatedness biased bidders	Home bias	76***	74	78

Table 29 Evidence on Regional industrial relatedness bias. Summary of findings.

Effect size and variable	Regional scale	Estimate (95% CI)		
Median difference of Tanimoto	Netherlands	0.080***	0.094	0.076
	Intra-province	0.087***	0.097	0.076
	Intra-COROP region	0.082***	0.095	0.075
	Intra-municipality	0.082***	0.096	0.076
Mean of Absolute IR bias	Netherlands	0.099***	0.095	0.093
	Intra-province	0.095***	0.087	0.104
	Intra-COROP region	0.093***	0.084	0.103
	Intra-municipality	0.080***	0.073	0.104
Mean of Relative IR bias	Netherlands	36.4***	32.9	39.8
	Intra-province	44.2***	38.0	50.2
	Intra-COROP region	46.7***	40.3	54.0
	Intra-municipality	47.4***	35.8	58.6
Percentage of IR biased bidders	Netherlands	76***	74	78
	Intra-province	77***	74	80
	Intra-COROP region	76***	72	80
	Intra-municipality	77***	70	84

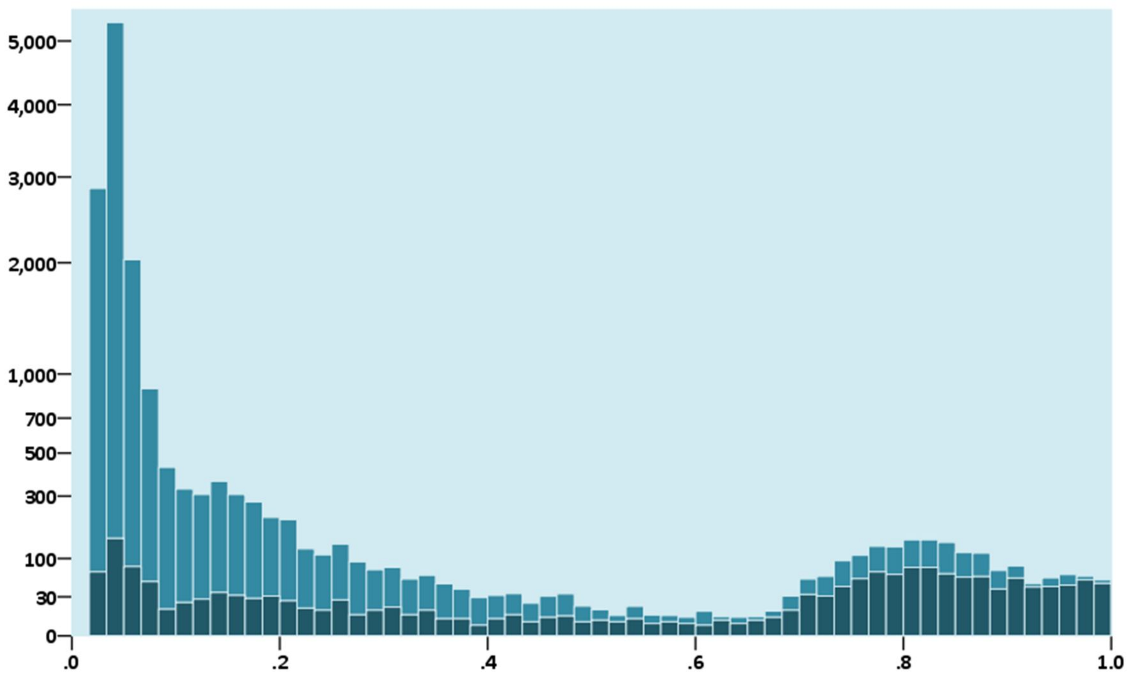
8.3 M&A formation model

The two previous sections presented evidence on the home bias and industrial relatedness bias of the bidder. In this section the logistic regression model, which was already used for this purpose, is assessed. First, I want to know what the contribution of geographical proximity, industrial relatedness and the control variables is on M&A formation (*research question B4*). **Figure 19** shows the predicted probabilities of the final model. These are, however, not the right values as most non-deals are missing. The probabilities need to be corrected. As the values corrected by ReLogit are about the same as the values given by estimating the model with the full sample (see 7.2) it is sufficient to examine the latter ones. **Figure 20** shows the distribution of those probabilities. The mean is 0.549%. As the deal probability is 1,855/1,608,863 = 0.115% if deals occur randomly, the modeled likelihood is significantly higher.

A logistic regression model is of less value as long as we do not know how well it fits the data. Not detecting potential outliers can distort the validity of the inferences drawn from the modeling (Sarkar et al., 2011). In order to evaluate the model three issues were assessed: (a) testing its generalizability , (b) identifying cases that exert an undue influence on the model, i.e. outliers that bias the model, and (c) identifying outliers for which the model fits poorly.

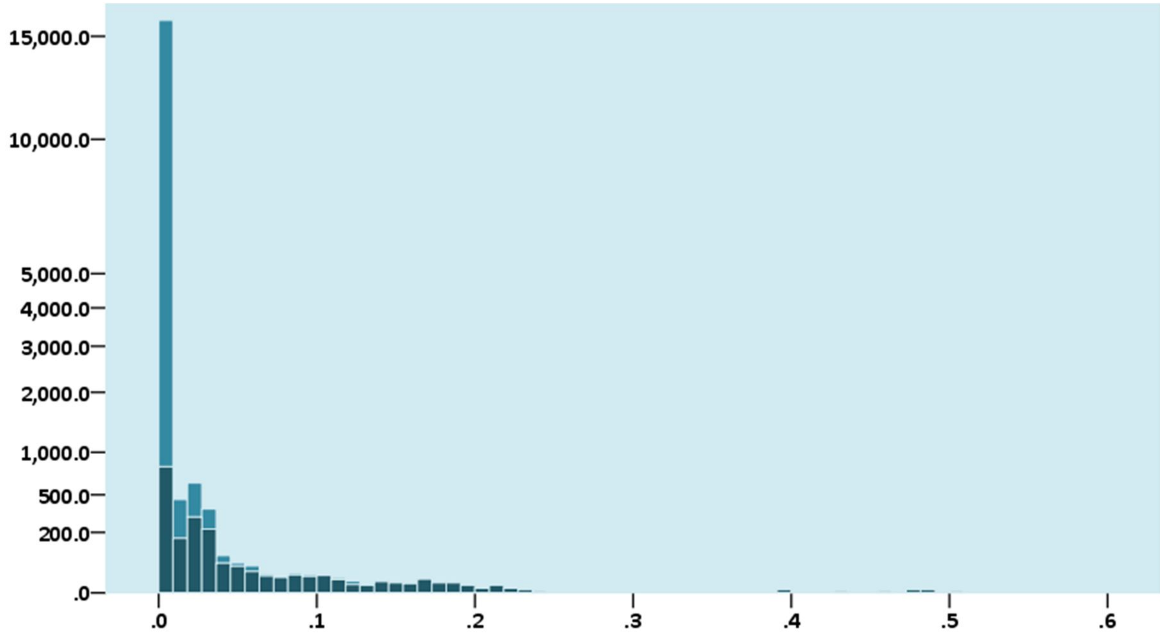
Tabachnick and Fidell (2007) suggest to cross-validate the model by splitting the data randomly into a selection of 80% and run the model on this training sample and its counterpart, the holdout sample, with the remaining 20% of the data. I have done that for the final model and **Appendix Table 24** shows that there are no severe differences in the values of the R^2s and *b*s. Also the effect sizes and significance values match very well with the overall sample. This means the model is generalizing quite well. Cases that exert an undue influence are very unlikely as there are no cases with extreme values; Tanimoto coefficients are always within 0 and 1 and also distance never exceeds 330km. One statistic to estimate the overall influence of a single case on the model is Cook's distance. Cook and Weisberg (1982) suggest that this statistic should not exceed 1. Indeed, there were no influential cases or outsiders which bias the model, as the largest value was 0.10.

Figure 19 Uncorrected predicted probabilities of M&A formation.



Histogram (#) of uncorrected predicted probabilites. Bright colouring indicates the fraction of non-deals and dark colouring the fraction of real deals. The y-axis is quadraticly scaled.

Figure 20 **Actual predicted probabilities of M&A formation.**



Histogram (#) of uncorrected predicted probabilites. Bright colouring indicates the fraction of non-deals and dark colouring the fraction of real deals. The y-axis is quadraticly scaled.

This does not mean, however, that there is a good fit between data and model. There are two ways to assess the goodness of fit. First, we can analyse in how far predicted group membership and real group membership, i.e. whether a deal is real or a non-deal, match. For each case deal likelihood can be estimated. The only difficulty is to decide on the optimal cut-point, which allocates a case either to the group of deals or non-deals. **Figure 19** shows that there are many deals that had a very low likelihood being a deal and there were some cases that had a very large likelihood being a deal but actually were non-deals. These are cases which were not rightly predicted by the model. However, there were not only very low and very high probabilities. Many cases had a moderate probability being a deal and therefore setting a cut-point is a subjective decision. If the cut-point was 0.5 97.8% of the non-deals were rightly predicted but only 58.2% of the deals. Predicting non-deals is not a problem, especially as there were much more non-deals than deals. More important is how well the model predicts real deals. Shifting the cut-point down the rightly predicted fractions are approximating each other at a cut-point of about 0.085, where for both groups about 80% of the cases were rightly predicted.

Another possibility is to examine standardized and studentized residuals and deviance statistics. Residuals are the difference between predicted and observed probability value for each observation. Standardized residuals are residuals that are divided by a common, average standard deviation of the residuals. Stundentized residuals are each divided by its own standard deviation. Field (2009) suggests watching out for residuals with an absolute value larger than 1.96, 2.58 and 3.29. The fractions of cases with these large values should not exceed 5%, 1% and 0%, respectively. These limits were kept for standardized residuals and deviance, but not for studentized residuals. There were too many cases with too large or too small standardized residuals (**Table 30** and **Figure 21**). These are about the same cases as deals with very low deal likelihood or non-deals with very high deal likelihood.

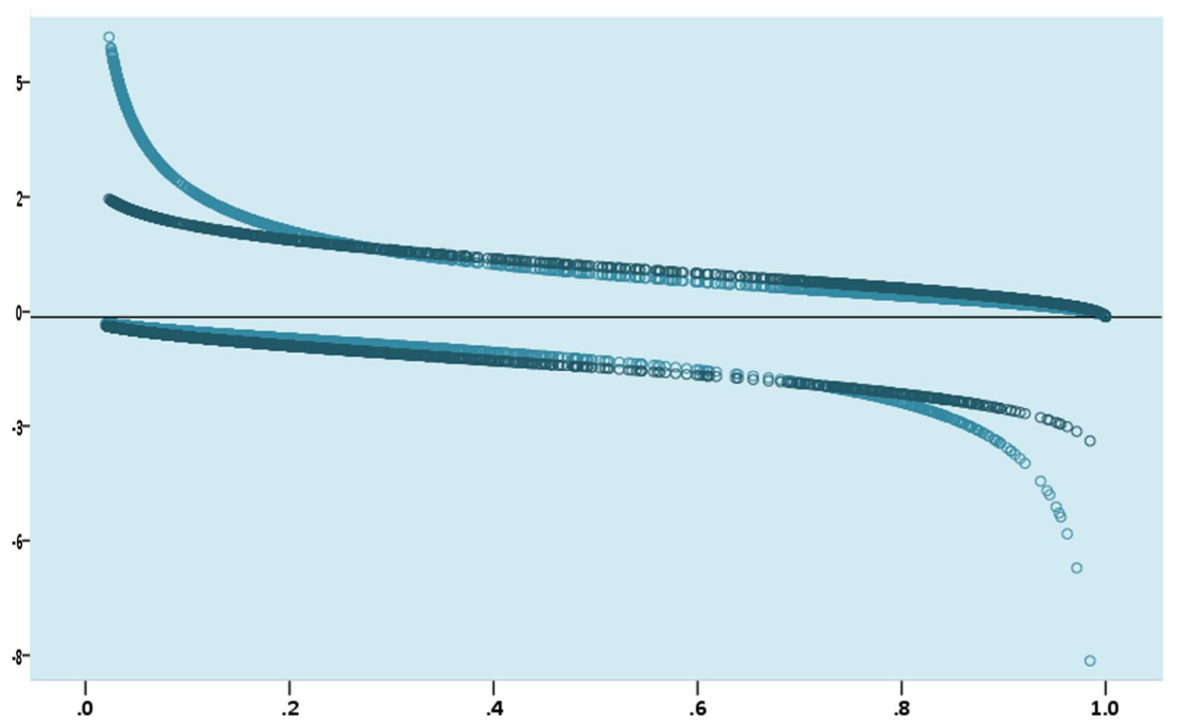
I further analysed those cases with a standardized residuals value larger than 3 because unfortunately exceptions do not prove the rule. Within these exceptional cases there were 375 deals and 15 non-deals. The non-deals were all within the same group and have high Tanimoto coefficients. We can speculate about the reasons for non-merging, such as better offers from other firms, managers that do not like each other, etc., however, much more important is to examine the 375 deals which were predicted wrongly by the models. These deals count for about 20% of the 1,855 deals. It appears that most of these deals were highly industrial unrelated deals, all within the same section and with very low Tanimoto coefficients. There are two main reasons why two firms announce a deal although the models predict the contrary. First, important predictors are missing, as discussed in 3.4. Second, the target is a private equity firm that invests in all kinds of industries. Indeed, in 66 of the 363 cases NACE group **649: Other financial service activities, except insurance and pension funding** is assigned to the bidder, which might be an indication for that.

Table 30 Large residuals in the logistic regression model.

Size of residuals	SRE	ZRE	DEV	Recommended maximum
Outside ± 1.96	4.25	5.71	4.25	5.00
Outside ± 2.58	0.76	3.99	0.75	1.00
Outside ± 3.29	0.00	3.18	0.00	0.00

Percentages of Studentized Residuals (SRE), Standardized Residuals (ZRE) and Deviances (DEV) outside the critical thresholds.

Figure 21 The residuals in the final logistic regression model.



Y: Standardized (dark) and studentized residuals (bright). X: Deal likelihood.

9 Discussion of results

This chapter discusses and interprets the results presented in the previous chapter and states how much evidence was found for the hypotheses.

9

Purpose of the empirical tests was to estimate structural differences between deals and non-deals based on the involved company's locations and industrial activities. Non-deals are pairs of companies that were not involved in any M&A transaction and for which no announcement of such has taken place. Deals describe company dyads that either were involved in a transaction or for which a transaction was announced. The information on location and industrial activity was the basis to investigate differences in geographical proximity and industrial relatedness between the involved companies. Both dimensions were analysed on an average level and from the perspective of the bidder (home bias and industrial relatedness bias). Furthermore, both dimensions were incorporated into a logistic regression model that explains M&A formation. As geographical proximity and industrial relatedness are not the only factors that determine which company dyad available on the M&A market would announce or engage in a transaction, I discussed in the theoretical part several other relational factors as possible facilitators. In the remainder of this chapter I discuss the empirical results on geographical proximity, industrial relatedness and the performance of the M&A formation model.

All tests of *hypotheses A1* and *A2* which address the effect of geographical proximity performed as expected. There were no unexpected results. The tests were performed under different assumptions. Overall, the home bias was always smaller if assuming that bidders can only select between targets of the bidder's real target's industry (assumption SG) compared to the more optimistic results under the assumption that any industry can be chosen (assumption ALL). For example, while the deals are 7.9 times more likely within the same municipality under assumption ALL it is only 4.6 times under assumption SG. This is because assumption SG controls for localization effects, which are due to clustering. As bidder and target are often active in the same industry (see **Figure 11**) and industries are often clustered in spaced, potential M&A partners are often nearby. In consequence, this leads to more geographical proximate non-deals than if all possible non-deals are considered.

Beside these two assumptions *hypotheses A1* and *B1* are tested by different methods. The weakest method is the bivariate comparison, resulting in results that are not controlled for other variables or the characteristics of the individual choice set of the bidder. The biases are given by the median differences. However, they only express the average home bias. They do not reveal anything about the deviation of home bias per individual bidder and do not consider their individual choice set, which is certainly different for a bidder located in the North of Groningen and for one located within the Randstad. The univariate analyses of the newly developed absolute and relative individual bias measures control for this and are therefore regarded to be more reliable. In fact, they only include bidders that could choose between at least 3 or 5 different potential targets. While the absolute bias measures did not consider the minimum and maximum distance, and Tanimoto difference, to the potential targets the relative home bias measure did so. The third method is logistic regression, which does not control for the characteristics of the choice set but for other control variables. While the bivariate and univariate methods are applied to different spatial scales the logistic regression model is only estimated for all deals.

The positive effect of geographical proximity on a national level is not a surprising finding, because it is known that a domestic home bias inheres in bidders (Eun and Mukherjee, 2006; Grote and Umber, 2006; Chakrabarti and Mitchell, 2008). Striking is, however, that home bias even prevails within regions within the Netherlands (*research question A3*). That means that geographical proximity even has significant effects when considering deals and non-deals in which both companies are located within the same region. The tests of the hypothesis revealed a significant effect for deals within the same province, COROP region and even municipality. Under assumption ALL the results are highly significant for all three sub-levels and highly robust. The effect was found by comparing the medians, by investigating the individual home bias measures and by the logistic regression model. The estimates of the logistic regression model show that distance, and even its logarithmic term of it, remains significant when including all three regional level dummies. Under assumption SG, however, the results are only robust on the province and COROP region level. On the municipality level the effect was only significant when tested by the median difference. In this case the home bias was 600 to 1,400 m (with 95% confidence). The result of the individual home bias measure is insignificant (-400 to 1,200 m). This means, that localization effects are sufficient to explain home bias on a municipality level.

While the study on home bias in M&As of Grote and Umber (2006) has only presented numbers on the absolute home bias, this research also considered the characteristics of the choice set of every bidder, what allows a standardization of the bias estimates. Consequently standardized values allow a comparison of the home bias effect for bidders that engaged in deals on different spatial scales. The relative home bias indicates whether a bidder has selected the most distant target (-100) or the closest target available (100). In this sense the average values can be interpreted. Under assumption SG, the highest average relative home bias is for bidders that bid on targets of the same province with a value between 31.4 and 46.8 (with 95% confidence). However, this average value does not significantly differ from the ones for the whole Netherlands (28.8 to 35.1) and for deals within the same COROP region (10.2 to 37.9). As already mentioned the home bias for inter-municipality deals is not significantly positive (-9.6 to 45.5).

The advantages of the individual measures are not only that the available choice set can be considered and that it can be standardized, these measures also reveal the proportion of home biased and distance biased bidders. Also here intra-regional deals were tested. As expected most bidders are home biased. On a national level and under assumption SG the proportion is 67 to 71% (with 95% confidence). For intra-municipality deals the proportion of home biased bidders is not significantly larger than the proportion of distance-biased bidder, as it is 40 to 71% (with 95% confidence). This is in line with the insignificant average home bias for those cases.

The tests of the effects of industrial relatedness follow quite strictly the approach used for the tests of geographical proximity. Also here a categorical and continuous measure is available, an individual bias measure was developed and the same testing methods were used. Almost needless to say that the assumption that bidders can only select between companies of the same industry as their real target, however, does not make sense in this case. However, the robustness tests on different spatial scales were exerted as well. While it is self-evident that the absolute home bias massively differs on different spatial scales this is not the case for industrial relatedness.

The effect of belonging to the same NACE group, division or section is easily understandable. If both companies share the same NACE group they are 26.1 times more likely to announce a M&As, if they belong to the same NACE division but different groups it is only 4.1 and only 1.7 if they merely share the same section. These are the results of the bivariate analysis on a national level. In contrast, the interpretation of the differences in the Tanimoto coefficient and the absolute industrial relatedness bias is not intuitive. An interpretation requires the understanding of the formula how the coefficient is constructed. Therefore it makes much sense to use the relative industrial relatedness bias. The average is quite similar to the average of the relative home bias measure: 32.9 to 39.8 (on a 95% confidence level). Unsurprisingly the majority of bidders have a positive relatedness bias (*hypothesis B2*). The share of bidders that selected a less related target than random selection would suggest is not more than 24% (on a 95% confidence level).

Interesting is that industrial relatedness bias is stronger when controlling for the characteristics of the bidder's choice set (univariate analysis). This is different than for the home bias, where the effect was weaker. When controlling for geographical proximity and some control variables in the logistic regression model the estimates of the dummy variables are much higher, namely 89.4, 12.1 and 5.2

instead of 26.1, 4.1 and 1.7. That means that local deals are even more related, or in other words that both dimensions are enforcing each other. This is also supported by some additional tests that compare the home bias of industrially biased and industrially unbiased bidders (not reported). The estimations suggested that local deals are more related than non-local deals. Therefore it would also being worth to investigate the links between geographical proximity, industrial relatedness, home bias and industrial relatedness bias in future research. The role of some control variables is significant but unclear. Those variables were constructed because the data allowed it, but not with any specific theoretical reasons. The estimations revealed that if both companies have at least one subsidiary or if both companies are active in at least two business activities the chance of a deal decreases. When excluding multiple bidders from the analysis another variable turns significant. If both companies are unlisted deal likelihood decreases as well. These findings can be explored in future research.

When comparing the contribution of industrial relatedness and geographical proximity on M&A formation the estimations clearly show that the impact of industrial relatedness is much larger (*research question B3*). It is also shown that both dimensions are reinforcing each other. The indicators of geographical proximity always remain significant. This means that industrial relatedness is not a substitute for geographical relatedness.

The logistic regression model reveals some important insights, not only in deal announcement likelihood, but also in the contribution of the different variables used for measuring geographical proximity and industrial relatedness. It allows measuring the contribution of distance, of belonging to the same region, of the Tanimoto coefficient and of belonging to the same NACE level and their relationships to each other. When modelling the effect of distance on deal announcement a logarithmic relationship between distance and the dependent variable was found. In other words, the performance of the model was better with the logarithmic term than with the standard term. The reason is the well-known effect that the chance of collaboration decreases with increasing distance, i.e. the increase in chance is much higher when comparing 3km to 2km than comparing 165km to 164km of distance. Also Eun and Mukherjee (2006) found this effect, although they used classes with unequal sizes instead of a logarithmic term. An interesting finding is that the effect of the logarithmic term (but also the standard) term of distance remains significant when estimating it in conjunction with the categorical variables, or vice versa, respectively. This means that not regions are deciding but also pure distance has an effect. This is in line with the finding that distance has an effect even on deals on very small regional scales. If one can chose between modeling the effect of geographical distance as the logarithmic term of distance and as regional dummies the logarithmic term of distance performs better. If one can chose between distance and the regional dummies the regional dummies perform better.

While both measurements of geographical proximity contribute to M&A formation independently this is not the case for the two measurements of industrial relatedness. They serve as substitutes. It is furthermore rather surprising that the dummy variables clearly outperform the Tanimoto coefficient. The Tanimoto coefficient has no added value in explaining M&A formation. The main reason is the high correlation between the two measurements. The Tanimoto coefficient can almost perfectly predict whether two companies share the same NACE group or not. It could be that the Tanimoto coefficient contributes more to the model if estimated on a NACE class (4 digits) level. The failure of the Tanimoto coefficient to explain M&A formation does, however, not imply that its development is useless. It is necessary to estimate the individual industrial relatedness bias.

The assessment of the final model revealed that some cases do not fit the model. Those cases did not show any significant commonalities in the data. It is suspected that especially a variable that indicated buyer-supplier relationships is missing, but also other relational variables as discussed in the theoretical section might contribute to the model. As financial bidders bid on unrelated targets, these deals could be responsible for the weak fit as well. Consequently the presented model would only explain horizontal and conglomerate M&As. However, as already explained in the methodological section these variables are either not available (e.g. complete data of firm sizes), inconstructible on the basis of single-firm data (e.g. buyer-supplier relationship), hard to collect (similarities in management styles) or hard to measure. For example, similarities in management styles can only be detected by intensive interviews. Also identifying potentials of financial synergies require a deep understanding and investigation. One solution might be to estimate a model with fewer cases but more and better variables.

Another issues related to the logistic regression estimations is the fact that M&As are rare events. Even given exclusively companies that are active on the M&A market the average chance of a deal is small, namely 0.115%, if deals occur randomly. The model predicts an average deal probability of 0.549%, with some outliers that even reach up to about 50% (*research question B4*). The small probabilities depend on the number of potential targets per bidder, or to say it differently, on the number of potential deals. Considering only potential targets that were actually considered by the bidder (by asking the responsible manager of the bidder company for example) there would, however, already be a inherit bias. Therefore I considered, given some assumptions and criteria, all available potential targets per bidder. Only in this way it was possible to estimate valid home and industrial relatedness biases, resulting in very small probabilities. As the probabilities are very small the model cannot be used for predictions, i.e. making valid projections concerning an outcome of a particular individual. It is only possible to examine how the outcome correlates to certain variables on a group level (Osborn, 2000).

All estimations were exerted for companies of all kinds of industries over a time span of eight years. Next to the extension of explanatory variables future research can address two other dimensions: (a) the examination of the temporal and (b) the examination of the industrial dimension of the outcomes of this study. Merger wave literature suggested that M&A behaviour changes significantly over time. Indeed, **Appendix Figure 5** shows that home bias significantly decreased in 2008 compared to the previous years, where the average home bias was stably ranging from about 30 to 40. In order to find out whether this was an abnormal shift or not research needs to be done on a longer time span. For the industrial relatedness bias no significant temporal trends could be found (**Appendix Figure 6**).

Industrial relatedness bias and geographical proximity bias might not only differ in time but also between different industries. **Appendix Figure 7** shows that the distribution of home bias differed between some industries while it did not between others. For example, Independent-Samples Mann-Whitney U Test, Independent-Samples Kolmogorov-Smirnov Test and Independent Samples Median Test showed all highly insignificant similarities between **F: Construction** and **G: Wholesale and retail trade; Repair of motor vehicles and motorcycles**. Between **F: Construction** and **J: Information and Communication** all three tests showed highly significant differences. The histograms in **Appendix Figure 7** show that the distribution of industrial relatedness bias had one common feature between industries. In most industries the mostly related target was acquired. For example, Independent-Samples Mann-Whitney U Test was insignificant for **F: Construction** and **G: Wholesale and retail trade; Repair of motor vehicles and motorcycles**, while Independent-Samples Kolmogorov-Smirnov Test and Independent Samples Median Test did show significant differences. However, **K: Financial and insurance activities** showed a significantly higher share of unrelated targets. All three tests between, for example, **F: Construction** and **K: Financial and insurance activities** showed highly significant differences.

10 Conclusion

Having discussed the findings of this research, this chapter draws the conclusions.

10

The first part of the aim was to test and estimate the home bias and industrial relatedness bias of bidders. Very robust evidence is presented. Sophisticated statistical analyses show highly significant and partly very strong structural differences between deals and non-deals. The existence of a domestic home bias is not surprising as for both phenomena a number of reasons have been discussed: strategic reasons, information advantages, familiarity and localization economies. It was possible to control for the effect of localization economies. The estimation differences between the models under different assumption concerning their choice set as well as some further research on the role of localization economies (see **Additional Research 1** in the **Appendix**) show that this effect is very prominent, although it could not be estimated to what extent because the remaining explanations could not be disentangled within this quantitative research project. Their contribution certainly differs per bidder. For example, familiarity has probably a higher impact for small bidders, where the manager is strongly embedded in the local context, than for large companies. Strategic reasons could differ in sectors. A decided analysis of the underlying explanations requires in-depth studies.

As I do not distinguish between the discussed factors that are responsible for the role of geographical proximity and industrial relatedness it might also be interesting to model the home bias measure as dependent variable and test the effect of various firm and regional characteristics on the bias. Eun and Mukherjee (2006) and Böckerman and Lehto (2006) already found that large firms are less home biased. This is because decision making processes in small companies follow a different logic than in large ones. Böckerman and Lehto (2006) furthermore found that old firms and multi-locational firms are less home biased. Eun and Mukherjee (2006) found that home bias differs in regions. That means that also regional factors could play a role.

Besides home bias the thesis delivers evidence on the so-called industrial relatedness bias. The evidence is extremely robust and results are highly significant. The main rational is that the involved companies can realize synergy effects, but also information advantages might play a role. To my knowledge, this kind of bias has never been estimated before and might therefore be an important theoretical contribution to the diversification and M&A literature. Another important contribution is the development of the relative bias measures. The estimation of relative bias values allows comparing the findings of this research conducted in the Netherlands with findings in other countries, if the method is adopted. So far, this is unfortunately not possible because previous studies only presented absolute values.

One explanatory part of the aim was to estimate the contribution of geographical proximity and industrial relatedness to M&A formation and to investigate their interaction. It appears that industrial relatedness is a crucial facilitator of M&A formation and geographical proximity enforces M&A formation additionally. This means both factors function in conjunction and cannot be substituted. Interestingly, the dummy variables indicating belonging to the same industry contribute more to the explanation of M&A formation than the Tanimoto coefficient, which was initially believed to show a better performance than the dummy variables. Another explanatory finding was that not only a domestic home bias exists, but also a regional home bias. Bidders are even home biased if they select targets from the same province, COROP regional and municipality. For deals within the same municipality mainly localization effects are responsible. For deals within the other regions I assume that the same causes hold as for the domestic home bias. Nonetheless, this is an interesting outcome which asks for further research.

In the theoretical framework other relational factors that facilitate M&A formation were discussed. The identified factors can be incorporated into the taxonomy of the proximity concept, as described by Knoben and Oerlemans (2006) or Boschma (2005), which is certainly useful if more variables are included. It must be clear, however, that M&A formation is about impossible to predict, even if all discussed variables are considered. A large part of the explanation is rooted in history and coincidences. As empirical research is reductionist by design such individual factors cannot be examined. Also other focal factors which might play a role, such as the price of the target, could not be incorporated.

Although taking a relational view on M&A formation, the research followed a bidder dominated design. While the M&A formation model and the estimations of average differences between deals and non-deals do not make an assumption about who initiated the deal, the individual bidder bias measurements does so. They support the view of the dominance of the bidder. This bidder dominance is inherent in almost all works in the M&A literature, which does surely not justify it. It is realistic that in most cases the bidder initiates the deal and only in few cases the target has the leading role. As there is no information on the activeness or passiveness of the bidders and targets it is reasonable to assume that it is always the bidder. It would be desirable for future research to attach value to the seller as well.

Having evidence provided for Dutch domestic deals the question remains whether the findings of this research might be true for other countries and settings as well. From an empirical and statistical point of view the results cannot be inferred to M&As in other locations because this requires a random sample of deals from different countries. However, the effect of geographical proximity was found on different spatial scales and even the largest scale (The Netherlands) is in fact rather small. Due to this strong evidence it is believed that home bias is also existent in other countries. It would also be reasonable because the underlying causes are not country-specific. The same holds for the industrial relatedness bias.

Software and Data

Software

ESRI Inc., 2009: ArcMap 9, Version 9.3.1, Redlands, CA.

IBM SPSS, 2009: PASW Statistics, Version 18.0, Chicago.

Microsoft Corporation, 2010: Access 2010, Version 14.0.

Microsoft Corporation, 2010: Excel 2010, Version 14.0.

StataCorpLP, 2009: Stata/SE, Version 11.0, Lakeway Drive College Station.

Tomz, M., King, G., and Zeng, L., 1999: RELOGIT: Rare Events Logistic Regression, Version 1.1, Cambridge.

Data

REACH: Company information by Bureau van Dijk. https://reach2.bvdinfo.com.

Thomson ONE Banker: M&A information by Thomson Financial. http://banker.thomsonib.com.

ZEPHYR: M&A information by Bureau van Dijk. https://zephyr2.bvdep.com.

References

Ahuja, G., & Katila, R. 2001. Technological Acquisitions and the Innovation Performance of Acquiring Firms: A Longitudinal Study. ***Strategic Management Journal***, 22: 197–220.

Alexandridis, G., Mavrovitis, C. F., & Travlos, N. G. 2010. How Have M&As Changed? Evidence from the Sixth Merger Wave. ***Working Paper***: 1–37.

Bachmann, R. 2001. Trust, Power and Control in Trans-organizational Relations. ***Organization Studies***, 22(2): 337–365.

Barney, J. B. 1988. Returns to Bidding Firms in Mergers and Acquisitions: Reconsidering the Relatedness Hypothesis. ***Strategic Management Journal***, 9: 71–78.

Berkovitch, E., & Narayanan, M. P. 1993. Motives for Takeovers: An Empirical Investigation. ***Journal of Financial and Quantitative Analysis***, 28(3): 347–362.

Bewick, V., Cheek, L., & Ball, J. 2005. Statistics review 14: Logistic regression. ***Critical care***, 9(1): 112–118.

Blackburn, V. L., & Lang, J. R. J. K. H. 1990. Mergers and Shareholder Returns: The Roles of Acquiring Firm's Ownership and Diversification Strategy. ***Journal of Management***, 16(4): 769–782.

Böckerman, P., & Lehto, E. 2006. Geography of Domestic Merger and Acquisitions (M&As): Evidence from Matched Firm-level Data. ***Regional Studies***, 40(8): 847–860.

Bogan, V., & Just, D. 2009. What Drives Merger Decision Making Behavior?: Don't Seek, Don't Find, and Don't Change Your Mind. ***Journal of Economic Behavior & Organization***, 72(3): 930–943.

Boschma, R. 2005. Proximity and Innovation: A Critial Assessment. ***Regional Studies***, 39(1): 61–74.

Bottazzi, G., & Pirino, D. 2010. Measuring Industry Relatedness and Corporate Coherence. ***Working Paper***: 1–27.

Bouwens, B. 2005. Below sea level: a second merger wave in the Netherlands? ***Working Paper***: 1–20.

Brakman, S., Garretsen, H., & Van Marrewijk, C. 2006. Cross-Border Mergers & Acquisitions: The Facts as a Guide for International Economics. ***CESifo Working Paper***(1823): 1–36.

Brar, G., Giamouridis, D., & Liodakis, M. 2009. Prediciting European Takeover Targets. ***European Financial Management***, 15(2): 430–450.

Brauer, M. 2006. What Have We Acquired and What Should We Acquire in Divestiture Research? A Review and Research Agenda. ***Journal of Management***, 32(6): 751–785.

Breslow, N. E., & Day, N. E. 1980. The Estimation of Choice Probabilities from Choice Based Samples. ***International Agency for Research on Cancer Scientific Publications***, 1(32): 388 pp.

Broekel, T., & Boschma, R. 2009. Knowledge networks in the Dutch aviation industry: the proximity paradox. ***Working Paper***: 1–34.

Brouthers, K. D., van Hastenburg, P., & van den Ven, J. 1998. If Most Mergers Fail Why Are They so Popular? ***Long Range Planning***, 31(3): 347–353.

Cai, Y., & Sevilir, M. 2009. Board connections and M&A transactions. ***Working Paper***: 1–48.

Cai, Y., & Tian, X. 2009. Firm Locations and Takeover Likelihood. ***Working Paper***: 1–40.

Capron, L. 1999. The Long-Term Performance of Horizontal Acquisitions. ***Strategic Management Journal***, 20(11): 987–1018.

Capron, L., & Shen, J. C. 2007. Acquisitions of Private vs. Public Firms: Private Information, Target Selection, and Acquirer Returns. ***Strategic Management Journal***, 28: 891–911.

Cartwright, S. 2006. 30 Years of Mergers and Acquisitions Reserach. Recent Advances and Future Opportunities. ***British Journal of Management***, 17(S1): S1-S5.

Cartwright, S., & Price, F. 2003. Managerial Preferences in International Merger and Acquisitions Partners Revisited: How are they influenced? In A. Gregory (Ed.), ***Advances in Mergers and Acquisitions. Volume 2***: 81–95. Oxford: JAI Press (Elsevier Science ltd).

Cefis, E., Marsili, O., & Schenk, H. 2009. The effects of mergers and acquisitions on the firm size distribution. ***Journal of Evolutionary Economics***, 19: 1–20.

Chakrabarti, A., & Mitchell, W. 2008. The Persistent Effect of Geographic Distance in Acquisition Target Selection. ***Working Paper***: 1–50.

Chapple, L., & Treepongkaruna, S. 2006. The Impact of Target Board Recommendations in Australian Takeovers. ***Working Paper***: 1–46.

Chatterjee, S., Lubatkin, M. H., Schweiger, D. M., & Weber, Y. 1992. Cultural Differences and Shareholder Value in Related Mergers: Linking Equity and Human Capital. ***Strategic Management Journal***, 13(5): 319–334.

Chen, H., Gompers, P., Kovner, A., & Lerner, J. 2010. Buy local? The geography of venture capital. ***Journal of Urban Economics***, 67: 90–102.

Colcera, E. 2007. ***The Market for Corporate Control in Japan***: ***M&As, Hostile Takeovers and Regulatory Framework***. Berlin, Heidelberg, New York: Springer.

Cook, R. D., & Weisberg, S. 1982. ***Residuals and influence in regression***. New York: Chapman & Hall.

Coval, J. D., & Moskowitz, T. J. 1999. Home Bias at Home: Local Equity Preference in Domestic Portfolios. ***The Journal of Finance***, 54(6): 2045–2073.

Cumming, D., & Dai, N. 2010. Local bias in venture capital investments. ***Journal of Empirical Finance***, 17(3): 362–380.

Eun, C. S., & Mukherjee, S. 2006. The Geography of M&A: Contours and Causes. ***Working Paper***: 1–49.

Fan, J. P. H., & Lang, L. H. P. 2000. The Measurement of Relatedness: An Application to Corporate Diversification. ***The Journal of Business***, 73(4): 629–660.

Field, A. 2009. ***Discovering Statistics Using SPSS***. London: Sage.

Flanagan, D. J. 1996. Announcements of purely related and purely unrelated mergers and shareholder returns: reconciling the relatedness paradox - includes appendix. ***Journal of Management***.

Garcia, E. Cosine Similarity and Term Weight Tutorial; http://www.miislita.com/information-retrieval-tutorial/cosine-similarity-tutorial.html#Cosim.

Gorton, G., Kahl, M., & Rosen, R. 2005. Eat or Be Eaten: A Theory of Mergers and Merger Waves. ***NBER working paper***.

Graebner, M. E. 2004. Momentum and serendipity: How acquired leaders create value in the integration of technology firms. ***Strategic Management Journal***, 25: 751–777.

Graebner, M. E., & Eisenhardt, K. M. 2004. The Seller's Side of the Story: Acquisition as Courtship and Governance as Syndicate in Entrepreneurial Firms. ***Administrative Science Quarterly***, 49(3): 366–403.

Grote, M. H., & Umber, M. P. 2006. Home biased? A spatial analysis of the domestic merging behavior of US firms. ***Working Paper***: 1–56.

Haleblian, J., & Finkelstein, S. 1999. The Influence of Organizational Acquisition Experience on Acquisition Performance: A Behavioral Learning Perspective. ***Administrative Science Quarterly***, 44(1): 29–56.

Hamers, L., Hemeryck, Y., Herweyers, G., Janssen, M., Keters, H., Rousseau, R., & Vanhoutte, A. 1989. Similarity Measures in Scientometric Research: The Jaccard Index Versus Salston's Cosine Formula. ***Information Processing & Management***, 25(3): 315–318.

Hasbrouck, J. 1985. The Characteristics of Takeover Targets: Q and Other Measures. ***Journal of Banking & Finance***, 9: 351–362.

Haunschild, P. R., & Beckman, C. M. 1998. When Do Interlocks Matter?: Alternate Sources of Information and Interlock Influence. ***Administrative Science Quarterly***, 43(4): 815–844.

Hillberry, R. 2003. Intranational home bias: Some explanations. ***Review of Economics and Statistics***, 84(4): 1089–1092.

History. 1998. Daimler-Benz announces purchase of Chrysler Corp.; http://www.history.com/this-day-in-history/daimler-benz-announces-purchase-of-chrysler-corp.

Homberg, F., Rost, K., & Osterloh, M. 2009. Do synergies exist in related acquisitions?: A meta-analysis of acquisition studies. ***Review of Managerial Science***(3): 75–116.

Hosmer, D. W., & Lemeshow, S. 2000. ***Applied logistic regression***. New York: Wiley.

Huberman, G. 2001. Familiarity Breeds Investment. ***The Review of Financial Studies***, 14(3): 659–680.

Hussinger, K. 2010. On the importance of technological relatedness: SMEs versus large acquisitions targets. ***Technovation***, 30: 57–64.

Ishii, J., & Xuan, Y. 2010. Acquirer-Target Social Ties and Merger Outcomes. ***Working Paper***: 1–48.

Kang, J. K., & Kim, J. M. 2008. The Geography of Block Acquisitions. ***The Journal of Finance***, 63(6): 2817–2858.

Katz, J. S. 2005. Geographical proximity and scientific collaboration. ***Scientometrics***, 31(1): 31–43.

Kenny, D. A., Kashy, D. A., & Cook, W. L. 2006. ***Dyadic Data Analysis:*** Guilford.

King, D. R., Dalton, D. R., Daily, C. M., & Covin, J. G. 2004. Metha-analyses of Post-acquisition Performance: Indications of Unidentified Moderators. ***Strategic Management Journal***, 25: 187–200.

King, G., & Zeng, L. 2001a. Explaining Rare Events in International Relations. ***International Organization***, 55(3): 693–715.

King, G., & Zeng, L. 2001b. Logistic Regression in Rare Events Data. ***Political Analysis***, 9(2): 137–163.

Knoben, J., & Oerlemans, L. A. G. 2006. Proximity and inter-organizational collaboration: A literature review. ***International Journal of Management Reviews***, 8(2): 71–89.

Laulajainen, R. 1988. The Spatial Dimension of an Acquisition. ***Economic Geography***, 64(2): 171–187.

Limmack, R. J., & Mcgregor, N. 1995. Industrial relatedness, structural factors and bidder returns. ***Applied Financial Economics***, 5(3): 179–190.

Lipton, M. 2006. Merger Waves in the 19th, 20th and 21st Centuries. ***Working Paper***: 1–21.

Lubatkin, M. 1983. Mergers and the Performance of the Acquiring Firm. ***The Academy of Management Journal***, 8(2): 218–225.

Luypaert, M., & Huyghebaert, N. 2007. Determinants of growth through mergers and acquisitions: An empirical analysis. ***Working Paper***: 1–55.

Manski, C. F., & Lerman, S. R. 1977. The Estimation of Choice Probabilities from Choice Based Samples. ***Econometrica***, 45(8): 1977–1988.

Martin, J. D., & Sayrak, A. 2003. Corporate diversification and shareholder value: a survey of recent literature. ***Journal of Corporate Finance***, 9(1): 37–57.

Matsusaka, J. G. 1993. Takeover Motives during the Conglomerate Merger Wave. ***The RAND Journal of Economics***, 24(3): 357–379.

Matsushima, N. 2001. Horizontal mergers and merger waves in a location model. ***Australian Economic Papers***, 35: 265–286.

Maynard, T. H. 2009. ***Mergers and Acquisitions***: ***Cases, Materials, and Problems***. New York: Aspen Publishers.

McFarland, D. D. 1975. Models of Marriage Formation and Fertility. ***Social Forces***, 54(1): 66–83.

Menard, S. 2002. ***Applied Logistic Regression Analysis***. Teller Road: Sage Publications.

Miller, D. J. 2006. Technological Diversity, Related Diversification, and Firm Performance. ***Strategic Management Journal***, 27: 601–619.

Mitchell, W., & Shaver, J. M. 2003. Who Buys What? How Integration Capability Affects Acquisitions Incidence and Target Choice. ***Strategic Organization***, 1(2): 171–201.

Mizruchi, M. S. 1996. What Do Interlocks Do? An Analysis, Critique, and Assessment of Research on Interlocking Directorates. ***Annual Review of Sociology***, 22: 271–298.

Montgomery, C. A. 1994. Corporate Diversification. ***Journal of Economic Perspectives***, 8(3): 163–178.

Morck, R., Shleifer, A., & Vishney R. W. 1990. Do Managerial Objectives Drive Bad Acquisitions? ***The Journal of Finance***, 45(1): 31–48.

Morosini, P., Shane, S., & Singh, H. 1998. National Cultural Distance and Cross-Border Acquisition Performance. ***Journal of International Business Studies***, 29(1): 137–158.

Nachum, L., & Wymbs, C. 2005. Product differentiation, external economies and MNE location choices: M&As in Global Cities. ***Journal of International Business Studies***, 36: 415–434.

Osborn, J. W. 2000. Prediction in multiple regression. ***Practical Assessment, Research & Evaluation***, 7(2).

Palepu, K. G. 1986. Predicting takeover targets: a methodological and empirical analysis. ***Journal of Accounting and Economics***, 8: 3–35.

Park, H. M. 2008. Univariate Analysis and Normality Test Using SAS, Stata and SPSS. ***Working Paper***: 1–41.

Paulson, E. 2001. ***The Technology M&A Guidebook***. New York: Wiley.

Pearson, C. H. Latitude, Longitude, And Great Circles; http://www.cpearson.com/excel/LatLong.aspx, 13 Feb 2011.

Pehrsson, A. 2006. Business relatedness measurements. ***European Business Review***, 18(5): 350–363.

Poast, P. 2010. (Mis)Using Dyadic Data to Analyze Multilateral Events. ***Political Analysis***, 18(4): 403–425.

Porrini, P. 2004. Can a Previous Alliance Between an Acquirer and a Target Affect Acquisition Performance? ***Journal of Management***, 30(4): 545–562.

Powell, R. G. 1997. Modelling takeover likelihood. ***Journal of Business Finance & Accounting***.

Powell, R. G. 2004. Takeover Prediction Models and Portfolio Strategies: A Multinominal Approach. ***Multinational Finance Journal***, 8(1 and 2): 35–72.

Real, R., & Vargas, J. M. 1996. The Probabilistic Basis of Jaccard's Index of Similarity. ***Systematic Biology***, 45(3): 380–385.

Rodrígues-Pose, A., & Zademach, H. M. 2003. Rising metropoli: The geography of mergers and acquisitions in Germany. ***Research Papers in Environmental and Spatial Analysis***(78): 1–38.

Rodrígues-Pose, A., & Zademach, H. M. 2006. Industry Dynamics in the German Merger and Acquisitions Market. ***Tijdschrift voor Economische en Social Geografie***, 97(3): 296–313.

Rossi, S., & Volpin, P. 2003. Cross-Country Determinants of Mergers and Acquisitions. ***CEPR Discussion Paper*** (3889).

Sarkar, S. K., Midi, H., & Rana, S. 2011. Detection of Outliers and Influential Observations in Binary Logistic Regression: An Empirical Study. ***Journal of Applied Sciences***, 11(1): 26–35.

Schenk, H. 2008a. Firms, managers, and restructuring - implications of a social economics view. In J. Davis & W. Dolfsma (Eds.), ***Companion to Social Economics***. Cheltenham: Edward Elgar.

Schenk, H. 2008b. Mergers and Concentration Policy. In P. Bianchi & S. Labory (Eds.), ***International Handbook of Industrial Policy***: 153–179. Cheltenham: Edward Elgar.

Schenk, H. 2008c. The Merger Paradox. In C. Wankel (Ed.), ***21st Century Management. A Reference Handbook, v. I***: 345–354. Los Angeles: Sage.

Schoenberg, R., & Reeves, R. 1999. What Determines Acquisition Activity within an Industry? ***European Management Journal***, 17(1): 93–98.

Servaes, H. 1996. The Value of Diversification During the Conglomerate Merger Wave. ***The Journal of Finance***, 51(4): 1201–1225.

Seth, A., Song K. P., & Pettit, R. 2000. Synergy, Managerialism or Hubris? An Empirical Examination of Motives for Foreign Acquisitions of U.S. Firms. ***Journal of International Business Studies***, 31(3): 387–409.

Shen, J. C., & Reuer, J. J. 2005. Adverse Selection in Acquisitions of Small Manufacturing Firms: A Comparison of Private and Public Targets. ***Small Business Economics***, 24: 393–407.

Song, M. H., & Walking, R. A. 1993. The Impact of Managerial Ownership on Acquisition Attempts and Target Shareholder Wealth. ***Journal of Financial and Quantitative Analysis***, 28(4): 439–457.

Tabachnick, B. G., & Fidell, L. S. 2007. ***Using multivariate statistics***. New York: Harper and Row.

Ter Wal, A. 2010. ***The Structure and Dynamics of Knowledge Networks: A Proximity Approach:*** PhD Thesis.

Tsagkanos, A., Georgopoulos, A., Siriopoulos, C., & Koumanakos, E. 2008. Identification of Greek Takeover Targets and Coherent Policy Implications. ***Review of Development Economics***, 12(1): 180–192.

UCLA: Academic Technology Services. Annotated SPSS Output Logistic Regression; http://www.ats.ucla.edu/stat/spss/output/logistic.htm.

Weitzel, U., & McCarthy, K. 2009. Theory and Evidence on Mergers and Acquisitions by Small and Medium Enterprises. ***Discussion Paper*** (21).

Wolf, H. C. 2000. Intranational Home Bias in Trade. ***The Review of Economics and Statistics***, 82(4): 555–563.

Yamaha Corporation. European Sales Subsidiaries to Be Consolidated into One Company: Press Release; http://www.yamaha.com/news_release/2008/20080827.html.

Yang, H., & Lin, Z. 2005. Interfirm Distances as Channels for External Resources: The Choice Between Allicances and Mergers & Acquisitions. ***Working Paper***: 1–7.

Yeung, H. W. C. 2005. The Firm as Social Networks: An Organisational Perspective. ***Growth and Change***, 36(3): 307–328.

Zademach, H. M. 2005. ***Spatial Dynamics in the Markets of M&A***: ***Essays on the Geographical Determination and Implications of Corporate Takeovers and Mergers in Germany and Europe***. München: Herbert Utz Verlag.

Zajac, E. J., & Westphal, J. D. 1996. Director Reputation, CEO-Board Power, and the Dynamics of Board Interlocks. ***Administrative Science Quarterly***, 41(3): 507–529.

Zutshi, R. K., & Tan, W. L. 2009. Impact of culture on 'partner selection criteria' in East Asian international joint ventures. ***International Entrepreneurship and Management Journal***, 5(369-393).

Appendices

Additional research 1: The effect of localization economies

Hypothesis

It was argued that the home bias in M&As is associated, among others, with localization economies. The existence of localization economies can be tested for regions as the units of analysis by following hypothesis: *On average, the proportion of intra-regional deals is higher than suggested by the number of potential targets within and outside the region*.

Method

Localization effects are the ratio of actual and expected deals within a region *i*. In other words, we have to divide the actual proportion of intra-regional deals $Intra_i$ by the expected proportion of intra-regional deals $P(Intra_i)$. $Intra_i$ is calculated with D_{intra} as real intra-regional deals and D_{all} as all deals undertaken by all bidders of region *i*. The expected proportion of intra-regional deals for region *i* corresponds with the probability of intra-regional deals when taking the number of potential targets into account. $(P)Intra_i$ is calculated with $PT\ ALL_{intra}$ as intra-regional potential targets and $PT\ ALL_{domestic}$ as the number of all potential targets per bidder in region *i*. If we consider all deals we get this formula

$$Loc\ ALL_i = \left(\frac{\sum_{intra=1}^{n} D_{intra}}{\sum_{all=1}^{n} D_{all}} \Big/ \frac{\sum_{intra=1}^{n} PT\ ALL_{intra}}{\sum_{all=1}^{n} PT\ ALL_{domestic}} \right) 100.$$

The value of this measure can vary between 0 and infinitive. If it is between 0 and 1 there are no localization effects, if it is larger than 1 there are. As done in the tests before it is also possible to assume that bidder can only select between companies of the same industry as their real target. Then we get the formula

$$Loc\ SG_i = \left(\frac{\sum_{intra=1}^{n} D_{intra}}{\sum_{all=1}^{n} D_{all}} \Big/ \frac{\sum_{intra=1}^{n} PT\ SG_{intra}}{\sum_{all=1}^{n} PT\ SG_{domestic}} \right) 100.$$

Alternatively, I suggest two further, modified, measures. If we control for the individual home bias of every bidder and we can still find values larger than one the evidence for localization effect would be even stronger. Therefore we multiply the localization effects with the average home bias of all bidders per region *i*. Also here we can use the basic and the alternative assumption. Consequently, the two modified measure are

$$Loc\ SG_i = \left(\frac{\sum_{intra=1}^{n} D_{intra}}{\sum_{all=1}^{n} D_{all}} \Big/ \frac{\sum_{intra=1}^{n} PT\ ALL_{intra}}{\sum_{all=1}^{n} PT\ ALL_{domestic}} \right) \left(1 + \frac{\sum_{b=1}^{n} Rel\ HB\ ALL_b}{n} \right) 100 \text{ and}$$

$$Loc\ SGHB_i = \left(\frac{\sum_{intra=1}^{n} D_{intra}}{\sum_{all=1}^{n} D_{all}} \Big/ \frac{\sum_{intra=1}^{n} PT\ SG_{intra}}{\sum_{all=1}^{n} PT\ SG_{domestic}} \right) \left(1 + \frac{\sum_{b=1}^{n} Rel\ HB\ SG_b}{n} \right) 100.$$

As the number of bidders differs per region (see **Appendix Table 1**) a threshold of 10 bidders per region is used. This leads to an exclusion of 5 COROP regions and 248 municipalities. Unfortunately, the effect of the number of bidders per region on localization effects is rather serious (see **Appendix Table 2**). There seems to be a logarithmic relationship (see **Appendix Figure 1**). As the correlation is least serious and less significant on a municipality level, the localization effect measure is most valid and reliable for this level.

Appendix Table 1 The number of bidders per region.

Number of bidders	Municipality		COROP region		Province	
Min	1	(10)	2	(10)	31	(31)
Max	168	(168)	261	(251)	371	(371)
Mean	6.4	(26.0)	46.4	(52.1)	154.6	(154.6)
SD	13.7	(28.3)	55.4	(57.0)	129.0	(129.0)
Number of regions	291	(43)	40	(35)	12	(12)

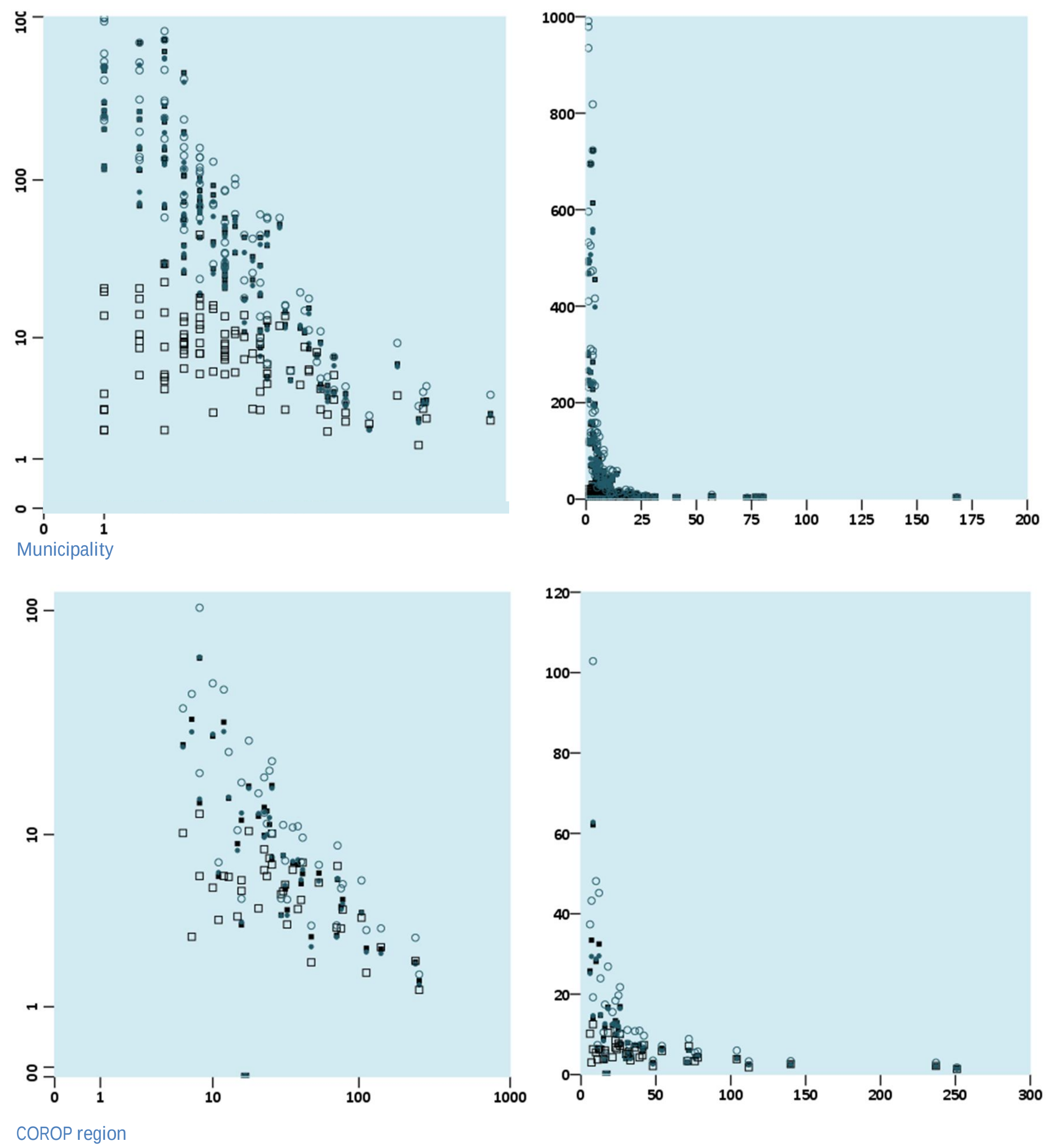
The numbers in parentheses indicate the numbers for the sample in which only regions with at least 10 bidders are selected.

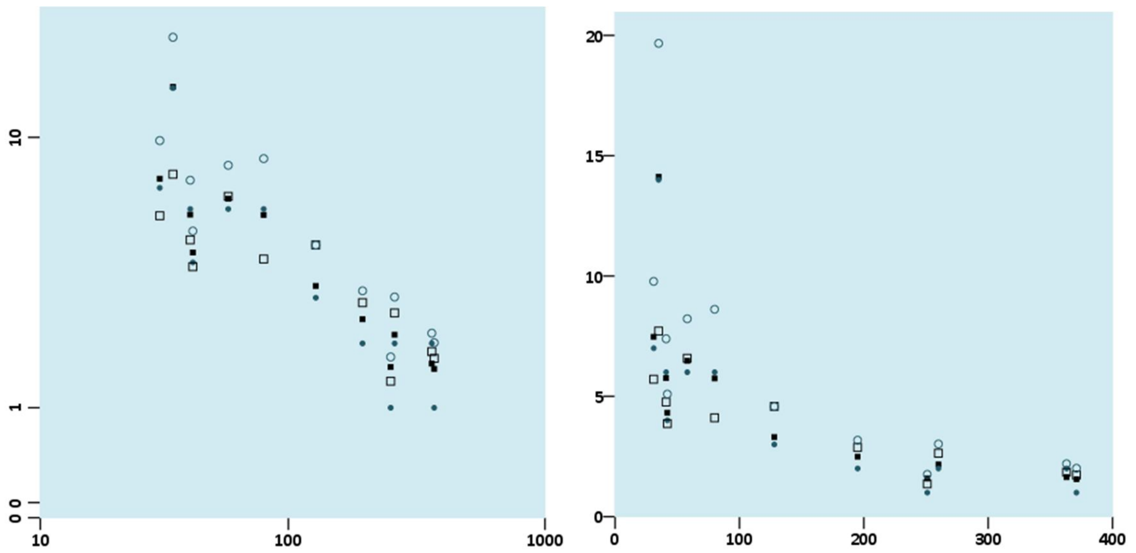
Appendix Table 2 The number of bidders and their impact on the localization effect measures.

Regions	Loc ALL	Loc SG	Loc ALL HB	Loc ALL SG
Province	-0.70***	-0.84***	-0.72***	-0.73***
COROP regions	-0.43***	-0.51***	-0.47***	-0.46***
Municipalities	-0.26*	-0.25*	-0.26*	-0.26*

Pearson correlation coefficients between number of bidders per region and agglomeration effects. Significance is 2-tailed.

Appendix Figure 1 **Relation between the number of bidders and localization effects.**





Province

Scatter plots of number of bidders (x-axis) and agglomeration effects (y-axis) for intra-regional deals with hollow circles as Agglo ALL, filled circles as Agglo ALLHB, hollow rectangles as Agglo SG and filled rectangles as Agglo SGHB.

Results

I test whether localization effects on average are larger than 1 for the level of provinces (**Appendix table 3**), COROP regions (**Appendix table 4**) and municipalities (**Appendix table 5**). In order to test whether localization effects are larger than 1 on average confidence intervals are calculated by comparing the means of many random samples. This method called bootstrapping is very useful for not normally distributed data, such as these localization effects. The largest and also most reliable effect is on the municipality level. The lowest lower bound of the confidence intervals is 3.6, which means that we can confirm the hypothesis. Localization effects are weaker on a COROP regional and provincial level, but note that these effects depended strongly on the number of bidders in those regions, i.e. those regions cannot be compared among each other. The map in **Appendix Figure 2** shows the strength of these effects for all 43 municipalities with at least 10 bidders.

Appendix table 3 Localization effects in provinces.

	Loc ALL			Loc SG			Loc ALLHB			Loc SGHB		
Min	2			1			1			1		
Max	20			8			14			14		
Mean	6.29	3.76	9.72	3.98	2.83	5.23	4.50	2.57	7.00	4.72	2.86	7.18
SD	5.06	2.20	7.65	2.00	1.16	2.56	3.68	1.68	5.45	3.62	1.57	5.40
Skewness	1.80	-0.47	2.42	0.44	-0.63	1.31	1.63	-0.60	2.26	1.70	-0.47	2.32
Kurtosis	4.00	-2.20	6.64	-0.64	-2.02	2.24	3.37	-2.25	5.93	3.53	-2.20	6.22

The 95% confidence intervals are each based on 5,000 bootstrap samples. N=12.

Appendix table 4 Localization effects in COROP regions.

	Loc ALL			Loc SG			Loc ALLHB			Loc SGHB		
Min	0			0			0			0		
Max	48			10			29			32		
Mean	11.77	8.63	15.45	5.10	4.29	5.93	8.29	6.30	10.56	8.45	6.32	10.84
SD	10.87	5.81	14.36	2.34	1.76	2.83	6.72	3.85	8.74	6.99	3.86	9.25
Skewness	2.06	0.64	2.85	0.15	-0.57	0.75	1.78	0.41	2.44	1.90	0.43	2.65
Kurtosis	4.64	-0.69	10.10	0.09	-0.91	1.15	3.63	-0.92	7.95	4.25	-0.94	9.16

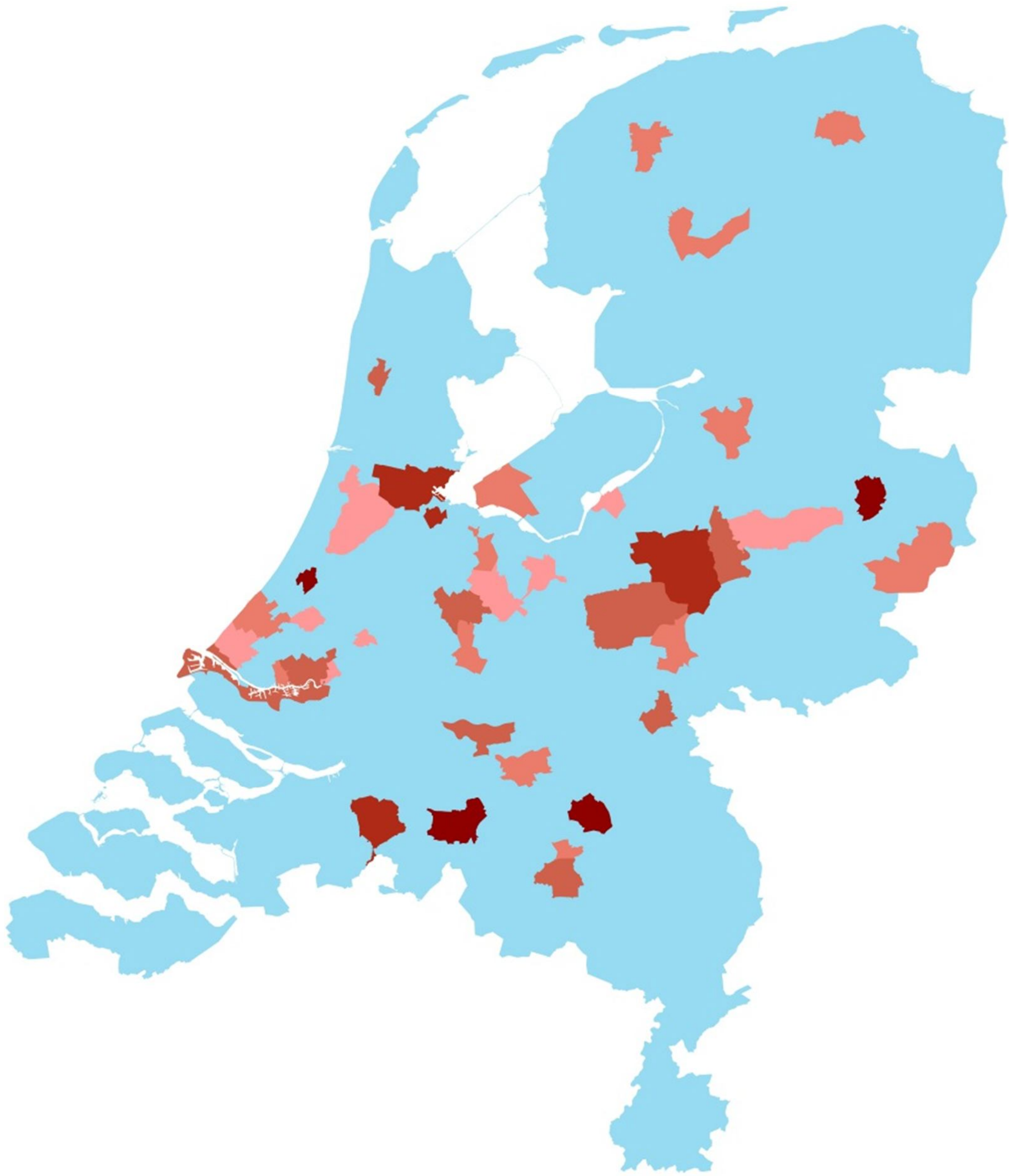
The 95% confidence intervals are each based on 5,000 bootstrap samples. N=35.

Appendix table 5 Localization effects in municipalities.

	Loc ALL			Loc SG			Loc ALLHB			Loc SGHB		
Min	0			0			0			0		
Max	61			14			51			53		
Mean	15.01	9.63	20.77	4.60	3.56	5.75	11.40	7.50	15.71	11.86	7.78	16.37
SD	18.07	11.70	22.19	3.58	2.65	4.28	13.30	8.53	16.55	13.87	8.74	17.33
Skewness	1.63	0.91	2.70	0.89	0.27	1.39	1.67	0.97	2.67	1.69	1.00	2.75
Kurtosis	1.54	-0.75	8.34	0.45	-0.88	2.42	1.96	-0.50	8.13	2.07	-0.41	8.50

The 95% confidence intervals are each based on 5,000 bootstrap samples. N=40.

Appendix Figure 2 Localization effects.



The map indicates the localization effects (Loc SG HB) in all munipalities with at least 10 bidders. The darker the red color the stronger the localization effect.

Additional research 2: Board interconnections in M&As

Aim

Aim of this research is to identify board interconnections between bidders and targets.

Method

I consider companies as socially related if bidder and target have shared a common board member prior to deal announcement. The REACH database contains not only executive directors but also employees in other leading positions. There is information on the function of every person. In the analyses I include inside and outside directors if the values of the function contain "*Board of Directors*", "*Raad van Bestuur*", "*Executive Board*" and "*Supervisory board*". Non-executive directors ("*Advisory board*", "*Dagelijks Bestuur*" and specific departments, such as "*Marketing*") are not included because they are not expected to influence M&A behaviour. Due to technical reasons I include not more than 56 persons per company. Whereas there was no information about all board members for 34 firms in one case there was information on more than 1000 persons. The data seems to be inconsistent.

Results

For only 9 out of 2113 deals a board interconnection was found. This number is not valid, however, as for some persons the application and resignation date was unknown. Thus, there might be more interconnections, which were just remained undetected due to missing data. Nonetheless, 5 of the 9 deals were taking place between companies located within the same city.

Appendix Table 6 Sensitivity test concerning the temporal scope of non-deal construction.

Independent variables	Model with full sample, *x=549 days, y=549 days*					Model with sample, *x=549 days, y=0 days*					Model with sample, *x=0 days, y=549days*				
	B	SE	Odds Ratio (95% CI)			B	SE	Odds Ratio (95% CI)			B	SE	Odds Ratio (95% CI)		
lnDistance	-0.12***	0.02	0.88	0.85	0.91	-0.18***	0.02	0.84	0.80	0.87	-0.14***	0.02	0.87	0.84	0.90
P	0.56***	0.08	1.75	1.51	2.03	0.51***	0.08	1.67	1.43	1.94	0.53***	0.08	1.70	1.46	1.98
C	1.05***	0.08	2.85	2.42	3.36	0.92***	0.09	2.52	2.12	3.00	1.03***	0.09	2.80	2.36	3.32
M	1.51***	0.11	4.54	3.65	5.64	1.29***	0.12	3.64	2.88	4.60	1.53***	0.12	4.64	3.68	5.84
Tanimoto	-3.78***	0.33	0.02	0.01	0.04	-4.03***	0.35	0.02	0.01	0.03	-3.55***	0.34	0.03	0.01	0.06
S	1.71***	0.08	5.52	4.76	6.40	1.70***	0.08	5.47	4.71	6.35	1.71***	0.08	5.54	4.77	6.43
D	2.60***	0.11	13.40	10.87	16.52	2.61***	0.11	13.55	10.98	16.72	2.57***	0.11	13.07	10.59	16.12
G	5.12***	0.09	168.12	142.16	198.82	5.17***	0.09	175.94	148.06	209.07	5.08***	0.09	161.39	136.18	191.27
UU	-0.24***	0.08	0.78	0.67	0.92	-0.22***	0.08	0.80	0.68	0.94	-0.26**	0.09	0.77	0.65	0.91
SiSi	-0.14***	0.05	0.87	0.78	0.96	-0.16***	0.05	0.85	0.77	0.95	-0.11**	0.05	0.89	0.80	0.99
MuMu	-0.18**	0.08	0.84	0.71	0.98	-0.18**	0.08	0.83	0.71	0.98	-0.20**	0.08	0.82	0.69	0.96
UnUn	-0.25***	0.05	0.78	0.70	0.87	-0.22***	0.06	0.80	0.72	0.89	-0.26***	0.06	0.77	0.69	0.86
DiDi	-0.51***	0.08	0.60	0.51	0.71	-0.51***	0.08	0.60	0.51	0.71	-0.51***	0.08	0.60	0.51	0.70
Intercept	-3.60***	0.14	0.03			-2.68***	0.14	0.07			-2.91***	0.05	0.05		
Deals	1,855					1,855					1,855				
Non-deals	1,607,008					753,935					853,073				
Omnibus test, χ² (df)	6597.88*** (13)					6557.40*** (13)					6571.74*** (13)				
Hosmer and Lemeshow test, χ² (df)	33.26*** (8)					29.41*** (8)					21.85*** (8)				
Dispersion Φ	4.16					3.68					2.73				
R^2_L	0.771					0.748					0.752				
R^2_{CS}	0.004					0.009					0.008				
R^2_N	0.231					0.255					0.251				

Appendix Table 7 The choice of the right endogenously stratified sample.

Independent variables	Sample (N=3,710)					Sample (N=11,755)					Full sample (N=1,608,863)				
	B	SE	Odds Ratio (95% CI)			B	SE	Odds Ratio (95% CI)			B	SE	Odds Ratio (95% CI)		
lnDistance	-0.47***	0.07	0.62	0.54	0.72	-0.50***	0.06	0.61	0.54	0.68	-0.12***	0.02	0.88	0.85	0.91
P	0.18	0.17	1.20	0.87	1.67	0.26**	0.12	1.30	1.02	1.65	0.56***	0.08	1.75	1.51	2.03
C	0.42*	0.23	1.52	0.97	2.39	0.38**	0.16	1.46	1.06	2.00	1.05***	0.08	2.85	2.42	3.36
M	0.59*	0.33	1.80	0.95	3.44	0.50**	0.24	1.65	1.03	2.65	1.51***	0.11	4.54	3.65	5.64
Tanimoto	1.39	0.99	4.00	0.58	27.84	-0.50	0.58	0.61	0.20	1.89	-3.78***	0.33	0.02	0.01	0.04
S	1.61***	0.13	5.01	3.90	6.43	1.66***	0.09	5.26	4.41	6.28	1.71***	0.08	5.52	4.76	6.40
D	2.24***	0.22	9.40	6.09	14.51	2.50***	0.15	12.21	9.16	16.27	2.60***	0.11	13.40	10.87	16.52
G	4.20***	0.25	66.42	40.94	107.74	4.55***	0.15	94.97	71.00	127.03	5.12***	0.09	168.12	142.16	198.82
UU	0.17	0.15	1.18	0.87	1.60	-0.07	0.12	0.94	0.74	1.19	-0.24***	0.08	0.78	0.67	0.92
SiSi	-0.06	0.10	0.94	0.77	1.14	-0.04	0.08	0.96	0.82	1.11	-0.14***	0.05	0.87	0.78	0.96
MuMu	-0.38**	0.15	0.69	0.51	0.92	-0.34***	0.12	0.71	0.56	0.89	-0.18**	0.08	0.84	0.71	0.98
UnUn	0.09	0.10	1.10	0.91	1.33	0.03	0.08	1.03	0.89	1.20	-0.25***	0.05	0.78	0.70	0.87
DiDi	-0.70***	0.16	0.50	0.36	0.69	-0.53***	0.12	0.59	0.46	0.75	-0.51***	0.08	0.60	0.51	0.71
Intercept	3.91***	0.35	49.69			2.35***	0.25	10.43			-6.73***	0.15	0.00		
Deals	1,855					1,855					1,855				
Non-deals	1,855					9,900					1,607,008				
Omnibus test, χ² (df)	2039.83*** (13)					4173.38*** (13)					6597.88*** (13)				
Hosmer and Lemeshow test, χ² (df)	3.44 (8)					8.38 (8)					33.26*** (8)				
Dispersion Φ	0.40					1.05					4.16				
R_L^2	0.607					0.597					0.771				
R_{CS}^2	0.423					0.299					0.004				
R_N^2	0.564					0.514					0.231				

Appendix Table 8 ReLogit correction of the endogenously stratified and the full sample.

Independent variables	Sample (N=11,755) Corrected with ReLogit					Full Sample (N=1,608,863) Corrected with ReLogit				
	B	Robust SE	Relative Risk (95% CI)			B	Robust SE	Relative Risk (95% CI)		
lnDistance	-0.50***	0.06	0.61	0.54	0.68	-0.12***	0.02	0.88	0.85	0.92
P	0.26**	0.12	1.31	1.03	1.66	0.56***	0.08	1.75	1.51	2.02
C	0.38**	0.16	1.45	1.07	1.98	1.05***	0.08	2.84	2.42	3.38
M	0.51**	0.25	1.67	1.02	2.68	1.51***	0.11	4.56	3.66	5.62
Tanimoto	-0.49	0.72	1.00	1.00	1.00	-3.77***	0.41	0.96	0.96	0.97
S	1.66***	0.09	5.22	4.37	6.17	1.71***	0.08	5.53	4.75	6.39
D	2.50***	0.14	12.10	9.10	15.97	2.60***	0.11	13.35	10.73	16.14
G	4.54***	0.16	89.40	65.62	117.48	5.12***	0.10	150	120	179
UU	-0.06	0.13	0.94	0.71	1.23	-0.24***	0.09	0.78	0.66	0.93
SiSi	-0.04	0.07	0.95	0.83	1.11	-0.14***	0.05	0.87	0.78	0.96
MuMu	-0.34***	0.12	0.71	0.56	0.89	-0.18**	0.09	0.83	0.71	0.99
UnUn	0.03	0.08	1.03	0.88	1.21	-0.25***	0.06	0.78	0.70	0.87
DiDi	-0.53***	0.13	0.59	0.44	0.76	-0.51***	0.08	0.60	0.51	0.72
Intercept	-5.21***	0.32				-6.73***	0.15			
Deals	1,855					1,855				
Non-deals	9,900					1,607,008				

For the estimations of relative risks lnDistance and Tanimoto are hold at the mean and all others, i.e. all binary variables, at 0. The relative risk indicates deal likelihood if the independent variable is increased by 1.

Appendix Table 9 Logistic regression model estimations without multiple bidders.

Independent variables	Sample (N=11,755)					Sample without multiple bidders (N=7,343)				
	B	SE	Relative Risk (95% CI)			B	SE	Relative Risk (95% CI)		
lnDistance	-0.50***	0.06	0.61	0.54	0.68	-0.61***	0.07	0.55	0.47	0.63
P	0.26**	0.12	1.30	1.02	1.65	0.21	0.16	1.23	0.90	1.68
C	0.38**	0.16	1.46	1.06	2.00	0.34*	0.21	1.41	0.94	2.12
M	0.50**	0.24	1.65	1.03	2.65	0.31	0.31	1.36	0.75	2.50
Tanimoto	-0.50	0.58	0.61	0.20	1.89	0.89	0.80	2.45	0.50	11.85
S	1.66***	0.09	5.26	4.41	6.28	1.52***	0.12	4.56	3.59	5.79
D	2.50***	0.15	12.21	9.16	16.27	2.53***	0.17	12.54	8.91	17.65
G	4.55***	0.15	94.97	71.00	127.03	4.59***	0.19	98.08	66.96	143.66
UU	-0.07	0.12	0.94	0.74	1.19	-0.52**	0.24	0.59	0.37	0.96
SiSi	-0.04	0.08	0.96	0.82	1.11	0.01	0.10	1.01	0.84	1.22
MuMu	-0.34***	0.12	0.71	0.56	0.89	0.40***	0.16	1.49	1.09	2.04
UnUn	0.03	0.08	1.03	0.89	1.20	0.03	0.10	1.04	0.85	1.26
DiDi	-0.53***	0.12	0.59	0.46	0.75	0.64***	0.17	1.90	1.35	2.66
Intercept	2.35***	0.25	10.43			-0.11	0.40	0.90		
Deals	1,855					1,162				
Non-deals	9,900					6,181				
Omnibus test, χ^2 (df)	4173.38*** (13)					2815.246*** (13)				
Hosmer and Lemeshow test, χ^2 (df)	8.38 (8)					6.891 (8)				
Dispersion Φ	1.05					0.86				
R_L^2	0.597					0.553				
R_{CS}^2	0.299					0.318				
R_N^2	0.514					0.547				

For the estimations of relative risks lnDistance and Tanimoto are hold at the mean and all others, i.e. all binary variables, at 0. The relative risk indicates deal likelihood if the independent variable is increased by 1.

8 Empirical results
8.1 Home bias evidence

Appendix Table 10 Home bias measured by SPCM.

PCM	Deals	Assumption ALL			Assumption SG		
		Non-deals	OR	RR	Non-deals	OR	RR
S	64.8%	86.9%	0.28	0.75	85.0%	0.33	0.76
P	12.1%	7.7%	1.64	1.57	7.9%	1.61	1.54
C	10.6%	3.8%	2.98	2.77	4.5%	2.55	2.38
M	12.5%	1.6%	8.94	7.94	2.7%	5.16	4.64
Deals	1,855	1,607,008			35,067		

The most common test for significance between two categorical data, Pearson's Chi-Square Test, reveals that there is a highly significant association between Deal and PCM; χ^2 (3) = 1766.29, p < 0.001 under assumption ALL; χ^2 (3) = 788.73, p < 0.001 under assumption SG.

Appendix Table 11 Home bias measured in distance mean differences (ALL).

Regional scale	Deals	Non-deals	Home bias (km)
Netherlands	60.85	86.12	25.27
Same province	15.16	24.38	9.22
Same COROP region	7.56	12.83	5.27
Same municipality	2.90	4.58	1.68

Appendix Table 12 Home bias measured in distance mean differences (SG).

Regional scale	Deals	Non-deals	Home bias (km)
Netherlands	60.85	78.32	17.47
Same province	15.16	20.85	5.69
Same COROP region	7.56	10.96	3.40
Same municipality	2.90	3.83	0.93

Appendix Table 13 Home bias measured by distance median differences (ALL).

Regional scale	Deals	Non-deals	Difference	Home bias (km) (95% CI)		
Netherlands	49.61	79.62	30.01	27.93***	25.54	30.32
Same province	10.97	21.50	10.53	11.23***	8.55	13.91
Same COROP region	5.42	10.62	5.20	4.37***	3.46	5.30
Same municipality	1.76	4.10	2.34	1.73***	1.39	2.08

The 95% interval is calculated by the non-parametric Independent-Samples Hodges-Lehman Median Difference. Due to limited computational power the difference for all deals in the Netherlands, the same provinces and COROP regions had to be calculated on a sample. The significance level is calculated by the Independent-Samples Median Test. N=11,756.

Appendix Table 14 Home bias measured by distance median differences (SG).

Regional scale	Deals	Non-deals	Difference	Home bias (km) (95% CI)		
Netherlands	49.61	69.00	19.39	19.85***	12.68	26.85
Same province	10.97	18.89	7.92	5.73***	4.66	6.84
Same COROP region	5.42	8.27	2.85	2.65***	1.96	3.33
Same municipality	1.76	3.49	1.73	0.94***	0.56	1.37

The estimate and its 95% interval is calculated by the non-parametric Independent-Samples Hodges-Lehman Median Difference. Due to limited computational power the difference for all deals in the Netherlands had to be calculated on a sample.
The significance level is calculated by the Independent-Samples Median Test. N=11,756.

Appendix Table 15 The joint effect of PCM and Distance.

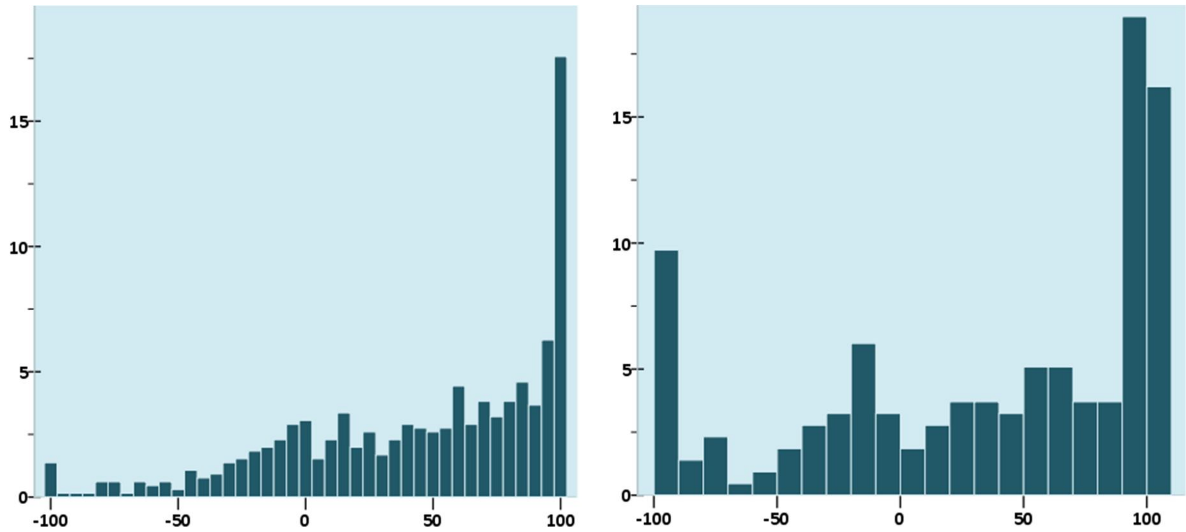
Independent variables	Model A					Model B					Model C				
	B	SE	Odds Ratio (95% CI)			B	SE	Odds Ratio (95% CI)			B	SE	Odds Ratio (95% CI)		
P						0.78***	0.11	2.19	1.77	2.71	0.49***	0.12	1.63	1.29	2.07
C						1.30***	0.12	3.67	2.88	4.69	0.93***	0.14	2.53	1.92	3.33
M						2.35***	0.14	10.44	7.97	13.67	1.92***	0.16	6.80	5.01	9.23
Distance	-0.011***	0.001	0.989	0.988	0.991						-0.005***	0.001	0.995	0.993	0.997
S	1.67***	0.09	5.33	4.48	6.32	1.64***	0.09	5.16	4.33	6.14	1.65***	0.09	5.21	4.37	6.21
D	2.54***	0.14	12.64	9.56	16.72	2.51***	0.14	12.33	9.30	16.35	2.52***	0.15	12.43	9.35	16.51
G	4.54***	0.15	93.28	70.09	124.14	4.51***	0.15	91.26	68.39	121.79	4.53***	0.15	92.80	69.48	123.95
Tanimoto	-0.40	0.57	0.67	0.22	2.04	-0.27	0.58	0.76	0.25	2.35	-0.40	0.58	0.67	0.22	2.07
UU	-0.05	0.12	0.95	0.75	1.20	-0.02	0.12	0.98	0.78	1.24	-0.04	0.12	0.96	0.76	1.22
SiSi	-0.07	0.08	0.93	0.81	1.08	-0.03	0.08	0.97	0.83	1.12	-0.05	0.08	0.95	0.82	1.11
MuMu	-0.30**	0.12	0.74	0.59	0.92	-0.34***	0.12	0.71	0.57	0.89	-0.33***	0.12	0.72	0.57	0.90
UnUn	0.02	0.07	1.02	0.88	1.17	0.02*	0.08	1.03	0.88	1.19	0.03*	0.08	1.03	0.89	1.19
DiDi	-0.51***	0.12	0.60	0.47	0.76	-0.54***	0.12	0.58	0.46	0.74	-0.54***	0.12	0.58	0.46	0.74
Intercept	0.95***	0.20	2.59			0.98***	0.20	2.67			1.17***	0.20	3.21		
Deals	1,855					1,855					1,855				
Non-deals	9,900					9,900					9,900				
Omnibus test, χ² (df)	4074.37*** (12)					4074.37*** (12)					4106.97*** (13)				
Hosmer and Lemeshow test, χ² (df)	60.44*** (8)					58.19 (8)					6.06 (8)				
Dispersion Φ	7.56					7.28					0.76				
R^2_L	0.615					0.603					0.599				
R^2_{CS}	0.285					0.293					0.295				
R^2_N	0.490					0.503					0.507				

Appendix Table 16 The effect of (Distance-mean)² and LnDistance.

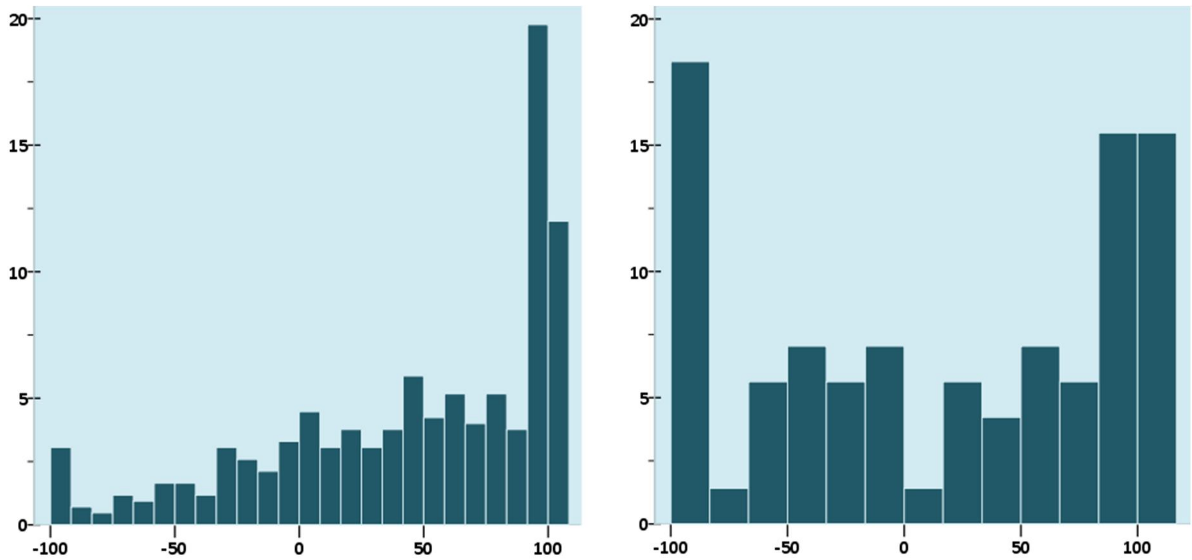
Independent variables	Model D					Model E					Model F				
	B	SE	Odds Ratio (95% CI)			B	SE	Odds Ratio (95% CI)			B	SE	Odds Ratio (95% CI)		
P											0.26**	0.12	1.30	1.02	1.65
C											0.38**	0.16	1.46	1.06	2.00
M											0.50**	0.24	1.65	1.03	2.65
Distance	-.012***	.001	.988	.987	.989										
(Distance-mean)²	.000***	.000	1.00010	1.00009	1.00012										
lnDistance						-0.62***	.031	.540	.508	.574	-0.50***	0.06	0.61	0.54	0.68
S	1.66***	0.09	5.28	4.44	6.29	1.66***	0.09	5.28	4.42	6.30	1.66***	0.09	5.26	4.41	6.28
D	2.50***	0.14	12.15	9.15	16.14	2.50***	0.15	12.23	9.18	16.29	2.50***	0.15	12.21	9.16	16.27
G	4.55***	0.15	94.46	70.85	125.94	4.56***	0.15	95.74	71.59	128.03	4.55***	0.15	94.97	71.00	127.03
Tanimoto	-0.51	0.57	0.60	0.20	1.82	-0.54	0.58	0.58	0.19	1.80	-0.50	0.58	0.61	0.20	1.89
UU	-0.07	0.12	0.94	0.74	1.18	-0.07	0.12	0.93	0.74	1.18	-0.07	0.12	0.94	0.74	1.19
SiSi	-0.05	0.08	0.95	0.82	1.10	-0.05	0.08	0.95	0.82	1.11	-0.04	0.08	0.96	0.82	1.11
MuMu	-0.31**	0.12	0.74	0.59	0.92	-0.34***	0.12	0.71	0.57	0.90	-0.34***	0.12	0.71	0.56	0.89
UnUn	0.02	0.08	1.02	0.88	1.18	0.03	0.08	1.03	0.89	1.19	0.03	0.08	1.03	0.89	1.20
DiDi	-0.53***	0.12	0.59	0.46	0.75	-0.53***	0.12	0.59	0.46	0.75	-0.53***	0.12	0.59	0.46	0.75
Intercept	0.76***	0.20	2.14			2.60***	0.23	13.52			2.35***	0.25	10.43		
Deals	1,855					1,855					1,855				
Non-deals	9,900					9,900					9,900				
Omnibus test, χ² (df)	4063.90*** (11)					1995.80*** (10)					4173.38*** (13)				
Hosmer and Lemeshow test, χ² (df)	23.46*** (8)					10.95** (8)					8.38 (8)				
Dispersion Φ	2.94					1.37					1.05				
R^2_L	0.604					0.594					0.597				
R^2_{CS}	0.292					0.298					0.299				
R^2_N	0.502					0.513					0.514				

8.1.2 Individual home bias

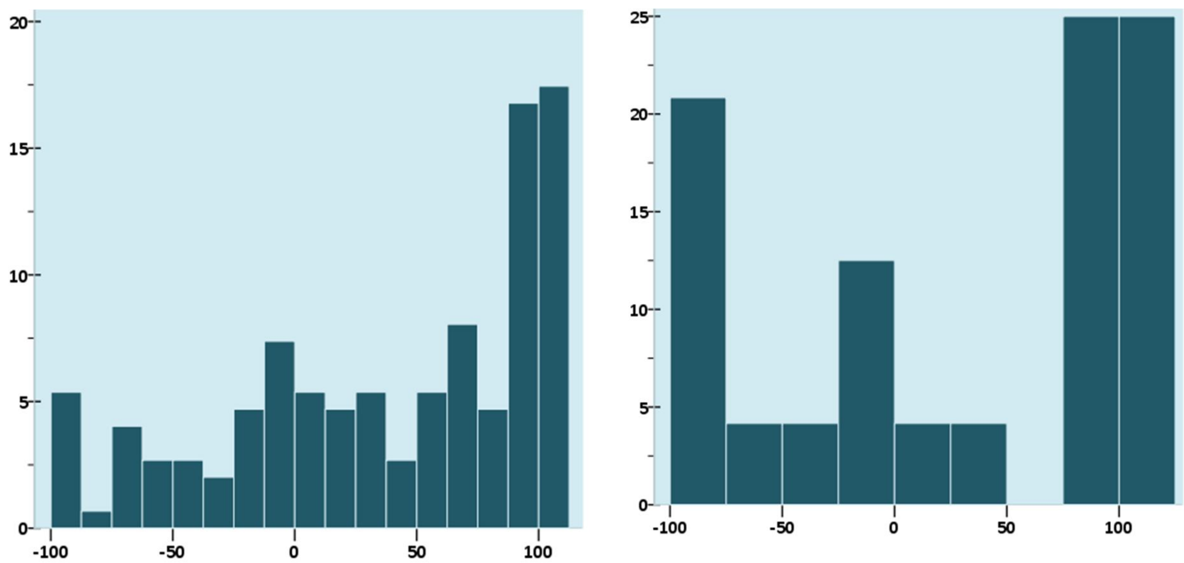
Appendix Figure 3 Distribution of individual HB for bidders involved in intra-regional deals.



Histogram (%) of home bias under assumption ALL (left) (N=654) and SG (right) (N=216) for deals within the same province.



Histogram (%) of home bias under assumption ALL (left) (N=425) and SG (right) (N=71) for deals within the same COROP region.



Histogram (%) of home bias under assumption ALL (left) (N=149) and SG (right) (N=24) for deals within the same municipality.

Appendix Table 17 Absolute home bias for deals on different regional scales (ALL).

Absolute home bias (%)	Domestic			Same province			Same COROP			Same municipality		
Min	-165.0			-59.1			-20.4			-8.1		
Max	159.8			66.9			28.1			6.5		
Mean (95% CI)	25.50	23.01	27.94	10.25	8.98	11.51	5.19	4.45	5.95	1.02	0.62	1.38
SD	51.74	49.84	53.40	15.83	14.65	16.91	7.86	7.29	8.46	2.53	2.16	2.91
Skewness	-0.46	-0.57	-0.37	-0.49	-0.82	-0.15	0.21	-0.04	0.45	-0.89	-1.25	-0.33
Kurtosis	0.20	-0.01	0.43	1.53	0.72	2.35	0.48	0.04	0.96	1.79	0.43	2.94
N	1,855			654			428			166		

Only cases with at least 5 potential targets are included. Bootstrapping is based on 1,000 samples.

Appendix Table 18 Relative home bias for deals on different regional scales (ALL).

Relative home bias (%)	Domestic			Same province			Same COROP			Same municipality		
Mean (95% CI)	33.7	31.4	36.2	44.8	40.7	48.9	43.7	38.7	49.1	38.67	29.17	47.30
SD	49.8	48.5	51.1	50.6	47.9	53.2	56.3	52.6	59.9	61.95	56.11	67.10
Skewness	-0.42	-0.49	-0.35	-0.80	-0.94	-0.65	-0.88	-1.04	-0.73	-0.76	-1.02	-0.51
Kurtosis	-0.76	-0.86	-0.64	-0.16	-0.48	0.17	-0.16	-0.52	0.26	-0.58	-1.02	0.00
Proportion of bidders with positive home bias in % (95% CI)	72	70	74	78	72	82	77	73	81	72	65	79
N	1,855			654			428			166		

Only cases with at least 5 potential targets are included. Bootstrapping is based on 1,000 samples.

8.2 Industrial relatedness bias

Appendix Table 19 Industrial relatedness bias measured by USDG.

SDG	Deals	Non-deals	Odds ratio	Relative risk
U	28.2%	88.3%	0.05	0.31
S	14.1%	8.3%	1.82	1.70
D	5.8%	1.4%	4.23	4.05
G	51.9%	2.0%	53.22	26.14
No. of deals	1,855	1,607,008		

Appendix Table 20 Multivariate analyses of Tanimoto and USDG.

Independent variables	Model I					Model J					Model K				
	B	SE	Odds Ratio (95% CI)			B	SE	Odds Ratio (95% CI)			B	SE	Odds Ratio (95% CI)		
P	0.27**	0.11	1.32	1.06	1.63	0.26**	0.12	1.30	1.02	1.66	0.26**	0.12	1.30	1.02	1.65
C	0.39**	0.14	1.47	1.11	1.95	0.38**	0.16	1.46	1.06	2.01	0.38**	0.16	1.46	1.06	2.00
M	0.54**	0.22	1.72	1.13	2.62	0.51**	0.24	1.66	1.04	2.67	0.5**	0.24	1.65	1.03	2.65
LnDistance	-0.46***	0.05	0.63	0.57	0.70	-0.50***	0.06	0.61	0.55	0.68	-0.50***	0.06	0.61	0.54	0.68
S						1.65***	0.09	5.19	4.36	6.17	1.66***	0.09	5.26	4.41	6.28
D						2.49***	0.15	12.04	9.05	16.02	2.50***	0.15	12.21	9.16	16.27
G						4.46***	0.10	86.27	71.40	104.24	4.55***	0.15	94.97	71.00	127.03
Tanimoto	18.74***	0.54	137*106	47*106	395*106						-0.50	0.58	0.61	0.20	1.89
UU	-0.31**	0.12	0.74	0.58	0.93	-0.08	0.12	0.93	0.73	1.17	-0.07	0.12	0.94	0.74	1.19
SiSi	-0.10	0.07	0.90	0.79	1.03	-0.04	0.08	0.96	0.82	1.11	-0.04	0.08	0.96	0.82	1.11
MuMu	-0.34***	0.11	0.71	0.58	0.87	-0.34***	0.12	0.71	0.56	0.89	-0.34***	0.12	0.71	0.56	0.89
UnUn	-0.17**	0.07	0.85	0.74	0.97	0.03	0.08	1.03	0.89	1.19	0.03	0.08	1.03	0.89	1.2
DiDi	-0.33**	0.12	0.72	0.57	0.92	-0.51***	0.12	0.60	0.47	0.76	-0.53***	0.12	0.59	0.46	0.75
Intercept	0.47**	0.23	1.61			2.29***	0.25	9.90			2.35***	0.25	10.43		
Deals	1,855					1,855					1,855				
Non-deals	9,900					9,900					9,900				
Omnibus test, χ^2 (df)	7217.32*** (10)					4172.63*** (12)					4173.38*** (13)				
Hosmer and Lemeshow test, χ^2 (df)	138.19*** (8)					7.94 (8)					8.38 (8)				
Dispersion Φ	17.25					0.99					1.05				
R^2_L	0.704					0.593					0.597				
R^2_{CS}	0.227					0.299					0.299				
R^2_N	0.391					0.513					0.514				

Appendix Table 21 Industrial relatedness bias measured by Tanimoto mean differences.

Regional scale	Deals	Non-deals	IR bias (Tanimoto)
Netherlands	0.111	0.008	-0.103
Same province	0.102	0.009	-0.093
Same COROP region	0.108	0.010	-0.098
Same municipality	0.101	0.013	-0.088

Appendix Table 22 Industrial relatedness bias measured by Tanimoto median differences.

Regional scale	Deals	Non-deals	Difference	IR bias (Tanimoto) (95% CI)		
Netherlands	0.0969	0.0002	0.0967	0.080	0.076	0.094
Same province	0.0879	0.0002	0.0877	0.087	0.076	0.097
Same COROP region	0.0976	0.0002	0.0974	0.082	0.075	0.095
Same municipality	0.0768	0.0003	0.0765	0.082	0.076	0.096

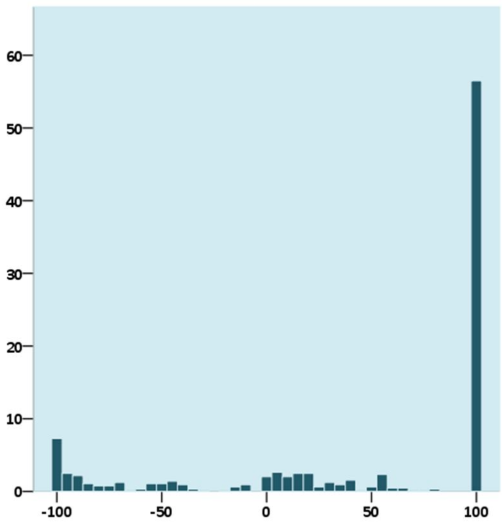
The 95% interval is calculated by the non-parametric Independent-Samples Hodges-Lehman Median Difference. Due to limited computational power the difference for all deals in the Netherlands, the same provinces and COROP regions had to be calculated on a sample. N=11,756.

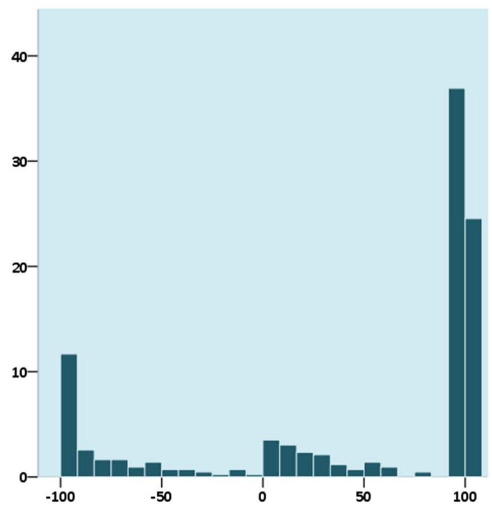
Appendix Figure 4 Industrial relatedness bias for bidders involved in intra-regional deals.

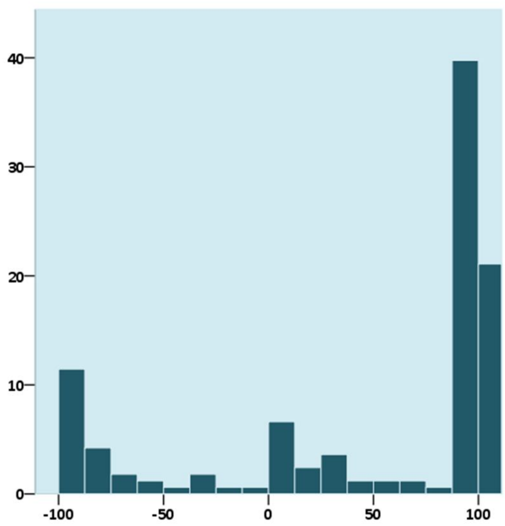
Top right:
Histogram (%) of industrial relatedness bias for deals within the same province (N=654).

Bottom left:
Histogram (%) of industrial relatedness bias for deals within the same COROP region (N=425).

Bottom right:
Histogram (%) of industrial relatedness bias for deals within the same municipality (N=149).







Appendix Table 23 Absolute industrial relatedness bias for deals on different regional scales.

Abs. IR bias (%)	Domestic			Same province			Same COROP			Same municipality		
Min	-0.065			-0.062			-0.036			-0.031		
Max	0.614			0.615			0.608			0.558		
Mean (95% CI)	0.099	0.095	0.093	0.095	0.087	0.104	0.093	0.084	0.103	0.08	0.073	0.104
SD	0.112	0.102	0.101	0.102	0.092	0.111	0.101	0.089	0.112	0.10	0.081	0.120
Skewness	1.40	1.40	1.46	1.40	0.98	1.71	1.46	0.89	1.83	1.67	0.65	2.09
Kurtosis	2.77	3.45	3.92	3.45	1.29	4.94	3.92	1.12	5.74	4.80	-0.19	7.01
N	1,855			654			428			166		
Only cases with at least 5 potential targets are included. Bootstrapping is based on 1,000 samples.												

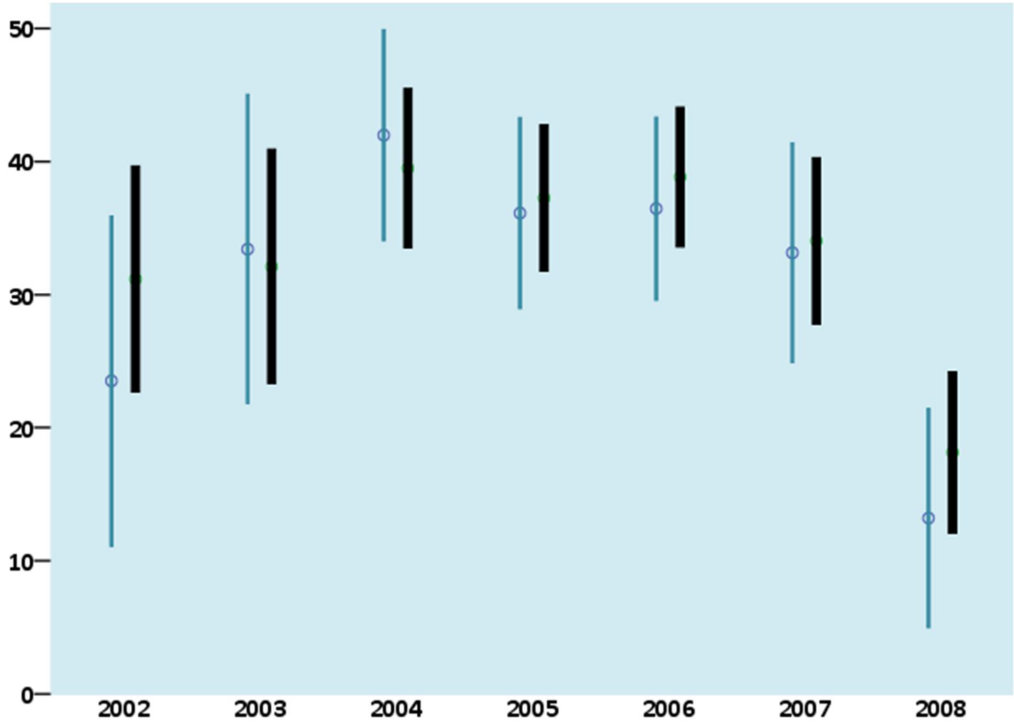
8.3 M&A formation model

Appendix Table 24 Cross-validation of the logistic regression model.

Independent variables	Sample (80%)					Sample (20%)				
	B	Robust SE	Relative Risk (95% CI)			B	Robust SE	Relative Risk (95% CI)		
lnDistance	-0.47***	0.06	0.63	0.55	0.71	-0.64***	0.13	0.53	0.41	0.68
P	0.29	0.14	1.33	1.02	1.75	0.14	0.29	1.15	0.65	2.03
C	0.40	0.18	1.50	1.05	2.14	0.24	0.37	1.27	0.62	2.59
M	0.80***	0.27	2.22	1.31	3.76	-0.66	0.56	0.52	0.17	1.56
Tanimoto	-0.55	0.63	0.58	0.17	1.98	-0.32	1.45	0.73	0.04	12.48
S	1.66***	0.10	5.23	4.28	6.40	1.70***	0.19	5.48	3.76	7.99
D	2.60***	0.16	13.48	9.81	18.52	2.08***	0.35	8.01	4.02	15.93
G	4.55***	0.16	94.95	68.84	130.96	4.60***	0.36	99.70	49.68	200.08
UU	-0.06	0.13	0.94	0.73	1.23	-0.09	0.27	0.91	0.53	1.56
SiSi	0.01	0.09	1.01	0.85	1.19	-0.25	0.17	0.78	0.56	1.09
MuMu	-0.28**	0.13	0.76	0.59	0.98	-0.59**	0.25	0.55	0.33	0.91
UnUn	0.02	0.08	1.02	0.86	1.20	0.10	0.17	1.10	0.79	1.54
DiDi	-0.60***	0.14	0.55	0.42	0.72	-0.22	0.29	0.80	0.45	1.41
Intercept	2.30***	0.28	10.02			2.54***	0.58	12.67		
Deals	1,506					234				
Non-deals	7,898					2,002				
Omnibus test, χ^2 (df)	3414.97***(13)					770.85*** (13)				
Hosmer and Lemeshow test, χ^2 (df)	8.73 (8)					17.34**(8)				
Dispersion Φ	1.09					2.16				
R^2_L	0.587					0.610				
R^2_{CS}	0.305					0.280				
R^2_N	0.520					0.492				

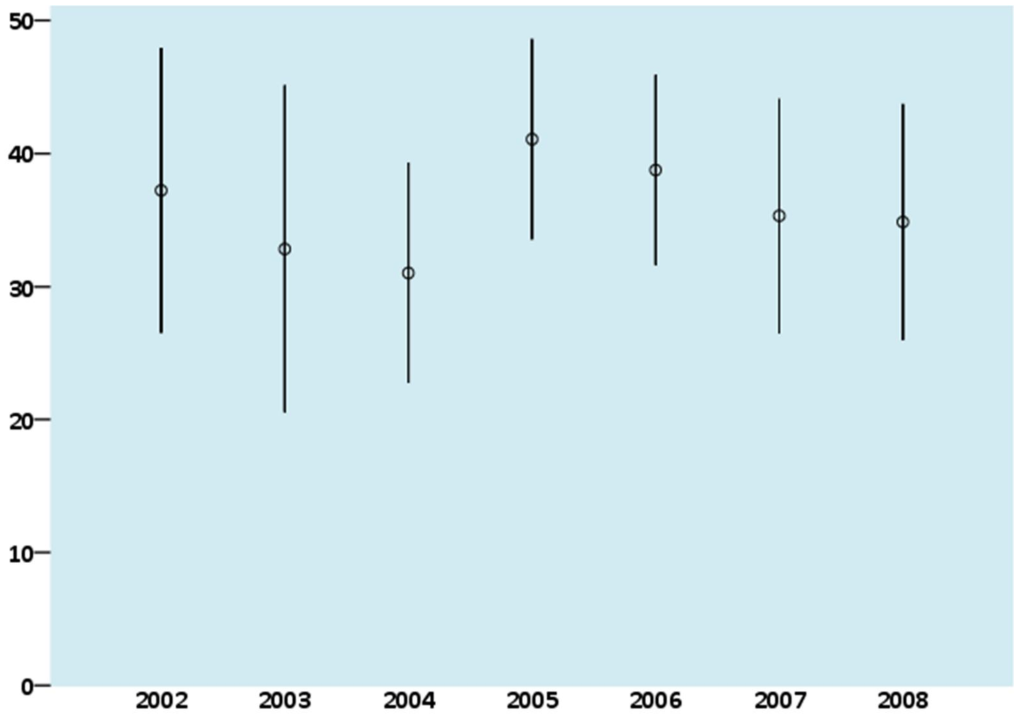
9 Discussion of results

Appendix Figure 5 Temporal differences in domestic home bias.



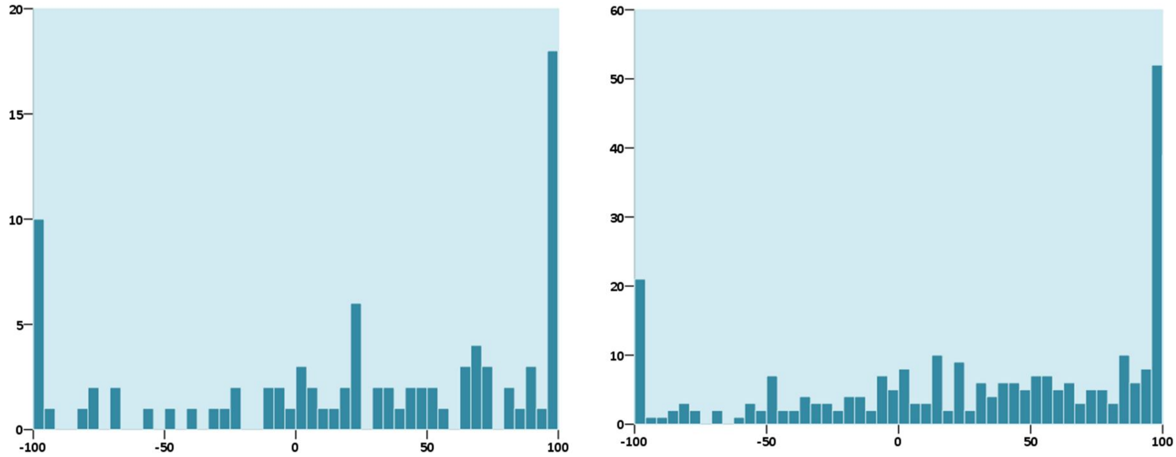
95% confidence intervals for relative domestic home bias under assumption ALL (grey) and SG (black) per year, N=1,571.

Appendix Figure 6 Temporal differences in industrial relatedness bias.

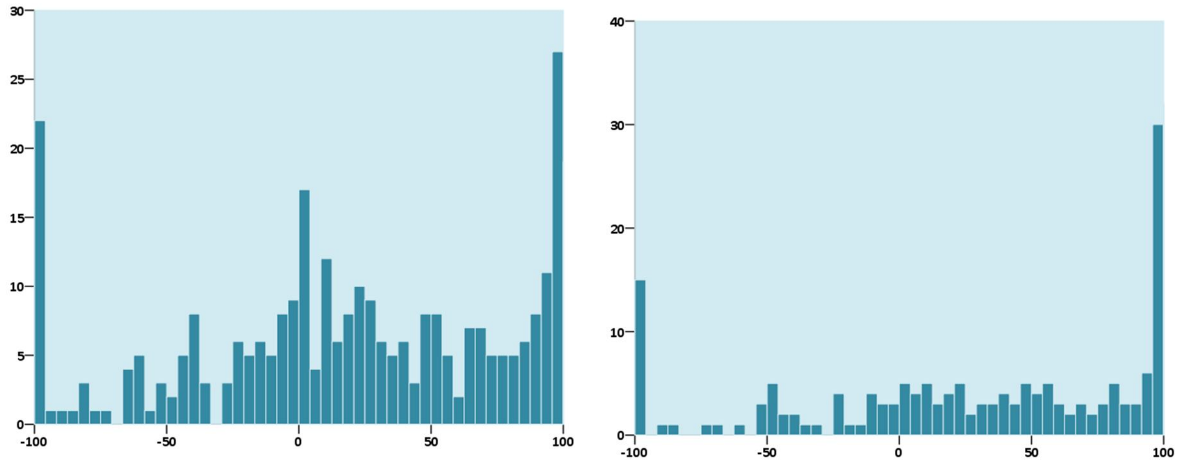


95% confidence intervals for relative domestic industrial relatedness bias under, N=1,855.

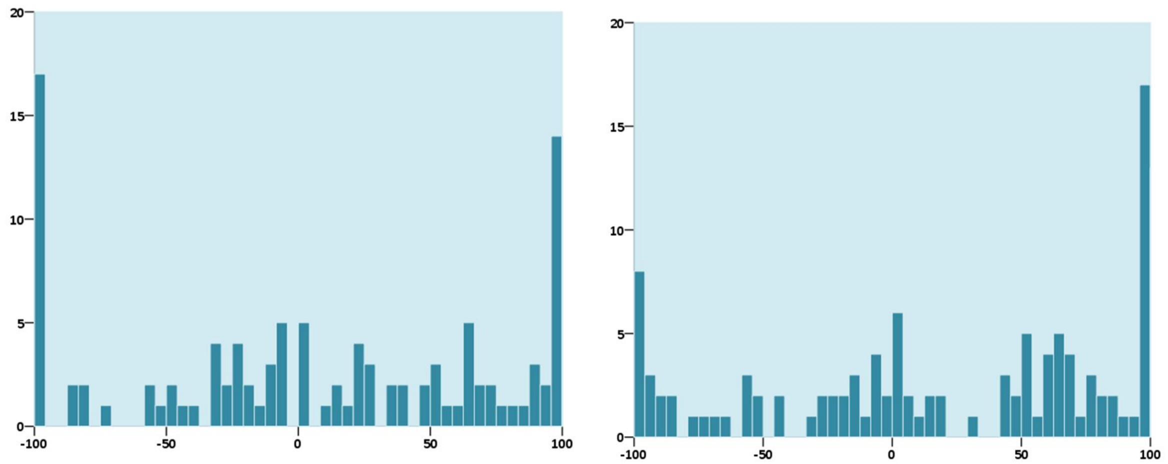
Appendix Figure 7 **Individual home bias for bidders of different industries.**



Histogram (%) of home bias under assumption SG for F: Construction (left) and G: Wholesale and retail trade (right).

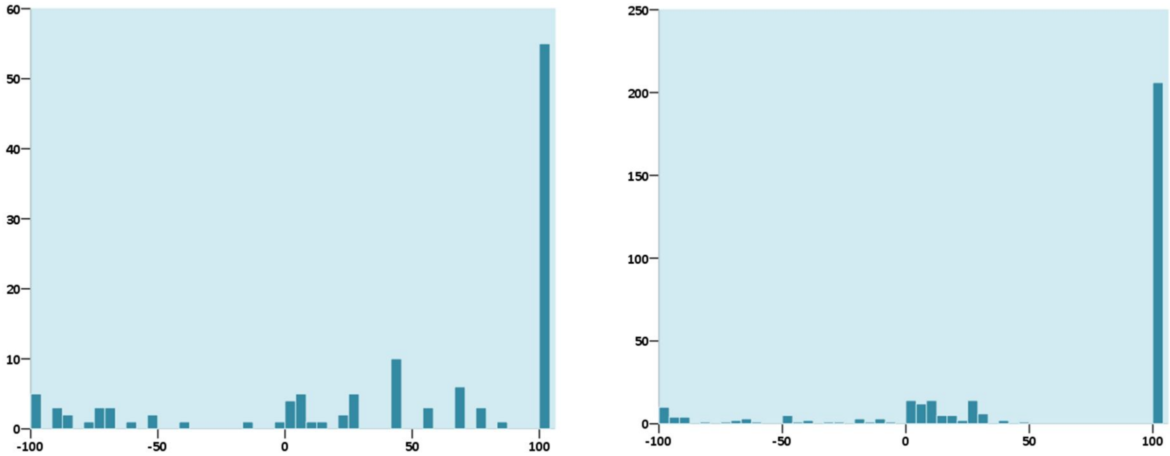


Histogram (%) of home bias under assumption SG for J: Information and Communication (left) and K: Financial and Insurance activities (right).

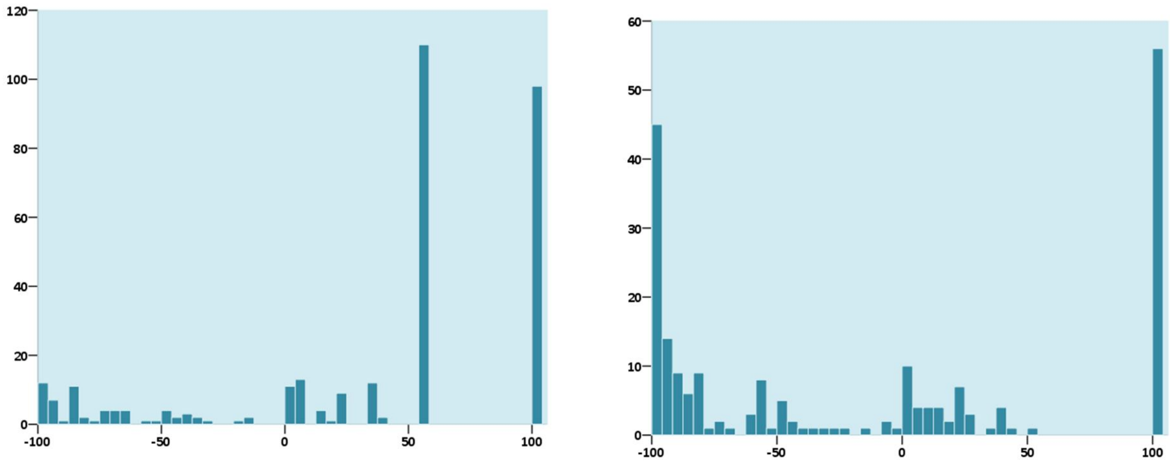


Histogram (%) of home bias under assumption SG for M: Professional activities (left) and N: Administrative activities (right).

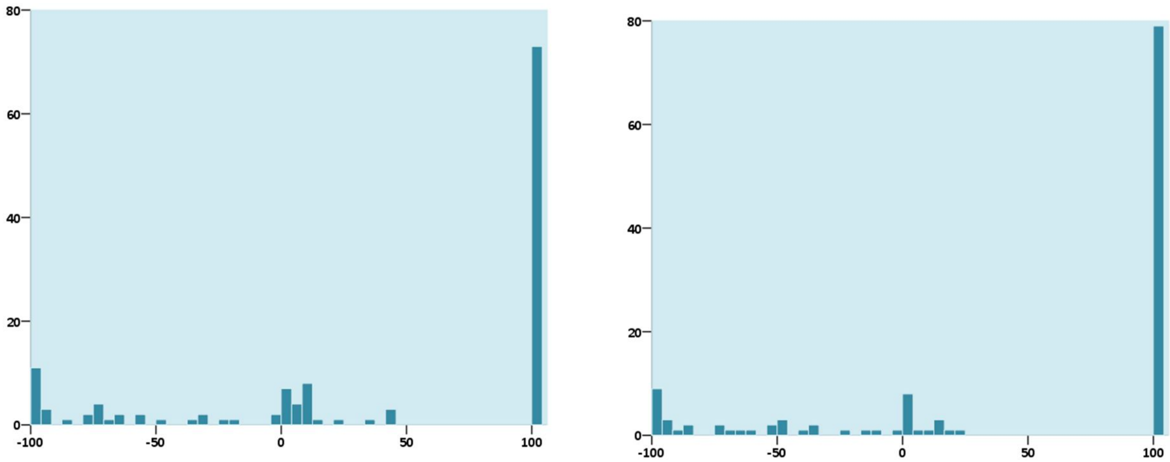
Appendix Figure 8 **Individual industrial relatedness bias for bidders of different industries.**



Histogram (#) of industrial relatedness bias for F: Construction (left) and G: Wholesale and retail trade (right).



Histogram (#) of industrial relatedness bias for J: Information and Communication (left) and K: Financial and iInsurance activities (right).



Histogram (#) of industrial relatedness bias for M: Professional activities (left) and N: Administrative activities (right).