ESSAYS ON THE ECONOMICS OF MERGERS & ACQUISITIONS

Dissertationsschrift
zur Erlangung des Grades
Doctor rerum politicarum (Dr. rer. pol.)
an der Universität Passau
Wirtschaftswissenschaftliche Fakultät

2012

vorgelegt von

Stefan Ebner

Erstgutachter: Prof. Dr. Michael Pflüger
Zweitgutachter: Prof. Dr. Lutz Bellmann

Disputation: Dienstag, 8. Januar 2013
Dissertationsort: Passau

To my parents.

Acknowledgments

My first thanks go to Michael Pflüger, University of Passau, for the supervision of this dissertation. I am thankful for his encouragement and the time he has spent to support and improve my research work. I am also grateful to my second advisor, Lutz Bellmann from the Institute of Employment Research (IAB) Nuremberg. I thank him for allowing me to work in his research department "Establishments and Employment". He and his whole department always gave me a feeling of being most welcome.

I acknowledge the financial support from the Passau Graduate School of Business and Economics for a scholarship, and the Bavarian Graduate Program in Economics (BGPE) for financing my participation in several research courses. Moreover, I am also grateful to a number of colleagues. In particular, I thank Hans-Dieter Gerner from the IAB for many stimulating discussions about my research work, and Viktoria Nußbeck who always organized my visitations at the IAB. Special thanks go to several colleagues at Michael Pflüger's chair at the Economics Department at the University of Passau for their support: Malte Mosel, Philipp Ehrl, Bohdan Kukharskyy, Stephan Russek, and Elisabeth Riesinger. I also thank Christian Pescher from the Ludwig-Maximilians-University Munich for his support, and Dennis C. Mueller for providing me longitudinal data about M&A.

The greatest thanks go to my parents. They always supported and encouraged me during the years of my education, and I will always be deeply indebted to them. For this reason, it was always clear to me that this thesis will be dedicated to them. Finally, I am grateful to my girlfriend Donya who helped me to come out of the deep valleys I had to walk through during writing this thesis.

Stefan Ebner

Passau, August 2012

Contents

List of Figures

List of Tables

Preface

In general, there are two possible ways for firms to grow. They can grow internally, i.e. they expand their production capacity by employing more workers and enlarging other production factors. Alternatively, they can grow externally through mergers and acquisitions (M&A), that is, they acquire a combination of existing production factors (Glaum and Hutzschenreuter, 2010). Both strategies have advantages and disadvantages. Buying a firm with already existing employees, technology, and products saves substantial time compared to internal growth. This is one of the main arguments for external growth in the form of M&A. Moreover, if firms need a technology for their growth strategy, which is protected by patents, or they need resources which are not easily available, the firms' goals usually cannot be reached with an internal growth strategy. In addition, through internal growth new production capacity will be generated which may increase the competition in the market. Instead, with an external growth strategy, the overall capacity in the market will be held constant, and the firms can gain market power, and thus, the competition in the market may probably even decline. A disadvantage of external growth is that acquisitions are not scaleable, and they require large investments and involve high risks. The acquired firms must be taken over with all their characteristics and employees, that is, they may also exhibit excess capacity or run inefficient plants. Contrarily, internal growth allows expanding gradually, enlarging the firms only with employees that are required, modern technology, efficient processes, etc.

Both strategies are subject to research in economics for decades. Contrary to internal growth, an external growth through acquisitions only redistributes existing property. Since no additional production capacity is generated, acquisitions should not affect growth in the economy as a whole. Nevertheless, it seems plausible to assume that mergers affect the firms' performance if they reallocate the combined firms' resources, causing synergy effects in the form of cost reduction, increased sales, or they increase market power. In economics, research questions about the effects of M&A on firms' performance were first discussed in the 1960ies and 70ies with a focus on the USA, and came up in Europe in the 80ies. Since both the number and volume of worldwide M&A has sharply increased within the last two decades, and hence, the phenomenon gained higher economic relevance, research activities have also become more intensive. There were some mega deals between large companies that attracted attention not only for researchers, but also for policy makers and the public, e.g. between the German steel companies Thyssen and Krupp in 1997,[1] the British mobile phone company Vodafone and the German industry company Mannesmann in 2000,[2] the Italian Bank UniCredit and the German bank HypoVereinsBank in 2005,[3] or the German automobile manufacturer Volkswagen and the German premium automobile manufacturer Porsche in 2012.[4] However, the increased M&A activity was not only caused by such large deals, but medium- and small-size firms were, and are, still engaged in M&A.

This dissertation builds on the economic research about M&A and the effects on the merging plants' performance. In particular, the objective of this thesis is to shed some light on questions about causal effects of M&A on plants' performance, taking firm heterogeneity into account. Since there is no typical merger (Tichy, 2001) it distinguishes between acquiring and target plants, and between horizontal and non-horizontal mergers. The thesis focuses on two major research questions: do plants with specific characteristics self-select in merger activity, and is there a causal effect of M&A on the merging plants' performance parameters, in particular on labor productivity, employment, and skill-intensity? The results allow drawing some conclusions about the reasons why plants merge.

The thesis consists of four chapters. All contributions have in common that they focus on questions about the effects of M&A on plant performance. That is, the thesis does not discuss questions about the effects of M&A on industry and aggregation concentration levels, or the effects of M&A on social welfare.

Each chapter in this thesis can be read separately, because they are based on stand-alone papers. Hence, all chapters have their own introduction and conclusion. The structure and storyline of this thesis and the interaction of the chapters are as follows: the first chapter is a survey about M&A and acts as an introduction to this research field. The second chapter describes propensity score matching as a newer microeconometric evaluation method and explains its implementation in the econometric computer software STATA. In a certain sense, the

[1] http://www.thyssenkrupp.com/de/konzern/geschichte_konzern_k5.html [July 17th 2012].

[2] Handelsblatt, February 2nd 2010: http://www.handelsblatt.com/unternehmen/it-medien/vodafone-mannesmann-die-mutter-aller-uebernahmeschlachten/3360804.html [July 17th 2012].

[3] Handelsblatt, July 27th 2008: http://www.handelsblatt.com/unternehmen/banken/hauptversammlung-hvb-das-ist-nicht-mehr-meine-bank/2995600.html [July 17th 2012].

[4] Spiegel Online, July 4th 2012: http://www.spiegel.de/wirtschaft/unternehmen/fusion-volkswagen-uebernimmt-porsche-schon-im-august-komplett-a-842658.html [July 17th 2012].

second chapter serves as a preparation for a better understanding of the econometric analysis performed in chapters three and four, which form the heart of the thesis. They both discuss questions about self-selection of plants into merger activity, and questions about causal affects on plants' performance. In particular, the third chapter focuses on the effects on merging plants' labor productivity, while the fourth chapter focuses on the effects on both employment and skill-intensity. Even if both chapters discuss the effects on different performance parameters, they are similar with respect to motivation, structure, and estimation strategy. Hence, there is some inevitable overlapping between these two chapters which are based on stand-alone papers as mentioned above.

The second, third, and fourth chapters use a new dataset, and, to the best of my knowledge, I am the first who was working with it so far. This dataset is a combined dataset from the IAB Establishment Panel and the M&A DATABASE from the University of St. Gallen, Switzerland. The IAB Establishment Panel is a representative employer survey for Germany, annually performed by the Institute of Employment Research (IAB) Nuremberg. The M&A DATABASE contains information about transactions for Germany, Austria, and Switzerland. The combined dataset contains German plants that merged domestically between 1996 and 2005, and also a control group of German plants that had not merged since 1980. The advantage of this dataset is the richness of plant-level variables, the differentiation between acquirers and targets, and between horizontal and non-horizontal mergers, and the availability of a control group, allowing a comprehensive econometric analysis. The number of observations differs between the three chapters: in the second chapter, I use all observations, but in the third and fourth chapter, some inevitable modifications of the dataset are necessary in order to perform the empirical methods (e.g. exclusion of observations with missing data for relevant variables).

Finally, I will now briefly explain the most important features, main contributions, and findings of the four chapters. In the first chapter, **"Mergers & Acquisitions: trends, reasons, effects"**, I review the literature about M&A. In particular, after some definitions and taxonomy, I describe the merger history and the current situation on the worldwide and German M&A market. Then, I discuss several theoretical hypotheses about motivations for M&A. In a further part, the chapter reviews empirical studies about effects of M&A on profitability, market share, market power, efficiency, productivity, employment, skill-intensity, and wages. However, the results do mostly not show a clear picture. For some performance parameters I observe different tendencies in the results between earlier and newer studies. Different data or estimation methods may be possible explanations. Studies about effects also provide evidence that mergers are differently motivated, i.e. no single hypothesis about motives of mergers explains all mergers.

In the second chapter, **"As easy as one, two, three... A guide to performing propensity score matching with STATA"**, I describe how to perform matching, which has become a popular tool in econometrics to evaluate treatment effects. In particular, the chapter is about the implementation of propensity score matching in the software STATA with the module PSMATCH2 developed by Leuven and Sianesi (2003). It is addressed to researchers not yet familiar with the method, and its objective is to quickly provide the basic understanding of the method and to simply explain how it can be performed in STATA. For this, the chapter describes the basic theoretical framework, and guides the reader step-by-step through the implementation of the method. Based on the dataset including merged and control plants, it presents the relevant commands in STATA, explains the corresponding results, discusses practical questions, and refers to further literature.

The third chapter, **"M&A and labor productivity: new evidence from micro-data for German plants"**, analyzes the impact of M&A on labor productivity of German merged plants in comparison to plants that were not involved in any M&A activity. I focus on two questions: does M&A impact the merging plants' productivity, and do more productive plants self-select in merger activity? Thereby, I apply a difference-in-differences propensity score matching approach, and I differentiate between subgroups of acquirers, targets, horizontally, and non-horizontally merging plants. I identify substantial pre-merger heterogeneity between plants and a strong support for a self-selection hypothesis of better performing plants into merger activities. I find a weak positive causal effect for acquiring plants, but the results show no support for a causal effect on the other subgroups.

The fourth chapter, **"Anybody afraid of M&A? Effects on German plants' employment and skill-intensity"**, analyzes the impact of M&A on employment and skill-intensity. The structure of the chapter is identical to the third chapter, and hence, the main research questions are: is there a self-selection of plants into M&A activity, and does M&A affect the plants' employment and the skill-intensity of the workforce? Again, I apply a difference-in-differences propensity score matching, and differentiate between acquirers and targets, and horizontal and non-horizontal mergers. In line with the third chapter, I confirm substantial pre-merger heterogeneity between plants. However, I do not find evidence for a causal effect of M&A on employment, but I find robust estimates that the skill-intensity of the target plants' workforce follows a U-shaped development path over time.

For a better reading, footnotes and equations are numbered independently in each chapter. Figures and tables are integrated in the text, and appendices, if necessary, can be found at the end of the chapters. Finally, I prefer American spelling to British spelling.

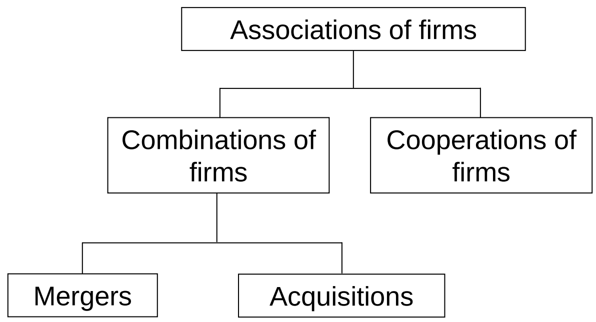
Chapter 1

Mergers & Acquisitions: trends, reasons, effects

1.1 Introduction

Different types of associations of firms exist which can be classified according to how intensively the firms are linked to each other, i.e. by how much firms reduce their economic independence (Pausenberger, 1989). In cooperations like membership in organizations, strategic alliances, or joint ventures firms preserve most of their independence. If, in contrast, the association of firms leads to a combination, at least one firm loses its economic independence. These combinations are either mergers or acquisitions. Figure 1.1 presents a classification.

Figure 1.1: **Different types of associations of firms**



Source: Pausenberger, 1989.

The usage of the terms "mergers and acquisitions" (M&A) is not consistent in the literature. Several definitions (e.g. from Vogel, 2002; Lucks and Meckl, 2002; Wirtz, 2003) characterize M&A as a transfer of leadership, managerial, and control authority. This requires to hold stakes or to invest in a firm's equity (Kirchner, 1991).[1] Parts of the literature do not distinguish between the terms "merger" and "acquisition", and they are often used as synonyms. However, differences exist and they will be briefly explained.

The difference between mergers and acquisitions: An acquisition describes the takeover of control of an existing, but formerly independent firm. Through acquisitions, firms fully give up their economic independence, but stay juridically independent. Two different modes of acquisitions exist. In acquisitions made by "asset deal" the acquirer buys all or the fundamental assets of the target, e.g. property, buildings, machines, etc. Acquisitions by "share deal" describe the acquisitions of shares in a capital company. In firms with other legal forms, the transfer involves the acquisition of stakes. In contrast to acquisitions, a merger creates

[1] It depends on a firm's legal form how much percent have to be held to constitute the stake. According to the German trade law ("Handelsrecht") 20% are sufficient, whereas the German law on stock companies ("Aktienrecht") requires more than 50%. In addition, a blocking minority of 25% is a possible way to influence corporate policy.

a new entity of formerly economically and juridically fully independent firms. There are two different ways to merge: either the target firm will be fully integrated into the acquiring firm, or both firms give up their legal entity and establish a new firm (Pausenberger, 1989; Glaum and Hutzschenreuter, 2010).

However, the main criteria by which to differentiate between mergers and acquisition, i.e. if firms preserve their legal entity or not after they combine, is obviously of minor importance in the literature. In the Anglo-American literature the terms "mergers" and "acquisitions" are virtually inseparable, which implies that differences disappear in practice (Grimpe, 2007). For this reason, I will use the terms "mergers", "acquisitions", "M&A", or "takeovers" as synonyms in the following text.

Different types and kinds of M&A: The M&A literature commonly distinguishes between three different types of mergers: horizontal, vertical, and conglomerate. In a horizontal merger firms which compete in the same market combine. In a vertical merger a firm combines with its supplier. And in a conglomerate merger firms of unrelated lines of businesses combine (Carlton and Perloff, 2005). Vertical and conglomerate mergers both decribe the combination of firms from different markets. Therefore, the literature uses the term "non-horizontal" M&A for both types of mergers (Church, 2004). However, the distinction between different types of mergers is not always clear cut and often depends on industry classifications.[2]

In vertical mergers, there is a distinction between forward integration, i.e. an upstream firm merges with a downstream firm, and backward integration, i.e. a downstream firm merges with an upstream firm (Church, 2008b). There are also different types of conglomerate mergers (Church, 2008a). First, mergers between firms producing complementary products means that consumers buy both products individually, but then assemble them for consumption. Second, mergers between firms producing neighboring products describe that goods are purchased by a common pool of buyers. That is, products are independent of each other, or they are weak substitutes, but they share the same distribution channels. And third, a merger between unrelated products means that products have no relation on either the demand or supply side.

The literature about M&A further distinguishes between friendly and hostile takeovers (Glaum and Hutzschenreuter, 2010). In friendly takeovers, the target firm's management sup-

[2]Pesendorfer (2003) analyzed horizontal mergers in the paper industry, and there are two kinds of firms. One group of firms produces finished cardboard boxes, and the other group produces linerboard, the raw material for cardboard boxes. However, the merger can also be considered as vertical, because the respective firms operate at different stages of the production chain. This makes clear that the cut between horizontal and vertical mergers also depends on industry classifications.

ports the acquisition process, e.g. by delivering all information about the target firm to acquirers. In contrast, takeovers are considered to be hostile if the management is not willing to negotiate about an acquisition. In this case, the potential acquirer submits an offer to shareholders to buy their shares, usually for a limited period of time, and mostly with a bonus to make the offer attractive. As a response, the target firm's management can try to impede the hostile takeover by making an alternative offer to their shareholders.

Finally, mergers can either be domestic or cross-border. In a domestic merger, the merging firms are located in the same country, whereas in cross-border M&A, the merging firms are located in different countries. Cross-border M&A are a way for firms to establish a foreign subsidiary, and they are an alternative to greenfield investments, where a new plant is built up from scratch (Barba Navaretti and Venables, 2004).

Firms in the M&A process: There are three different kinds of acquirers. First, acquirers buy other firms for strategic reasons. Popular motivations mentioned in the literature are market power, acquisition of complementary resources to strengthen the range of the firm's products and services, cost savings through economies of scale, financial synergies, etc.[3] The second group includes financial acquirers like investment funds, venture capital investors, or private equity companies. These acquirers buy firms with restructuring potential, and after restructuring, they resell them. The third group of acquirers are managers which buy the firm they work for in a so-called management-buy-out (MBO).

Owners of targets that are listed in a stock exchange are either private investors, institutional investors (funds, banks, insurances, etc.), entrepreneurs, or families of entrepreneurs. If the targets' shareholders sell their shares due to an attractive offer made by the acquirer, the acquirer takes control of the target. Owners of small or medium-size firms not listed on the stock exchange may also want to sell the firm, either because a multinational firm with larger personal and financial resources offers new growth opportunities, or, with respect to family owned firms, because there is no successor. In addition, firms disinvest and sell parts of their business for strategic reasons, e.g. because they put a higher focus on their core competences and get rid of business units (Glaum and Hutzschenreuter, 2010).

The structure of this survey is as follows: section 1.2 describes merger waves between 1895 and today. It also gives a brief overview of the current situation of the worldwide and German M&A market. Section 1.3 discusses the literature about reasons for mergers. In section 1.4, the effects of M&A on several performance parameters and are presented, and an overview of

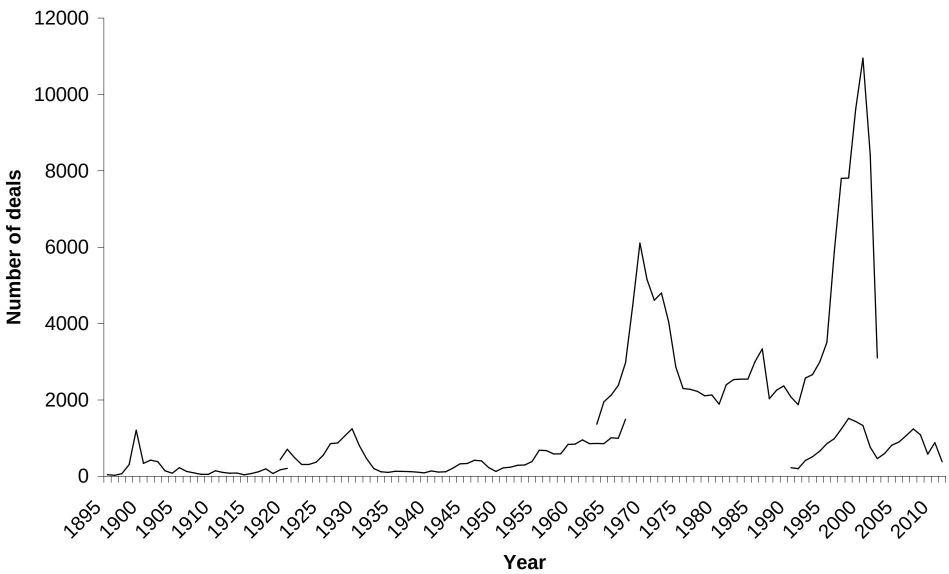
[3]See section 1.3 for a more comprehensive discussion about motives for mergers.

empirical studies is given. An explanation of the differing results of effects from mergers is also included. Section 1.5 presents the lessons learned and mentions further research questions.

1.2 M&A activities between 1895 and today

The longest data series about merger activity exists for the USA, i.e. the number of mergers is documented since 1895.[4] There are two characteristics of M&A activities over the last century. First, mergers come in waves, and second, M&A activities are correlated with stock market prices and economic activities (Mueller, 2003a). Figure 1.2 displays the number of deals in the USA between 1895 and 2010 and identifies six merger waves within the observation period with a peak in year 2000.

Figure 1.2: **Number of M&A deals in the United States between 1895 and 2010**



Note: the data stems from different sources. Hence, the number of deals differ when curves overlap. For example, the last curve is based on data from UNCTAD (2012) which only includes cross-border deals from large transnational purchaser companies. Since no deals from smaller firms are included in the data, the total number of deals is clearly lower. Source: 1895-1920: Nelson (1959); 1919-1967: Federal Trade Commission (FTC); 1963-2002: Town (1992); 1990-2010: UNCTAD (2012), World Investment Report (WIR).

1.2.1 Merger waves

The M&A literature about merger waves (e.g. Kleinert and Klodt, 2002; Hartford, 2005) has identified economic, technological, and regulatory changes as reasons for merger waves. These

[4] For West Germany, M&A data exists since the mid-1970ies, and for the European Union, data is available since the late 80ies.

changes cause a restructuring of sectors and economies which lead to combinations of firms. In this subsection, I briefly describe the six merger waves.

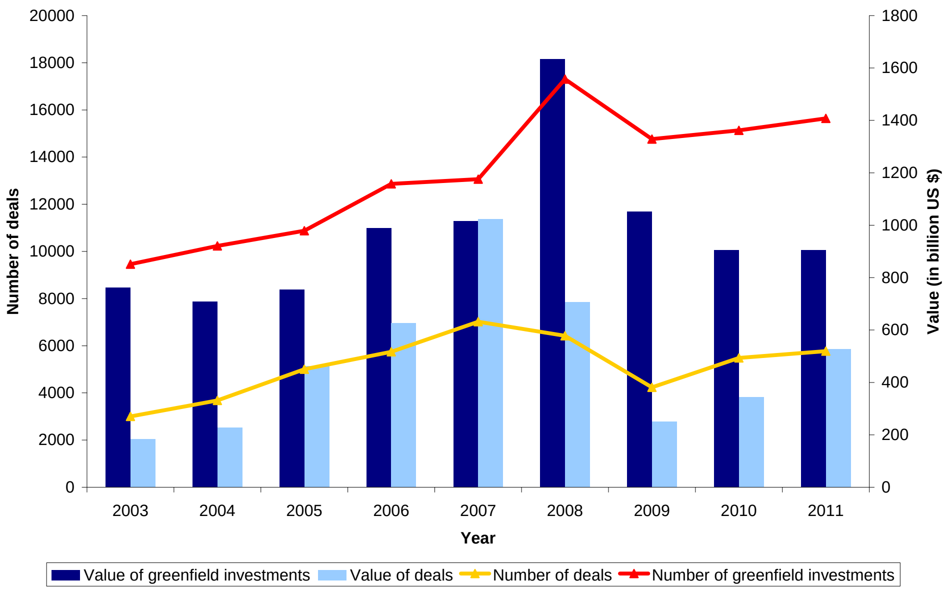
The first M&A wave occurred between 1897 and 1904 and was a reaction to the industrial revolution. The steam engine and the emergence of heavy industries enabled the exploitation of high scale economies in large firms. Large industry trusts emerged which are still characteristic for the old economy in the USA. In these times, M&A were mostly horizontally. However, the passing of the Sherman Act and the Clayton Act stopped the merger wave because M&A were impeded if they substantially increased a firm's market power. The second wave between 1920 and 1929 was dominated by vertical and conglomerate mergers. This was due to the Clayton Act which made horizontal mergers more difficult. The railway and utilities sectors were mostly affected by M&A in this time, and economies of scale were exploited due to new opportunities from these networks. The third merger wave started in 1965 and ended in 1973. The driving force behind the M&A activities were the exploitation of economies of scale from mass industry production, a diversification of products, and the acquisition of firms from other markets. The USA further controlled mergers by passing the Hart-Scott-Rudino Improvement Act of 1976, and Germany also introduced regulations in 1973 to control M&A activities. The fourth merger wave occurred between 1984 and 1988, but mostly in Europe. National firms prepared for the European single market and merged with European or international firms. Merging plants expected synergy effects due to a combination of production activities with related technologies. M&A were focused on technology intensive industries. As a reaction to increased merger activity, European antitrust laws were passed in 1989. The fifth wave started in 1995, and many merger activities were motivated by globalization and deregulation. Globalization creates larger markets, and firms followed this expansion, i.e. mergers were cross-border mergers. Another characteristic of this merger wave were large transactions, so-called "mega deals", and they occurred mostly in the telecommunication, pharma, oil, or banking sector. The fifth wave ended in year 2000 after the technology and dotcom bubble burst with a sharp decrease in the number of deals. However, only a few years later in 2002, the sixth merger wave started with numbers and volumes of deals similar to the preceding wave. The driving forces for mergers were institutional investors like hedge funds or private equity firms. Low interest rates also stimulated merger activities. The merger wave ended 2008 with the eruption of the financial crisis (Glaum and Hutzschenreuter, 2010).

Alternative explanations for merger waves focus on correlations between high stock prices and high numbers of acquisitions. This may be because a high number of acquisitions increases the valuation level of firms which are listed on stock markets. However, the causality could be reverse: if stock market prices are high, managers have an incentive to finance mergers by stocks, leading to higher merger activity (Rhodes-Kropf, Robinson, and Viswanathan, 2005; Mueller, Gugler, and Weichselbaumer, 2012).

1.2.2 The current M&A market

Worldwide trends: Figure 1.3 displays numbers and values of transnational firms' worldwide cross-border M&A and greenfield investments, the two main entry modes of foreign direct investment (FDI). The data refers to the period between 2003 and 2011, and is based on the World Investment Report (WIR) from the United Nations Conference on Trade and Development (UNCTAD).

Figure 1.3: **Number and value of worldwide cross-border M&A deals and greenfield investments between 2003 and 2011**



Source: UNCTAD (2012), World Investment Report (WIR).

The figure shows that both value and number of greenfield investments were higher than those of M&A in almost every year. The numbers and values for both greenfield investments and M&A peaked in years 2007 and 2008, and the following decline was a reaction to the financial crisis. In 2011, the latest year displayed in the figure, the worldwide value of cross-border M&A increased by 53% compared to the previous year and accounted for US-$ 526 billion. However, this amount was only half of the peak in 2007. The recent increase reflects both the growing value of assets on stock markets, and an increased financial capacity of acquirers to carry out these deals. This increase was driven by several mega deals in both developed countries and transition economies. In addition, corporate and industrial restructuring creates new opportunities for M&A deals, mostly for transnational corporations which have sufficient liquidity. In contrast, the values of greenfield investments stay constant, and the number of greenfield investments slightly increased in 2011 compared to 2010. These differing trends between greenfield investments and cross-border M&A emerge over time, because companies may consider both entry modes as alternatives to each other (UNCTAD, 2012).

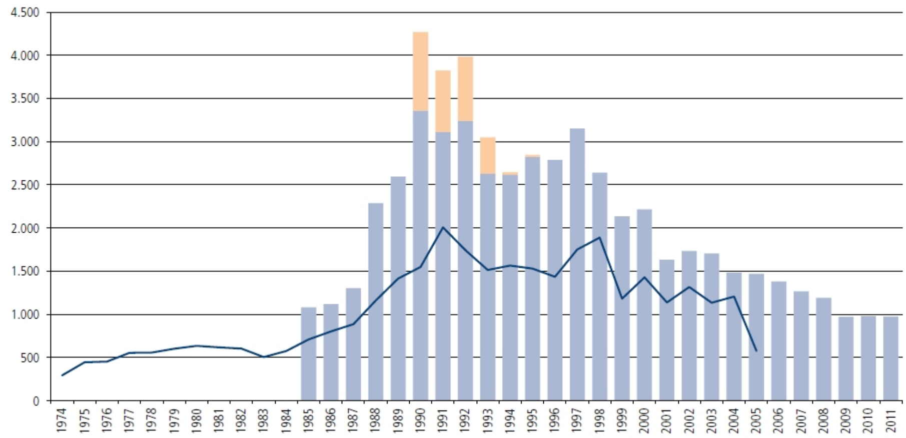
In addition to data from UNCTAD (2012) presented above, data from Dealogic delivers some further information about worldwide M&A transactions for 2011 (Spanninger, 2012). The sectors that accounted for the largest M&A volumes were real estate (10%), oil and gas (10%), and finance (9%). The USA was the largest market for M&A, accounting for 37% of the worldwide M&A volume, followed by China (7%), and Great Britain (5%). The European M&A market exceeded the US market between 2006 and 2008 in volume, but now lagged behind the USA again (29% of worldwide M&A volume). Within Europe, Great Britain accounted for the largest M&A volumes (17%), followed by France (12%) and Russia (10%). The volume for Germany was 7%. The emerging markets accounted for around 26% of the worldwide M&A volume.

The situation in Germany: Based on data from M&A DATABASE St. Gallen, figure 1.4 presents the development of transactions in Germany between 1974 and 2011, i.e. deals that involve German firms. Compared to the data from figure 1.2, Germany was less hit by the sharp decline in deals in year 2000. The aftermath of the financial crisis also affected Germany, but in contrast to the worldwide trend with a time lag: compared to the respective previous year, the number of deals with German firms involved declined by 6% in 2008 (1191 deals), and by another 18% in 2009 (972 deals), reaching a historical bottom line (Kunisch and Wahler, 2010). In 2010 and 2011, the number of deals remained at this low level with 979 deals in 2010, and 975 deals in 2011 (Spanninger, 2011b; 2012).

Table 1.1 presents statistics about the home countries of acquirers of German targets, and about home countries of targets acquired by German firms in 2011. More than half (51%) of all transactions with German firms involved were domestic M&A. For German acquirers, most foreign targets were located in neighboring countries (81 transactions): Switzerland (43), Austria (20), Netherlands (11), and France (7). Other important countries for German acquisitions were USA (28) and Great Britain (17). Foreign acquirers buying German firms were also mostly from neighboring countries (104), but from USA (45) and Great Britain (17) as well.

Table 1.2 displays the frequency of German mergers in different sectors and differentiates

Figure 1.4: **Number of M&A deals in Germany between 1974 and 2011**



Note: gray bar: data from M&A DATABASE; orange bar: sales by German Trust Agency (Treuhandanstalt); line: reported finished deals by German Federal Cartel Office (Bundeskartellamt, BKA). The 7th amendment of the law against restraints on competition (Gesetz gegen Wettbewerbsbeschränkungen, GWB) changed the rules for announcements of finished deals since 2005. This led to the decrease of reported deals by the BKA. The stagnant numbers of the M&A DATABASE are also explained by this modification. Source: M&A DATABASE, University St. Gallen; M&A REVIEW, 2/2012.

Table 1.1: **Acquirers and targets by countries for Germany, 2011**

Targets of German acquirers are from...		Acquirers of German targets are from...	
Country	Total number of mergers (in %)	Country	Total number of mergers (in %)
Germany	502 (51%)	Germany	502 (51%)
Switzerland	43 (4%)	Switzerland	53 (5%)
USA	28 (3%)	USA	45 (5%)
Austria	20 (2%)	Austria	29 (3%)
GB	17 (2%)	GB	17 (2%)
Netherlands	11 (1%)	France	12 (1%)
France/Italy	7 (1%)	Netherlands	10 (1%)

Source: M&A REVIEW, 2/2012; M&A DATABASE, University of St. Gallen.

between acquirers and targets for the year 2011. The majority of acquirers (19%) came from the financial sector. Most of their targets were either from the financial sector (34) as well, or from the service sector (42). The largest share of targets (22%) came from the service sector, followed by the chemicals and pharmaceuticals sector (8%), and the energy sector (8%). The table also shows that approximately 50% of all deals were intrasectoral, i.e. acquirers and targets belonged to the same sector.

Table 1.2: **Acquirers and targets by sectors for Germany, 2011**

	Sectors of targets																				
Sectors of acquirers	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	Sum	Share in %
01 Energy	44	2	3	3		2			2	1			5		1	1				64	6.6
02 Chemicals/Pharma	1	51	4	1	1	4	1	1	2	2			1			1				70	7.2
03 Steel/Metal	1		20	5	2	2							1			1				32	3.3
04 Manufacturing		1	4	29	1	1							3	1		1		3		44	4.5
05 Automobile		2	8	4	31	3			2	1			5				1			57	5.9
06 Electronics	5	2	12	1		14				1	1		4				1	3		44	4.5
07 Textile	1						10		1				1							13	1.3
08 Food		1				1		31	4	1			1							39	4.0
09 Trade	1	2	5			1	2	3	17	1			2	1						35	3.6
10 Finance	15	11	13	11	4	5	5	2	9	34		8	42	5	5	4		9		182	18.7
11 Insurance	2	1								2	11				1				1	18	1.8
12 Transport					1						1	23	4							29	3.0
13 Service	5	3	2	8	1	5	2	1	7	9		8	110	2				7	1	171	17.6
14 Media				2								1	10	17				1		31	3.2
15 Construction			1	2								1	3		11					18	1.8
16 Paper		3		1						1			1	1		6				13	1.3
17 Aerospace Techn.													1				3			4	0.4
18 Computer/IT	1					1			1			1	4	2	1			33		44	4.5
19 Others	3	3	2	4	4	3	5	1	3	4		2	13	4	2	5		8		66	6.8
Sum	79	82	74	71	45	42	25	39	48	57	13	44	211	33	22	18	5	64	2	974	100.0
Share in %	8.1	8.4	7.6	7.3	4.6	4.3	2.6	4.0	4.9	5.9	1.3	4.5	21.7	3.4	2.3	1.8	0.5	6.6	0.2	100.0	100.0

Source: M&A REVIEW, 2/2012; M&A DATABASE, University of St. Gallen.

1.3 M&A: different hypotheses about motivations and merging firms

The objective of this section is to understand why mergers occur. The literature has developed different theories about motives for mergers in the last decades, and they were surveyed by Müller-Stewens, Kunisch, and Binder (2010), Margolis (2006b), Jansen (2008), Scherer (2002), Tichy (2001) and others. A useful classification of hypotheses presented in the following subsections is to differentiate between firms that merge in order to maximize profits, and firms that merge for other reasons (Mueller, 2003a). I also discuss if firms self-select in M&A activity.

1.3.1 Profit maximizing motives

If firms maximize their profits, M&A should increase profits of the combining firms. Two ways are possible: firms can increase their revenues, and this may come from an increase of their market power. Alternatively, firms increase profits by cutting their costs, and this comes from increased efficiency. Thus, the most obvious motives for M&A are market power and efficiency increases, but there are differences between horizontal, vertical, and conglomerate mergers.

Increase of market power: Horizontal mergers can increase market power because they reduce the number of firms in the same market. In an oligopolistic industry, the merger then leads to higher prices. Several studies exist using models with a Cournot framework. For example, Salant, Switzer, and Reynolds (1983) presented a model for horizontal mergers and used a framework based on a symmetric Cournot equilibrium, homogeneous products, and identical constant unit costs for all firms. The model showed that mergers are never profitable for the merging firms, that is, horizontal mergers do not occur in this model. Other work using Cournot models exists from Perry and Porter (1985), Farrell and Shapiro (1990), Kamien and Zang (1990), and Baye, Crocker, and Ju (1996). In contrast, Deneckere and Davidson (1985) used a Bertrand framework, and the model predicted profits from mergers. The anticompetitive effects of horizontal M&A were modeled by Farrell and Shapiro (2001) who discussed the revised US Horizontal Merger Guidelines. Based on oligopoly theory regarding cost savings, competition, and consumer welfare, they concluded that any significant horizontal merger involves a loss of direct competition, and thus, is at least slightly anticompetitive. Similarly, and also based on oligopoly theory, Gugler, Mueller, Yurtoglu, and Zulehner (2003) pointed out that horizontal M&A may increase market power.

Vertical mergers are a way to increase market power by increasing entry barriers at one or more links in the vertical production chain (Comanor, 1967). Entry barriers can be established by foreclosing markets to competitors willing to enter the market. The literature distinguishes between input and customer foreclosure (Church, 2008b). Input foreclosure occurs if vertically integrated firms with market power at the upstream market stage do not sell to downstream rivals in the post-merger period any more, or they sell at higher prices, or offer lower quality. Anticompetitive effects emerge because of these higher prices, or due to lower quality for downstream rivals. Customer foreclosure means that the integrated downstream firm with market power at the downstream market stage no longer sources supply from independent upstream firms. This may lead to lower sales volumes and an increase in the average costs of upstream rivals, reducing its competitive pressure on the integrated upstream firm. As a consequence, the market power of the integrated upstream firm increases, and so do input prices. If higher input prices lead to higher prices in the downstream market, effects from customer foreclosure are anticompetitive. Anticompetitive effects may also emerge from coordination (Church, 2008b). Firms coordinate the increase in their prices in order to reduce the possibilities of substitution by customers to each other. Coordination effects arise from vertical mergers if post-merger firms are able to coordinate more effectively, i.e. it is easier to reach agreements on the coordinated outcome, or to make enforcements more effective. In this context, Nocke and White (2007) used models of symmetric upstream and downstream firms and demonstrated that vertical mergers will facilitate upstream collusion in an unintegrated industry. In a further paper, Nocke and White (2010) showed that vertical mergers with larger (in terms of capacity or products) downstream buyers are more likely to facilitate upstream collusion than mergers with smaller buyers.

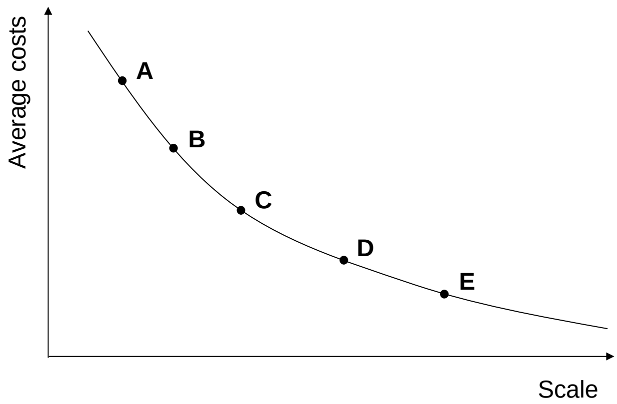
Conglomerate mergers also enable firms to gain market power. Firms make tacit collusion if they compete over time (multi-period, supergame situations), and thus, are willing to cooperate with their rivals and establish higher prices. This is beneficial to firms because the present discounted loss in profits over all future periods are higher than the gains from cheating today (Mueller, 2003a). Similarly, tacit collusion may also evolve if firms meet in different markets at the same time (multi-market contact). A high multi-market contact increases the costs for firms to cut prices in any given market. Hence, this can lead to more cooperative behavior (Gugler, Mueller, Yurtoglu, and Zulehner, 2003). Furthermore, conglomerate mergers may realize anticompetitive advantages because of an increase of the firms' portfolio or range of products (Church, 2008a). This allows the firm to engage in "contingent sale", i.e. the sale of one product in which the firm has market power is linked to the purchase of other products that were acquired in the merger. Examples for contingent sale are bundling or tying. A tying strategy means that customers who buy a product A also have to buy another product B, but product B is also individually available. In a bundling strategy two products, A and B, are sold only together in some fixed proportion, and they are not individually available (Nalebuff, 2003). Finally, anticompetitive effects emerge if conglomerate mergers lead to direct foreclosure, i.e. the conglomerates are able to post-merger make acquired complements not compatible with products of its rivals (Church, 2008a).

Efficiency increases: Increased profits from M&A can also come from increased efficiency. Again, there are differences between horizontal, vertical, and conglomerate mergers. Figure 1.5 presents the average costs of firms (A to E) which differ in size for an industry with significant scale economies. According to the figure firms exhibit smaller average costs if they increase their scale. Thus, horizontal mergers increase the scale of firms, leading to reduced average costs. If the decrease in average costs becomes smaller as the scale of the firms increase, as presented in the figure, cost reductions are larger for smaller firms compared to larger firms. Therefore, horizontal mergers should mostly be expected between smaller firms.[5] Additionally, horizontal mergers can lead to cost savings through a reorganization of production, or because the combination of formerly separated sales and distribution networks eliminates duplications (Pesendorfer, 2003).

With respect to vertical mergers, cost reductions come from the elimination of production

[5] However, this was not supported by empirical studies, e.g. from Mueller (1980a) for seven different countries.

Figure 1.5: **Industry with significant scale economies**



Source: Mueller (2003a).

steps. For example, consider a firm producing steel ingots that have to be cooled down before they are sold to another firm, which in turn has to reheat the steel in order to produce steel wire (Mueller, 2003a). Hence, vertical mergers may lead to higher quality, shorter lead times, improved quality control, reduced costs of inventory, optimized production runs, etc. (Riordan and Salop, 1995). Another argument for vertical mergers are transaction costs which emerge from the transfer of goods and services between firms. Thus, a vertical integration of firms reduces these transaction costs (Williamson, 1975). In addition, if two firms are in a mutual dependence due to a trading relationship, and they have to invest in assets that are specific to this relationship, i.e. the investment does not have any other use, firms behave opportunistic and underinvest. A vertical merger may solve this so-called "hold up problem" (Church, 2004).[6] In addition, a vertical merger is a way to eliminate double marginalization (Spengler, 1950). That is, if there are an upstream and a downstream monopolist, both set prices above marginal costs. If the downstream monopolist purchases inputs from the upstream monopolist, there is a markup on the markup (Church, 2004). With a vertical integration of the upstream monopolist the input will be transferred within the firm at marginal costs, eliminating the double marginalization.

The existence of economies of scope can also lead to conglomerate mergers. For example, two different products can be stored in a warehouse and delivered to a retailer. Since both products are stored and shipped together, the merged firms save costs (Mueller, 2003a).

Finally, the literature has identified several arguments why mergers can also create inefficiencies. For example, M&A may create diseconomies of bureaucracy, and the costs from the

[6] However, Grossman and Hart (1986) point out that integration does not entirely solve the hold up problem.

administration of larger units may more than offset the efficiency gains from economies of scale (Williamson, 1988). Vertical mergers can also lead to disadvantages like increased fixed costs if inputs are produced internally, less flexibility in changing business partners and relations, etc. (Porter, 1992). Moreover, synergy gains between targets and acquirers may not be realized because of geographical or cultural distances between the merging parties (Uhlenbruck, 2004).

1.3.2 Further motives

According to the arguments discussed above one should assume a counter cyclical pattern between M&A activities and economic growth. In a recession with low demand and high excess capacity, the competitive pressure is high, and firms want to cut costs. Thus, mergers that reduce costs should mostly occur in a recession. Similarly, in a recession the pressure to cut prices and to steal rivals' customers is high. Firms may see mergers as a way to eliminate rivals and increase their market power (Mueller, 2003a). However, empirical studies present a different picture and point to a positive correlation between merger activity and stock market prices over time (Glaum and Hutzschenreuter, 2010). This implies that other motives than those discussed above exist which explain the empirical findings about merger waves. Some motives will be briefly presented in this subsection (Mueller, 2003a).

The hypothesis about "empire building" states that managers merge because they pursue the firm's size growth instead of profit maximization. One explanation is the positive correlation between managers' income and the firms' size (Jensen and Murphy, 1990; Graßhoff and Schwalbach, 1997), i.e. managers may seek to increase the firm's size through M&A in order to expect an increase of their income afterwards. Other reasons why managers merge in order to increase the firm's size are the pursuit of non-monetary goals like increased power or prestige.

Another argument for mergers is the "free cash flow hypothesis" from Jensen (1986). If cash flow exists, and there are no more investment opportunities that are expected to increase the firm's value, this free cash flow should normally be paid to the firm's shareholders in order to maximize shareholder value. However, managers may instead prefer to keep the power over these financial resources, and one possibility is to use it for the takeover of other firms, even if the merger is not expected to increase the firm's value. This hypothesis is more relevant in mature industries with high cash flows and low investment needs.

The "hubris hypothesis" from Roll (1986) provides another explanation for mergers. When firms bid for a target, the acquirer with the most optimistic expectations about the target's future profits acquires it. Assuming that bidders have rational expectations, the bidder with the highest bid pays a price above the expected true value of the firm. This true value should be at the mean of the distribution. Thus, the winning bidder has probably bid too much, a phenomenon described by the literature as "winner's curse". However, it seems paradox that firms bid for targets even if they already know that the winner will lose. This can be explained by manager's hubris, i.e. they overestimate their M&A competence and claim to have better knowledge than the efficient capital market about the target's true value.

Mergers may also occur for "speculative motives", i.e. mergers are caused by promoters' profits, i.e. promoters (e.g. investment banks) approached corporate managers and suggested possible mergers. The promoters earn by fees charged from their advice and the services they rendered to finance and facilitate the merger.

According to the "adaptive (failing firm) hypothesis" mergers are seen as an alternative to bankruptcy (Dewey, 1961), i.e. mergers rescue firms from impending bankruptcy. Some mergers are part of a Darwinian process, i.e. underperforming firms disappear, even if they are not yet immediately bankrupt (Mueller, 2003a).

The "market for corporate control hypothesis" from Marris (1963, 1964) defines a valuation ratio V as the ratio of a firm's market value M and the book value of the firm's assets K. Under perfect competition this ratio should be one if managers maximize shareholder wealth. If, in contrast, managers pursue growth instead of shareholder wealth, V falls below its maximum value. According to Manne (1965) buyers in the market for corporate control acquire other firms as soon as V falls below its maximum value. This process ensures that corporate assets are allocated to the most competent managers and those who maximize shareholder wealth.

The "economic disturbance hypothesis" from Gort (1969) states that individual shareholders have expectations about a firm's future profits. These expected profits are associated with a price for shares. Shareholders are those individuals which expect these or higher profits. If those who did not hold shares change their expectations about the firm's future profits, and if they find that shares are undervalued, they buy the entire firm. If these individuals that did not hold shares before are managers of another firm, the transaction takes place as a merger.

There are two hypotheses about "financial efficiencies" due to mergers. First, larger firms face lower borrowing costs than smaller firms, and firms invest until the marginal return on investment equals the cost of capital. If the smaller and the larger firms merge, the smaller firm is able to make additional investments with lower returns at lower borrowing costs. Second, and according to the portfolio theory of Markowitz (1952), mergers may be motivated by reasons of risk pooling. A portfolio of assets generates the same average return at a lower risk (measured as variance) compared to the sum of its elements if the correlation of returns on assets is not perfect. A diversified firm can be thought of as a portfolio of assets of separate lines of businesses which faces a lower risk than the same number of businesses as stand-alone-firms.

According to the "capital redeployment hypothesis" multidimensional organized firms are able to establish an internal capital market, and thus, avoid the dangers of external capital markets. Mergers are a way to establish a diversified firm with such an internal capital market. This hypothesis goes beyond the hypothesis about saved borrowing costs from above. It points out that potential gains are generated from the ability of a central management to monitor the different investment opportunities of different divisions and to shift capital across them (Williamson, 1970; Weston, 1970).

Of course, no single hypothesis about M&A is able to explain all mergers, but all hypotheses are able to explain at least some of them (Steiner, 1975). In particular, some of the presented hypotheses are more consistent with the observed merger waves than others.[7] Mueller and Sirower (2003) tested which of the competing hypotheses received the most support. The test was based on a dataset of 168 large acquisitions between 1978 and 1990, and the authors found out that the mean lost for bidders is US-$ 50 million with a large variance of US-$ 3,580 billion. They wondered why managers are willing to play a game with negative expected winnings and such a high variance and concluded that managerial hubris is one plausible explanation. These managers believe they are able to see values in the target firm which other managers do not. The second hypothesis which received the most support from the tests is managerial discretion, i.e. mergers do not create any gains, and managers merge in order to increase the firms' size. Thus, prices paid as premium to targets' shareholders are losses to the bidder. In other words, managers gamble with other people's money.

In addition to the motives for M&A mentioned above, the literature also discusses reasons for cross-border mergers. Firms have different options to enter foreign markets, either through exporting, greenfield FDI, or cross-border M&A. They may prefer M&A for strategic reasons, e.g. because the foreign target possesses specific assets like better knowledge of the foreign local market. Several theoretical contributions discussed the firms' choice between different entry modes. For example, Barba Navaretti and Venables (2004) used a model to compare mergers with greenfield investments and argued that M&A are promoted when firm-level fixed costs are large relative to the market size of home and foreign country. A further contribution is provided by Görg (2000) who formalized a firm's decision between cross-border M&A and greenfield investment when entering foreign markets via FDI. He showed that in an asymmetric duopoly situation the new entrant will be best off by taking over an existing indigenous low technology firm, and thus, forming a duopoly with an indigenous high technology firm. Nocke and Yeaple (2007) developed a general equilibrium model with heterogeneous firms. They found out that cross-border mergers involve the most or the least efficient active firms, depending on whether firms differ in their mobile or immobile capabilities.

In this context, the literature identified positive effects from lower trade costs and liberal-

[7]See Mueller (2003a) for an assessment of each hypothesis and the compatibility with the picture of merger waves.

ization on cross-border M&A activity. Horn and Persson (2001) presented a model to analyze cross-border M&A in an international oligopolistic market. They found out that high trade costs may be conducive to national ownership of assets, and international firms may arise when trade costs are lower. This is in contrast to what the "tariff-jumping" argument suggests. Neary (2007) presented a two country model of oligopoly in general equilibrium for cross-border acquisitions and showed that trade and capital market liberalization leads to an international merger wave. Breinlich (2008) used the Canada-United States Free Trade Agreement of 1989 to estimate the impact of trade liberalization on M&A activity. He showed that freer trade leads to a significant increase in merger activity. Hijzen, Görg, and Manchin (2008) empirically analyzed the role of trade costs for cross-border M&A for 23 OECD countries. They found evidence that trade barriers negatively impact cross-border M&A activities. However, horizontal mergers are less negatively affected compared to non-horizontal mergers, supporting the tariff-jumping argument.

1.3.3 Which firms merge? The self-selection hypothesis

There is both theoretical and empirical support for a self-selection of firms into M&A activity, i.e. firms that merge seem to systematically differ in certain characteristics from firms that do not merge. Hence, the pre-merger performance of firms has become an important issue in the M&A literature.

From a theoretical perspective, acquiring firms may be overperforming firms because they are able to bear fixed costs that emerge due to the takeover process. This argument corresponds with the discussion from Melitz (2003), Helpman, Melitz, and Yeaple (2004), and others about heterogeneous firms. They found out that only the most productive firms engage in foreign activities like exporting or FDI. Contrary, acquirers may also be poor performing firms which merge with another firm in order to survive on the market. Spearot (2007a) added M&A into the model framework of heterogeneous firms from Melitz (2003) and Helpman, Melitz, and Yeaple (2004) and found that mid-productivity firms are most likely to acquire other firms. In a further paper from Spearot (2007b), using the same model framework, firms that acquired abroad were more productive than firms that acquired at home. In an empirical study about mergers in US manufacturing firms, Maksimovic and Phillips (2001) found that acquirers tend to be relatively more productive firms. These findings are consistent with the results from Andrade and Stafford (2004) for firms in different US industries. Earlier studies came to different results: Mueller (1980b) and Harris, Stewart, and Carleton (1982) provided evidence that acquiring firms' average profit rates do not differ from other firms. For conglomerate mergers, Weston and Mansinghka (1971) and Melicher and Rush (1974) found that acquiring firms had profits below average, and also lower profits than the acquired firms.

Turning towards targets, acquirers may select overperforming firms, so-called "cherries". Acquirers expect to benefit from the targets' assets and capabilities like advanced technology, management skills, and large market shares in order to realize efficiency gains (Balsvik and Haller, 2011). In contrast, acquirers may also focus on underperforming targets, so-called "lemons". Acquirers expect the targets' performance to improve after the merger due to new management which realizes the potential of the targets' assets. Empirical evidence on this issue came from Bellak, Pfaffermayr, and Wild (2006). They compared performance parameters for Austrian manufacturing firms and showed that lemons exhibit significantly higher profitability growth rates than cherries, but there were no differences with respect to employment and productivity growth. Salis (2008) found support for the cherry-picking argument based on data for Slovenia. Arndt and Mattes (2010) analyzed cross-border M&A and found that both cherries and lemons are acquired.

1.4 Effects of M&A on firms' performance: an empirical overview

There are three different sets of consequences from mergers (Mueller, 2003a). First, M&A may affect the performance of firms. Second, M&A can affect industry and aggregation concentration levels, and third, M&A may affect social welfare. In this section, I discuss the first group of effects. In particular I present a survey of studies that estimate the effects from M&A on firms' profitability, market share, market power, productivity, employment, skill-intensity, and wages,[8] and I provide an overview of the empirical results over the last decades. For each performance parameter, I will first provide an overview of earlier studies, and then, present the results from recent studies, mostly published since year 2000. I will show that the results are often ambiguous, and for some performance parameters, the results of newer studies differ slightly from those of earlier studies. Then, I briefly discuss further effects of M&A, and also present the results from event studies, an alternative approach to assess the effects of mergers. Finally, I discuss why empirical analysis does not give clear answers about effects of M&A.

1.4.1 Effects on profitability

From a theoretical point of view, the average merger should generate positive profits if managers maximize profits and have rational expectations. Moreover, it should be expected that managers

[8] There also exists research about effects on other performance parameters, e.g. R&D (e.g. see Tichy (2001) for an overview), but the literature is scarce.

undertake mergers deliberately and make the decision with great care (Mueller, 2003a).

Nevertheless, empirical studies provide a different picture and present ambiguous results. One of the most comprehensive studies about effects of mergers on profitability came from Ravenscraft and Scherer (1987) for the USA who analyzed 6000 observations between 1950 and 1977. The authors found that profits of acquired firms declined after the acquisition. Negative effects were also found in Reid (1971), Melicher and Rush (1973, 1974), and Mueller (1986) for the USA, Meeks (1977), Hughes (1989), Cosh, Hughes, and Singh (1980) and Kumar (1985) for the UK, in Peer (1980) for Holland, and Ryden and Edberg (1980) for Sweden. In contrast, there exist studies that found positive effects: Weston and Mansinghka (1971), Mueller (1980b) and Healy, Palepu, and Ruback (1992) for the USA, Baldwin (1998) for Canada, and Ikeda and Doi (1983) for Japan. No significant changes were found in Cable, Palfrey, and Runge (1980) for Germany, McDougall and Round (1986) for Australia, and Rhoades (1987) and Healy, Palepu, and Ruback (1997) for the USA.[9]

Table 1.3: **Recent studies about effects of M&A on profits**

Country	Authors	Period	Unit	Merger sample	Control group	Estimation method	Profit measure	Change
Austria	Bellak, Pfaffermayr, and Wild (2006)	1985-2002	Firms	60 foreign mergers in manufacturing firms	421 non-acquired firms	DiD PSM	Cash flow	Lemons: >0; cherries: <0
France	Bertrand and Zitouna (2008)	1993-2000	Firms	371 targets in horizontal M&A of manufacturing firms; domestic/ cross-border M&A	Non-acquired firms	DiD PSM	EBITDA	≈ 0
UK	Conyon, Girma, Thompson, and Wright (2004)	1979-1991	Firms	190 domestic related/ unrelated mergers in 140 manufacturing firms	236 non-acquired firms	Regression	Profit per worker	>0
USA	Bhuyan (2002)	1992	Firms	Vertical mergers in 43 food manufacturing industries	None	Regression	(Total sales - total costs)/ total sales	<0
	Pesendorfer (2003)	1984-1987	Firms	31 horizontal mergers in the paper industry	Non-acquired firms in the same industry	Regression	Profits before tax	>0
	Gugler, Mueller, Yurtoglu, and Zulehner (2003)	1981-1998	Firms	$\approx$2700 mergers in manufacturing and service sector	Non-acquired firms in the same industry	Regression	Profits before interest and taxes/ total assets	>0

Note: DiD PSM: difference-in-differences propensity score matching.

In addition to these earlier studies, table 1.3 presents an overview of recent empirical studies, and they have two major advantages: first, the quality and availability of data is better, and second, research about questions on causality and self-selection has improved due to advanced research designs (Angrist and Pischke, 2010). However, similar to earlier contributions the empirical results of these newer studies about effects of M&A on profits are also ambiguous,

[9] Comprehensive overviews of empirical studies mostly published before year 2000 can be found in Tichy (2001), Mueller (2003a), Gugler, Mueller, Yurtoglu, and Zulehner (2003), and Jansen (2008).

i.e. despite newer data and advanced methods research has not brought us forward in this issue. In subsection 1.4.9 I will discuss several reasons why the results differ from each other.

The conclusion from a large number of empirical studies from the last decades about effects from M&A on profits is that mergers do not necessarily increase profits, and it seems even more likely that a large proportion of mergers even decrease merging firms' profits. These findings support the view that not all mergers are motivated by profit maximizing reasons, but occur for other reasons as discussed above. Hence, these conclusions are similar to those from earlier surveys, e.g. from Bühner (2002).

1.4.2 Effects on market share, market power, and efficiency

Even if not all mergers are profitable, it is important to understand how M&A increase profits. Taking a simple model from Gugler, Mueller, Yurtoglu, and Zulehner (2003) there are two possibilities how profits can increase after mergers. First, mergers lead to efficiency increases with a fall in the merging firms' costs. In profit maximizing firms, lower marginal costs should lead to lower prices, increasing both sales and profits.[10] Alternatively, efficiency improvements may also occur in the form of better product quality, and this should also lead to a rise in firms' sales and profits. This situation is shown in the upper left cell of table 1.4. Second, increased profits can also come from increased market power. If firms have market power and are able to control prices, and if they maximize profits, an increase in prices can be expected, leading to a decrease in output and sales (lower left cell). By simply analyzing how output changed in comparison to the situation in which the firms had not merged, one can conclude if profits increased due to efficiency gains or due to market power increase.

Table 1.4: **Possible consequences of M&A**

	Profits > 0	Profits < 0
Sales > 0	Efficiency increase	Market power reduction (?)
Sales < 0	Market power increase	Efficiency decrease

Source: Gugler, Mueller, Yurtoglu, and Zulehner (2003).

Since one can expect that changes in sales will translate into corresponding changes in market shares (Gugler, Mueller, Yurtoglu, and Zulehner, 2003), studies analyzing market shares help to find out if mergers increase efficiency or market power. If mergers increase efficiency or improve the products' quality, the mergers are expected to increase the firms' market shares.

[10] A merger is also efficiency improving if it reduces fixed costs but not marginal costs. This should increase profits but not change sales. This case is not considered here (Gugler, Mueller, Yurtoglu, and Zulehner, 2003).

However, no changes were found in studies from Goldberg (1973) for a sample of advertising intensive firms, and in Rhoades (1987) for acquired banks. Mueller (1985, 1986) compared the market shares of acquired manufacturing firms between 1950 and 1972 to a group of non-acquired firms and found a decrease in market shares. Baldwin and Gorecki (1990) found evidence for a decrease of market shares for acquired Canadian plants in horizontal mergers, but no effects for other sorts of mergers. In a study about horizontal mergers in the US paperboard industry, Pesendorfer (2003) also found that merging firms lose market shares. With respect to changes in sales, Gugler, Mueller, Yurtoglu, and Zulehner (2003) found a decrease for merging firms on average. The decrease in sales was stronger in conglomerate mergers compared to horizontal mergers. These studies did not provide any evidence that M&A led to an increase in the firms' market shares, and thus, they did not provide evidence for an efficiency increase after M&A.

The model from Gugler, Mueller, Yurtoglu, and Zulehner (2003) also analyzed negative profits and changes in sales, displayed in the right part of table 1.4. If mergers are motivated by growth, or because managers suffer from hubris, efficiency must not necessarily decline or market power must not increase. Nevertheless, one can expect that these mergers lead to transaction costs which arise if two firms with different organizational structures and different company cultures are brought together, with a negative impact on efficiency. Therefore, efficiency decreasing mergers should both decrease profits and sales, as displayed in the lower right cell of table 1.4. Alternatively, mergers may also reduce profits but increase sales (upper right cell). This situation can be labeled with "market-power reduction", because it reflects the opposite to a market power increase. However, the combination of reduced profits but increased sales seems not really plausible,[11] and this scenario may most likely occur if managers maximize growth or sales instead of profits.

Gugler, Mueller, Yurtoglu, and Zulehner (2003) empirically tested these four hypotheses from table 1.4 and showed that the fractions of mergers resulting in efficiency increase, market power increase, and efficiency decrease all account for almost 30%. The puzzling situation with decreased profits and increased sales ("market-power reduction") accounts for around 15%. Summarizing, less than 30% of all mergers increase efficiency, and therefore, are welfare enhancing. Assuming that increased market power and decreased efficiency is welfare reducing, the majority of mergers reduces social welfare.

[11] For this reason, Gugler, Mueller, Yurtoglu, and Zulehner (2003) add a question mark to this categorization (see table 1.4).

1.4.3 Effects on productivity

Effects of M&A on productivity are of great interest, because it is a measure of economic efficiency. There are several ways to estimate productivity, e.g. as labor productivity, defined as output or sales per employee, or as total factor productivity (TFP). The latter is superior because it measures effects on all inputs, but the required data to estimate TFP is not always available.

If mergers are motivated by profit maximizing reasons, productivity changes after M&A are expected to be positive. Different studies exist that analyze effects on productivity. For example, Lichtenberg and Siegel (1992a) found that productivity in US manufacturing plants fell before the merger, but rose afterwards. Productivity increases after horizontal mergers are found in Baldwin (1998) for Canadian plants. McGuckin and Nguyen (1995) examined productivity effects for plants in the US food manufacturing industry. They observed post-merger productivity improvements for acquired plants but productivity losses for the buyer's existing plants. Contrary, Caves and Barton (1990) and Lichtenberg (1992a) found lower productivity in plants that were held by diversified firms compared to plants in undiversified firms. A survey of earlier studies about productivity effects of M&A is provided by Caves (1989). Summarizing, the majority of earlier studies concluded that M&A reduces productivity, and this is in line with the results from several studies identifying a negative impact of M&A on profits.

In contrast to these earlier studies, a large majority of recent studies found positive productivity effects, as shown in table 1.5. Hence, these results support profit maximizing hypotheses as explanations for mergers. However, this seems somewhat puzzling with respect to the results from studies about the effects on profits and sales. Several explanations are possible, e.g. different estimation methods, observation periods, datasets, etc., which will be explained in greater detail at the end of this section.

1.4.4 Effects on employment

Predictions about effects of M&A on employment are difficult. Mergers which are motivated by profit maximization are more likely to be followed by cost savings and employment reductions compared to mergers that are differently motivated (Conyon, Girma, Thompson, and Wright, 2002a). Ownership changes may lead to displacement of management, plant closure, etc. with negative effects on employment on the one hand. On the other hand, new ownership may also bring new capital inflows, expertise, etc. with positive employment effects (McGuckin

Table 1.5: **Recent studies about effects of M&A on productivity**

Country	Authors	Period	Unit	Merger sample	Control group	Estimation method	Profit measure	Change
Austria	Bellak, Pfaffermayr, and Wild (2006)	1985-2002	Firms	60 foreign acquisitions in manufacturing firms	421 non-acquired firms	DiD PSM	Labor productivity (value added per worker)	≈0
France	Bertrand and Zitouna (2008)	1993-2000	Firms	371 targets in horizontal M&A of manufacturing firms; domestic/ cross-border M&A	Non-acquired firms	DiD PSM	Total factor productivity	>0, stronger for cross-border M&A
Germany	Arndt and Mattes (2010)	1997-2003	Firms	158 cross-border M&A	Non-acquired domestic multinationals	DiD PSM	Total factor productivity	>0
	Mattes (2010)	2000-2007	Plants	352 foreign acquisitions of domestic plants	≈15000 non-acquired plants	Regression; DiD PSM	Labor productivity (sales per worker)	≈0
India	Petkova (2009)	2001-2006	Firms	150 cross-border M&A in manufacturing firms	1470 non-acquired plants	DiD PSM	Total factor productivity	>0
Indonesia	Arnold and Javorcik (2009)	1984-1994	Plants	185 plants that switched from domestic to foreign ownership	≈2000 non-acquired plants	Regression; DiD PSM	Total factor productivity	>0
Italy	Piscitello and Rabbiosi (2005)	1994-1997	Firms	113 foreign acquisitions of domestic manufacturing plants	374 non-acquired firms from a random control sample	Regression	Labor productivity (sales per worker)	>0
Slovenia	Salis (2008)	1994-1997	Firms	186 foreign acquisitions of manufacturing firms	≈1000 domestically-owned and non-acquired firms	DiD PSM	Total factor productivity	≈0
Sweden	Bandick (2011)	1993-2002	Firms	464 foreign mergers of manufacturing firms; MNEs/ non-MNEs; horizontal/ vertical mergers	≈4000 non-acquired firms	Regression; DiD PSM	Total factor productivity	Vertical acquisition: >0; Horizontal acquisition: ≈0
UK	Conyon, Girma, Thompson, and Wright (2002b)	1989-1994	Firms	129 foreign mergers	Random control sample of 642 firms	Regression	Labor productivity (sales per worker)	>0
	Griffith and Simpson (2004)	1980-1996	Plants	≈8.800 observations of plants changing ownership from domestic to foreign	≈4600 observations of plants changing ownership from foreign to domestic	Regression	Labor productivity (value added per worker)	>0
	Girma, Thompson, and Wright (2006)	1988-1996	Firms	542 foreign acquisitions (US and European MNEs) of domestic manufacturing plants	454 non-acquired firms	Regression; DiD PSM	Total factor productivity	>0
USA	Maksimovic and Phillips (2001)	1974-1992	Plants	17720 mergers in manufacturing industry	Non-merged plants	Regression	Total factor productivity	>0
	Ollinger, Nguyen, Blayney, Chambers, and Nelson (2006)	1977-1992	Plants	≈5.000 acquired and ≈12000 buyer plants in food industries	≈10000 non-acquired plants in the same industry	Regression	Labor productivity (total value of shipments (output) per worker)	>0

Note: DiD PSM: difference-in-differences propensity score matching.

and Nguyen, 2001). Another argument points out that new management is less committed to employees, and therefore, renegotiates explicit and implicit labor contracts and conditions. These renegotiations may be seen as a "breach of trust" by employees with negative employment effects (Shleifer and Summers, 1988). Effects on the workforce may also differ between different types of mergers. Conyon, Girma, Thompson, and Wright (2002a) argued that horizontal mergers lead to higher employment losses if there are increasing returns to scale. Dutz (1989) also expected negative effects from horizontal mergers in declining industries. Williamson (1975) expected negative effects from vertical integration due to layoffs in the sales function in the upstream firm, and in the procurement function of the downstream firm. In addition, because of a closer proximity between acquirer and target, domestic mergers make radical structural reforms with negative employment effects more likely than cross-border M&A (Lehto and Böckerman, 2008). Gugler and Yurtoglu (2004) argued that mergers are seen as a way to restructure and optimally adjust the workforce, and thus, higher employment losses can be expected in countries with rigid labor markets.

The empirical evidence about employment effects is ambiguous. In earlier studies, a reduction in employment was found in Baldwin (1998) who showed that M&A had a negative effect on employment of non-production workers. According to Bhagat, Shleifer, and Vishny (1990), 45% of US firms involved in hostile takeovers laid off workers. Brown and Medoff (1988) found negative effects for the state of Michigan, USA. Lichtenberg and Siegel (1990) showed that employment growth was lower in central office plants compared to production plants after M&A. In a further study, Lichtenberg and Siegel (1992a) found negative effects for larger firms in the US manufacturing sector. In contrast, several studies estimated positive effects. For example, McGuckin, Nguyen, and Reznek (1995) reported about rising employment in acquired plants in the USA, but they did not find significant effects at the firm-level. In McGuckin, Nguyen, and Reznek (1998), positive employment effects after M&A were found for the US food manufacturing industry.

Table 1.6 presents a list of newer studies about employment effects.[12] The results of the studies are also ambiguous, but with a tendency towards a decrease of employment. Hence, these negative employment effects provide support that mergers are rather driven by profit maximizing motivations than other reasons (Conyon, Girma, Thompson, and Wright, 2002a).

[12] See also Siegel und Simons (2008) for an overview of plant- and firm-level studies about effects of M&A on employment and wages.

Table 1.6: **Recent studies about effects of M&A on employment**

Country	Authors	Period	Unit	Merger sample	Control group	Estimation method	Change
Austria	Bellak, Pfaffermayr, and Wild (2006)	1985-2002	Firms	60 foreign acquisitions in manufacturing firms	421 non- merged firms	DiD PSM	≈0
Finland	Lehto and Böckerman (2008)	1989-2003	Plants	7923 foreign and domestic mergers	Non-merged plants	DiD PSM	<0
Germany	Arndt and Mattes (2010)	1997-2003	Firms	158 cross-border M&A	Non-merged domestic multinationals	DiD PSM	≈0
	Mattes (2010)	2000-2007	Plants	353 foreign acquisitions of domestic plants	≈12000 non-merged plants	DiD PSM	≈0
Sweden	Siegel and Simons (2008)	1985-1998	Employees	Employer-employee data for 19000 firms; differentiated in full/ partial acquisitions, divestitures; related/ unrelated	Employees in non-merged firms	Regression	<0
UK	Conyon, Girma, Thompson, and Wright (2001)	1983-1996	Firms	201 friendly and 39 hostile domestic mergers in 195 firms	238 non-merged firms	Regression	Hostile: <0; Friendly: <0
	Conyon, Girma, Thompson, and Wright (2002a)	1967-1996	Firms	442 domestic mergers in 277 firms; differentiated in related/ unrelated; hostile/ friendly	298 non-merged firms	Regression	<0
	Girma and Görg (2004)	1980-1993	Plants	239 foreign acquisitions in the electronics and 121 in the food industry	524 domestic plants in electronics and 241 plants in food industry	DiD PSM	<0
	Girma (2005)	1988-1998	Firms	542 foreign mergers in the manufacturing sector	454 matched non-merged domestic firms	Regression; DiD PSM	≈0
	Amess and Wright (2007)	1999-2004	Firms	1350 leveraged buyouts (LBOs)	4029 non-LBOs	Regression	≈0
USA	McGuckin and Nguyen (2001)	1977-1987	Plants	≈20000 merging plants in the manufacturing sector	≈300.000 non-merged plants	Regression	Typical plant: >0; Bigger plants: <0
	Gugler and Yurtoglu (2004)	Since 1970	Firms	646 mergers; differentiated in related/ unrelated; friendly/ hostile; tender/no tender; domestic/ cross-border mergers	≈10.000 non-merged firms	Regression	US: ≈0; UK: <0; Continental Europe: <0
	Ollinger, Nguyen, Blayney, Nelson, and Chambers (2005)	1977-1992	Plants	≈31.000 merged plants in food industry	Non-merged firms in the same industry	Regression	≥0

Note: DiD PSM: difference-in-differences propensity score matching.

1.4.5 Effects on skill-intensity

In the context of employment changes due to M&A it is an interesting question if the merging firms' skill-intensities change. From a theoretical point of view, merging firms may be able to use firm-level assets together (e.g. management, administration, marketing, R&D, etc.), and therefore, they may seek to rationalize these activities (Barba Navaretti and Venables, 2004). These firm-level activities are supposed to be mostly performed by white-collar workers. If these white-collar workers are higher skilled than production workers, and if production workers are not affected by M&A, a reduction of the merging firms' skill-intensities should be expected. In contrast, unskilled workers may suffer if mergers lead to reorganizational changes. Lindbeck and Snower (2000) argued that measures of reorganizational changes require specific skills of employees. If skilled workers have these skills, unskilled workers may be laid off, leading to higher skill-intensities in firms.

The empirical literature on this issue is scarce. For plants of the US manufacturing sector, Lichtenberg and Siegel (1990) found that white-collar workers suffer more from M&A compared to production workers. Similar results were found in Bhagat, Shleifer, and Vishny (1990) for the USA. In a study for Finland, Huttunen (2007) found that acquired plants reduce the share of highly educated workers after a change from domestic to foreign ownership. Contrary, Girma and Görg (2004) found negative employment effects after foreign takeovers, in particular for unskilled workers in the UK electronics industry. There are studies about hostile mergers that report about negative employment effects for higher skilled workers. For example, Franks and Mayer (1996) found higher resign rates for directors after hostile mergers for the UK. For the US and the UK, Hirshleifer and Thakor (1994) observed that mainly board members are displaced, and Bhagat, Shleifer, and Vishny (1990) reported that mostly white-collar workers are laid off after hostile M&A. Some further studies exist that examined management turnover after M&A: Martin and McConnell (1991) reported that 42% of top managers in targets have been replaced in the first, and 21% in the second post-merger year after hostile mergers. In friendly mergers, the replacement rate was 41% and 17%. This was confirmed by Franks and Mayer (1996). Kini, Kracaw, and Mianc (1995) found that 58% of the CEOs have been replaced after takeovers.

1.4.6 Effects on wages

In general, if M&A increase productivity, as found in several recent empirical studies listed in table 1.5, wages should be expected to rise if firms pay according to workers' marginal product. In addition, horizontal mergers can lead to changes in the structure of the product market. This may increase the amount of surplus available for wages. Vertical mergers may also generate a surplus because of reduced transaction costs or eliminated mark-ups (Conyon, Girma, Thompson, and Wright 2004). Negative effects can be expected if new management renegotiates about explicit and implicit employment contracts after mergers, as mentioned above (Shleifer and Summers, 1988).

Earlier studies exist from Brown and Medoff (1988) who found little positive effects on wages after M&A for small firms in the state of Michigan, USA. Another study is from McGuckin, Nguyen, and Reznek (1998). They identified an increase in wages after M&A for the US food and beverage manufacturing industry. Contrary, Lichtenberg and Siegel (1990) analyzed wage changes in the manufacturing sector of the USA, and they estimated a wage decrease in central offices, but only little effects in production plants.

Table 1.7: **Recent studies about effects of M&A on wages**

Country	Authors	Period	Unit	Merger sample	Control group	Estimation method	Change
Finland	Huttunen (2007)	1988-2001	Plants	284 foreign acquisitions of domestic plants	≈14.000 non-merged plants	Regression; DiD PSM	>0
Germany	Andrews, Bellmann, Schank, and Upward (2009)	2000-2004	Employees	Employer-employee data	Employees in non-merged plants	Regression	≈0
Indonesia	Lipsey and Sjöholm (2006)	1975-1999	Plant	1045 domestic and 1243 foreign takeovers	≈40.000 non-merged plants	Regression	>0
Sweden	Siegel and Simons (2008)	1985-1998	Employees	Employer-employee data for 19.000 plants; differentiated in full/ partial acquisitions, divestitures; related/ unrelated	Employees in non-merged firms	Regression	≥0
	Bandick (2011)	1993-2002	Firms	464 foreign acquisitions of manufacturing firms; MNEs/ non-MNEs; horizontal/ vertical mergers	≈4.000 non-merged firms	Regression; DiD PSM	≈0
UK	Conyon, Girma, Thompson, and Wright (2002b)	1989-1994	Firms	129 foreign mergers; 139 domestic mergers; horizontal/ vertical mergers	Random control sample of 642 firms	Regression	>0; Horizontal domestic M&A: <0
	Conyon, Girma, Thompson, and Wright (2004)	1979-1991	Firms	190 domestic related/ unrelated mergers in 140 manufacturing firms	236 non-merged firms	Regression	>0
	Girma and Görg (2007)	1980-1994	Plants	203 foreign acquisitions of domestic plants in the electronics and 100 in the food industry	Non-merged plants	DiD PSM	Acquisitions by US firms: >0; acquisitions by EU firms: ≈0
	Amess and Wright (2007)	1999-2004	Firms	1.350 leveraged buyouts (LBOs)	4.029 non-LBOs	Regression	<0
USA	McGuckin and Nguyen (2001)	1977-1987	Plants	≈20.000 merging plants in the manufacturing sector	≈300.000 non-merged plants	Regression	Typical plants: >0; Bigger plants: <0

Note: DiD PSM: difference-in-differences propensity score matching.

Results from newer studies about the impact of M&A on wages are presented in table 1.7. Despite little exceptions, most studies found that wages stay unchanged or increase after mergers. This is in line with results from newer studies about mostly positive productivity effects.

1.4.7 Further effects

In the literature on international trade and heterogeneous firms, mergers are discussed as a way to restructure industries. With an oligopoly framework, Neary (2007) showed that cross-border

M&A lead to a restructuring of industries with a specialization in the direction of comparative advantage, because low cost firms in one country buy high-cost firms in another country. In an empirical study about the role of M&A as a way of industry restructuring, Breinlich (2008) provided evidence that M&A are a channel for industrial restructuring if trade becomes freer. He also found that M&A transfers resources from less to more productive firms. Andrade and Stafford (2004) showed that mergers have both a contractionary and expansionary function in industrial restructuring, which changes over time. Especially in the 1970s and 1980s, industries with excess capacity were rationalized and restructured via mergers. Their findings supported the contractionary role of M&A which led to a more efficient allocation of resources and capacity within industries and the economy. In the 1990s, however, merger activity was highest in industries with high growth prospects and profitability, which underlined the expansionary role of M&A.

1.4.8 A different approach: event studies

Event studies are a different way to assess the success of mergers. The methodology goes back to Fama, Fisher, Jensen, and Roll (1969) and analyzes abnormal returns of stock prices around the announcement of an acquisition. This approach is based on the assumption of perfect capital markets which display all information available about past, present, and future at any time, and stock prices immediately adjust to new information. However, the analysis of share prices leads to some problems (Mueller, 2003a): when share prices change due to a merger, how should one know at which point of time this change occurs? And how to separate price changes caused by mergers from price changes caused by other factors? There are also justified doubts about the reliability of the efficiency of financial markets and their ability to correctly predict effects of mergers. In order to demonstrate the failure of financial markets to analyze mergers, Scherer (2002) presented two examples of abnormal positive returns for targets after a merger announcement, one of them ending in bankruptcy.

However, event studies are also not able to present an unambiguous picture about effects of M&A. For example, Jensen and Ruback (1983) summarized 13 earlier empirical studies and found that M&A increased stock value due to improved allocation. The observation period only covered several days or months around the merger announcement. In contrast, Tichy (2001) surveyed studies with long windows and concluded that results have a wide distribution with a negative mean after the acquisitions. The main findings were: there is a clear trend of declining abnormal bidder returns; cash financed mergers perform better than stock financed mergers; and bidders with a low book-to-market-ratio (so-called "glamor bidders") showed high abnormal announcement returns, but perform badly after the takeover.

There are studies that found increases in stock prices of target firms, but only little reaction of the acquirers' stock prices. For example, Andrade, Mitchell, and Stafford (2001) presented a survey about abnormal stock market returns of M&A deals of NASDAQ companies between 1973 and 1998, and they analyzed long run effects. They found that M&A created abnormal stock market returns of 0.4% for the acquiring firm, which is similar to other types of investments. However, there were large premiums for stockholders of target firms from 16 to 25%. Mueller (2003b) found that there were no abnormal returns to acquiring companies on average over a short period of time around the merger announcement. But over a longer observation period between one and three years, returns of acquirers declined on average relative to the market. Scherer (2002) summarized that after adding up gains for stockholders of targets and losses for stockholders of acquirers, the net effect of M&A depended on the relative size of the companies.

1.4.9 Why do the empirical results differ?

So far, empirical studies about effects of M&A do not present an unambiguous picture. Instead, studies even find opposing effects. There are some possible explanations for this phenomenon.[13]

- Different datasets can be an explanation for different results. They can differ from each other with respect to the number of observations included, the firms' size, the covered time periods, countries, industries, the availability and quality of control groups, etc. Moreover, different results can emerge from sample selection bias due to missing data (e.g. no data about smaller merging firms) (Schwert, 2000).

- Different empirical methods may estimate different effects. This is of great importance, because earlier studies do often not control for selection bias, while newer studies apply advanced econometric methods in order to identify causalities and control for selection effects. And even with newer techniques like matching,[14] results may differ because the performance of each step of the method is not yet standardized. Hence, different researchers may come to different results even if they use the same dataset (Angrist and Pischke, 2009).

[13] See Tichy (2001) for a more comprehensive discussion.
[14] See Caliendo (2006) for an introduction.

- Moreover, different types of mergers could also be a cause for ambiguous empirical results. For example, some studies focus on horizontal mergers (e.g. Pesendorfer, 2003), others on non-horizontal mergers (e.g. Bhuyan, 2002). There is some evidence that mergers between firms with similar products or markets perform better. For example, Ravenscraft and Scherer (1987) find that horizontal mergers are more profitable than vertical ones, and conglomerate mergers are the least profitable.

- There are studies that examined domestic M&A, while others examined cross-border M&A. Conn, Cosh, Guest, and Hughes (2001) found a better performance for cross-border M&A, while Black (2000) showed that domestic M&A perform better. The motivations behind mergers may be different and can cause different results, e.g. market extension in the case of cross-border M&A, or cost reductions in the case of domestic M&A (Capron, 1999). Similarly, effects from hostile mergers may also differ from friendly mergers (e.g. Martin and McConnell, 1991).

- Studies also differ with respect to the observation unit, i.e. between firm- (e.g. Bellak, Pfaffermayr, and Wild, 2006) and plant-level (e.g. Girma and Görg, 2004). Results may differ because plants are fully involved in a merger, whereas effects from M&A may disperse at firm-level if the firm is a multi-plant firm. Moreover, several studies analyzed effects for only one party involved in the merger, e.g. only for targets (e.g. Arndt and Mattes, 2010).

- The method of payment may also influence the success of mergers. Most studies point to a better performance of studies financed by cash (e.g. Conn, Cosh, Guest, and Hughes, 2001). Cash bids may be seen as an indicator of a good performance of the bidder (Loughran and Vijh, 1997).

1.5 Lessons learned and open research questions

The preceding sections provided a broad picture of research results about determinants and effects of mergers. I will try to draw some general conclusions, present important lessons that can be learnt, and finish with open research questions.

First, the large number of studies that found profit losses after mergers suggest that a high fraction of managers do not merge for profit maximizing reasons. Instead, other motives obviously play an important role, e.g. empire building, hubris, speculative motives, etc. This is in line with earlier surveys, e.g. from Tichy (2001) or Mueller (2003a).

Second, recent studies about effects on productivity, employment, and wages point to an increased importance of profit and efficiency maximizing motivations for mergers, and this is not fully in line with implications drawn from studies about effects on profits. In particular, there is evidence from newer studies that productivity increases after M&A, even if older studies show ambiguous effects. This may be due to newer research methods, different observation periods, or due to a larger proportion of firms than before that merge for efficiency reasons. Recent studies show a tendency towards employment losses after mergers. This also supports the relevance of profit and efficiency maximizing reasons for mergers, because employment losses are more likely if mergers are motivated by these reasons. However, studies do not point to mass layoffs after mergers, which is an often held fear by the public. Moreover, the fear of a large decrease in wages after mergers is also not confirmed by the data. Instead, the results point to positive effects. This is what could be expected if productivity effects are positive, and if wages are paid according to workers' productivity.

Third, a simple comparison of the empirical results is problematic. This is because studies highly differ from each other with respect to the underlying data, econometric methods, etc. Moreover, a differentiation between different types and motivations of mergers is necessary, because there is no typical merger (Tichy, 2001). Otherwise, we might compare the incomparable. With respect to hostile and friendly mergers, Morck, Shleifer, and Summers (1988) stated that "research results on friendly bids may have little to say about hostile bids, and vice versa".

Which research questions remain? The number of empirical studies about effects of M&A on several performance parameters is still to small to draw stylized facts (e.g. effects on R&D activities, investment, etc.).[15] This is also true for the differentiation between different types of mergers, i.e. horizontal, vertical, and in particular conglomerates. Moreover, while effects of cross-border M&A has gained much attention in the last years, almost no research work has focused on effects from domestic mergers, which account for a large fraction of all M&A.[16] In addition, studies often have a bias towards larger firms, and effects of smaller firms are less examined. This may also be a question of data availability. Finally, in addition to results from studies that are based on datasets with a large number of merged firms, studies analyzing individual mergers may provide further important information about the mechanisms at work, the motives and effects of mergers.

[15] Even if Tichy (2001) presented several stylized facts about effects of M&A on different performance parameters, I do not consider sufficient empirical support for several statements. This was also criticized by Lyons (2001).

[16] For Germany, around 50% of all M&A were domestic (Spanninger, 2012).

Chapter 2

As easy as one, two, three... A guide to performing propensity score matching with STATA

2.1 Motivation

Matching is a useful tool for situations in which the effect of a treatment on a group should be evaluated in comparison to the counterfactual situation in which the same group had not received the treatment. In other words, the intuition of the method is to mimic the situation in which the researcher is able to step back in time to observe the same group again, but now not participating in the treatment.

Matching has been applied in different fields of economics in order to discuss questions about self-selection and causality. For example, within international economics, Wagner (2002) was the first who used a matching approach to examine how a firm's productivity is affected if it starts to export. Later, Wagner (2007a) surveyed 45 microeconometric studies about exports and productivity for 33 countries which were published over the years 1995 to 2006, and several of these studies also applied matching strategies. In addition, studies exist that applied matching to evaluate the effects of firms' outward FDI on domestic performance. For example, Barba Navaretti, Castellani, and Disdier (2010) evaluated effects for Italian and French firms on several performance indicators, and Hijzen, Inui, and Todo (2007) investigated effects for Japanese firms. A similar study is from Jäckle and Wamser (2010) about German firms switching their status from national to multinationals. There are also studies that used matching techniques to study effects of foreign acquisitions on acquired firms' performance: with respect to productivity effects Girma, Thompson, and Wright (2006) for the UK, Salis (2008) for Slovenia etc.; with respect to effects on the firms' profits Bellak, Pfaffermayr, and Wild (2006) for Austria, Bertrand and Zitouna (2005) for France etc.; with respect to employment changes Girma (2005) for the UK, Arndt and Mattes (2010) for Germany etc.; and with respect to effects on wages Girma and Görg (2007) for the UK, or Bandick (2011) for Sweden etc.

The matching techniques were first applied in the areas of labor economics, mostly to evaluate the effectiveness of labor market programs (e.g. Stephan (2008) analyzed labor market programs in Germany). The method is also used in medical science to evaluate effects of therapies (e.g. see Austin (2007) for an evaluation of 47 articles published between 1996 and 2003 in the medical literature using propensity-score matching). There are further research fields in which the matching approach may be a useful strategy (e.g. see Brand and Halaby (2006) who investigated the effects of elite college attendance on educational and career achievement).

How is matching included in the econometrics literature? Generally, evaluation methods differ with respect to the data (Caliendo, 2006). The most compelling results are generated from experimental data, i.e. individuals are assigned randomly to a treatment, and randomization ensures the group of treated and the group of untreated[1] to have the same distribution of characteristics. That is, they only differ in their treatment status. Therefore, the data generates the correct missing counterfactual and eliminates the evaluation problem. This is the problem that arises because both states - treatment and no treatment - are not observable for the same individual at the same time. However, for most research questions only non-experimental data is available, i.e. data that is not generated by a controlled experiment. Since treated and untreated differ in more than their treatment status, a simple comparison of the outcome after treatment does not reveal the true impact of the treatment. In other words: since individuals do not randomly select in the treatment group, a comparison between treated and control suffers from selection bias. This selection problem requires the use of non-experimental estimation strategies. The performance of non-experimental estimators can be evaluated by using experimental data as a benchmark (e.g. LaLonde, 1986; Heckman, Ichimura, and Todd, 1997; Heckman, Smith and Clements, 1997).

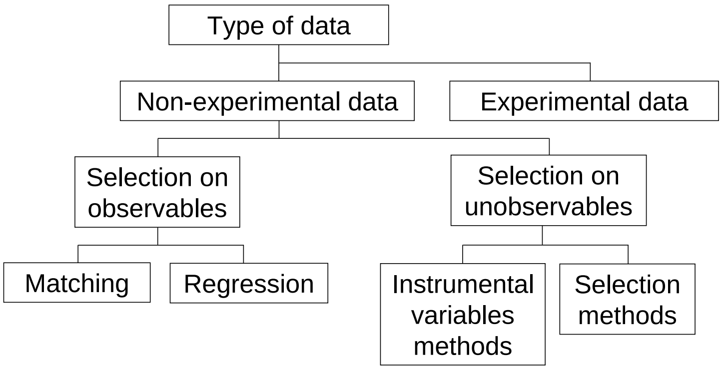
Figure 2.1 presents a classification of estimation strategies. As explained above, data is classified in experimental and non-experimental data. Methods dealing with non-experimental data are separated in two broad categories, depending in how they handle the selection effect (Caliendo, 2006). The first group of estimators is based on the assumption that the selection process is based on variables that are observable to the researcher. This assumption, also called "conditional independence assumption" (CIA), implies that systematic differences in the outcome between treated and untreated individuals with the same values for observable characteristics are caused by the treatment. There are two estimation methods that rely on the CIA and therefore, fall in this category: matching and linear regression. In order to justify the CIA, and to yield a good performance of the methods, rich data about observable characteristics and outcomes is required. Briefly explained, the major differences between matching and regression as estimation strategies when selection is on observables are, firstly, that matching is non- or semi-parametric, i.e. no functional form is required.[2] In contrast, linear regression requires a functional form. Hence, estimations from regression are biased if the functional form is not correctly specified. Secondly, matching requires finding a control to each treated observation, but for some treated, no controls may be found. Thus, the estimated effects from the matching

[1] I will use the terms "treated" and "participant" interchangeably, and I will also use the terms "control", or "non-participant" as synonyms for "untreated".

[2] If a parametric model for the estimation of the participation probability (e.g. with logit or probit models, see subsection 2.3.2) is combined with a non-parametric comparison of the outcomes, propensity score matching is semi-parametric. In contrast, exact (or cell) matching is completely non-parametric (Caliendo, 2006).

method only refer to those treated which find comparable controls.[3] In contrast, regression analysis produces estimates even if there are no untreated comparables to treated, i.e. the functional form fills in for the missing data and extrapolates for treated without comparable controls.[4] However, if selection on unobservable characteristics may also be relevant, estimators from the second group can be applied: these are instrumental variables (IV) methods, or selection models. Because this paper's focus is about matching as a strategy if selection is on observables, I will not further discuss these methods.[5]

Figure 2.1: **Estimation strategies**



Source: Caliendo (2006).

My paper adds to the existing literature about the implementation of propensity score matching (e.g. Caliendo (2006) and others). It is addressed to researchers not yet familiar with the method, and provides a stepwise description of how to perform propensity score matching with STATA, a widely used software program in econometrics. I use the module PSMATCH2, a matching program developed by Leuven and Sianesi (2003). The focus is on different practical questions: what is propensity score matching good for? How to perform propensity score matching in STATA? Which algorithm should be applied to match treated and controls to each other? How do the algorithms differ from each other? How can the quality of the process be assessed? And how do I know if the results are useful?

The first contribution in the evaluation literature which deals with the matching approach based on propensity scores came from Rosenbaum and Rubin (1983). The methodology was extended by Heckman, Ichimura, and Todd (1997) and Heckman, Ichimura, Smith, and Todd

[3] This so-called "common support condition" will be explained in greater detail later in this paper.

[4] See Angrist and Pischke (2009) and Caliendo (2006) for a more detailed discussion about similarities and differences between matching and regression.

[5] See Caliendo (2006) for more details about IV methods and selection models. In addition, see Caliendo and Hujer (2006) for a discussion of estimators if selection is on unobservables.

(1998). Dehejia and Wahba (1999, 2002) used data from National Supported Work Demonstration (NSW) data and different comparison groups from the Current Population Survey (CPS) and Panel Study of Income Dynamics (PSID) to perform matching with different matching algorithms, i.e. different ways how to match treated with controls. Their paper was a reply to LaLonde's (1986) critique on non-experimental estimators: LaLonde (1986) applied different standard evaluation estimators and showed that they produced different estimates. Different matching algorithms were introduced or analyzed by Heckman, Ichimura, and Todd (1997), Heckman, Ichimura, Smith, and Todd (1998), Imbens (2004), Smith and Todd (2005a), and Caliendo (2006). Moreover, there are contributions about certain aspects of the method like the estimation of the propensity score. This score determines which treated and controls will be matched to each other. For this, Rubin and Thomas (1996), Bryson, Dorsett, and Purdon (2002) as well as Augurzky and Schmidt (2001) discussed specifications of the estimation of the propensity score. Rosenbaum (2002) and Lechner (2008) addressed questions about robustness analysis. Reinowski (2008) discussed matching algorithms for small sample sizes. Imbens (2000), Hirano and Imbens (2004), and Imai and van Dyk (2004) extended the standard matching approach to continuous treatments.

This paper is closest to several contributions about practical implementation of the method. For example, Caliendo (2006) provided a comprehensive introduction to the method, but there are no descriptions about how to implement the method in a computer software. This is also true for Gensler, Skiera, and Böhm (2005) who described the theory about matching and provide a practical example. Becker and Ichino (2002) demonstrated their own developed software program *att** to perform propensity score matching in STATA. The intension of the paper of Essama-Nssah (2006) is similar to this contribution, but the author used the software EViews. In addition, there are also introductions to propensity score matching and its implementation in STATA with the software PSMATCH2 that are available in the form of downloadable internet presentations,[6] or as documents or handouts from research workshops.[7]

Despite the availability of several instructions and to the best of my knowledge, there is no paper that explains the implementation of the method in STATA[8] with the program PSMATCH2 step-by-step, i.e. with all relevant STATA-commands[9] and the respective outcomes, all displayed

[6]For example, see http://www.stata.com/meeting/germany10/germany10_sianesi.pdf [May 9th 2012] from Sianesi.

[7]For example, the workshops from Caliendo at the Institute of Employment Research (IAB) Nuremberg. However, the documents are not downloadable from the internet.

[8]I use STATA version 11.1.

[9]I present all necessary commands, but the reader should be familiar with the basics in STATA. Kohler and Kreuter (2008) provided an introduction to the software.

in tables and figures. Hence, my paper should provide a more comprehensive introduction than most existing internet presentations or workshop documents. Its objective is to make the researcher able to apply the method without any further teaching instructions. With respect to the description of the theoretical framework, I closely follow Caliendo (2006), but contrary to him, I do not go into details and restrict explanations to the most important aspects of the theory. The presented STATA-commands and the respective explanations in this paper are mostly guided by the descriptions in the help-file available from the program PSMATCH2. To demonstrate how the method works, I use a dataset with plants that merged and a control group of plants that had not merged, and I estimate the effect of mergers and acquisitions (M&A) on the plants' number of employees. The dataset is a combination of the IAB (Institute of Employment Research Nuremberg) Establishment Panel and the M&A DATABASE from St. Gallen.

The paper is structured as follows: the next section presents the theoretical framework of the method. Section 2.3 describes how to estimate the propensity score, and section 2.4 is dedicated to the explanation of several matching algorithms. Section 2.5 is about the quality of the procedure and presents robustness checks. Finally, section 2.6 concludes with a critical review of the matching method.

2.2 Theoretical framework

The focus of this section is to discuss the theory behind the matching method. The presentation closely follows Caliendo (2006). Matching is based on the theoretical framework developed by Roy (1951) and Rubin (1974), and it is known as the Roy-Rubin-Model. The central question in evaluating the effect of a treatment or of program participation is: what would have happened to an individual who received a treatment if it had not received the treatment? Of course, both states cannot be observed at the same time, and thus, the Roy-Rubin-Model is also known as the "potential outcome approach".

Average treatment effect on the treated (ATT): First, let Y be an outcome variable and D a treatment variable with $D = 1$ if the individual participated in a treatment, and $D = 0$ if there is no treatment. Hence, $Y_i(1)$ is the outcome if an individual i participated in the treatment D, and $Y_i(0)$ if the same individual did not participate in the treatment D. To measure the effect of the treatment it would be desirable to compare both outcomes for individual i, that is

$$\Delta = Y_i(1) - Y_i(0). \tag{1}$$

Of course, Δ cannot be identified because for one individual i only $Y_i(1)$ or $Y_i(0)$ can be observed at one point of time. The literature refers to this missing data problem as "fundamental evaluation problem", and the missing outcome as the "counterfactual outcome". The evaluation problem can be solved under the assumption of "unit homogeneity" (Holland, 1986). That is, the outcome of a participant will be compared to a non-participant who exhibits the same observable and non-observable characteristics before the treatment. However, the assumption of unit homogeneity can usually not be held because of heterogeneity between participants and non-participants with regard to their observable and unobservable pre-treatment characteristics. Therefore, the focus of interest has to switch from the observation of a single individual to the differences in mean values between groups of treated and untreated, because it is not possible to observe the counterfactual outcome for a single individual. This difference - known as the average treatment effect on the treated (ATT) - is

$$ATT = E[Y(1)|D=1] - E[Y(0)|D=1]. \tag{2}$$

However, the mean outcome of a treated without treatment, $E[Y(0)|D=1]$, can also not be observed. Therefore, the evaluation problem is to identify an adequate control group for the participants. Taking $E[Y(0)|D=0]$ for $E[Y(0)|D=1]$ may lead to a selection bias if individuals are not randomly assigned to the treatment, i.e. every individual does not face the same ex ante probability of being treated. The resulting estimate would then include a selection bias and can be written as

$$E[Y(1)|D=1] - E[Y(0)|D=0] = ATT + \underbrace{E[Y(0)|D=1] - E[Y(0)|D=0]}_{Selection\ bias}. \tag{3}$$

In non-experimental data individuals do not assign randomly to treatment. This implies that $E[Y(0)|D=1] - E[Y(0)|D=0]$ can be expected to be different from zero.

The matching method offers a way to estimate the ATT without selection bias. Therefore, to each individual of the treatment group an individual of the control group with identical - or at least similar - pre-treatment characteristics will be assigned. It should be clear that the identification of a causal effect is only possible if all pre-treatment characteristics, which are relevant for the treatment decision, are included. If there is a difference in the mean outcome between the treated and control group after treatment, it can be ascribed to the treatment.

Conditional independence assumption (CIA): To make a matching approach work, some important assumptions have to hold. First, the "conditional independence assumption" (CIA) states that - conditioning on the values of a set of observable characteristics X which are not affected by treatment - the outcome of both groups would be the same in the absence of treatment (Lechner, 1999):

$$Y(0), Y(1) \perp D | X, \ \forall X \tag{4}$$

with $\perp$ indicating independence.[10] This assumption is called "unconfoundedness" (Rosenbaum and Rubin, 1983), but the terms "conditional independence assumption" (Lechner, 1999) and "selection on observables" (Heckman and Robb, 1985) mean the same.[11]

Taking the mean, the CIA allows stating that

$$E[Y(0)|D = 1, X] = E[Y(0)|D = 0, X], \tag{5}$$

that is, given the CIA, the selection bias from equation 3 disappears after conditioning on the covariates X.

Propensity score: Matching is based on a set of several observable pre-treatment characteristics X. Even with a small number of characteristics and its different values it becomes difficult to find matching partners with equal characteristics. Then, treated individuals may remain unmatched. Rosenbaum and Rubin (1983) showed that the use of a propensity score $P(X)$ as a single index is also sufficient in order to reduce a "potentially high dimensional matching problem" (Heckman, Ichimura, Smith, and Todd, 1998). The propensity score $P(X)$ is a measure of the probability of participation for each individual conditional on observed characteristics X. The propensity score is then estimated from a probit or logit model, and the matching of individuals is then based on the propensity score of treated and untreated individuals. The ATT can now be estimated:

$$ATT = E[Y(1)|D = 1, P(X)] - E[Y(0)|D = 0, P(X)]. \tag{6}$$

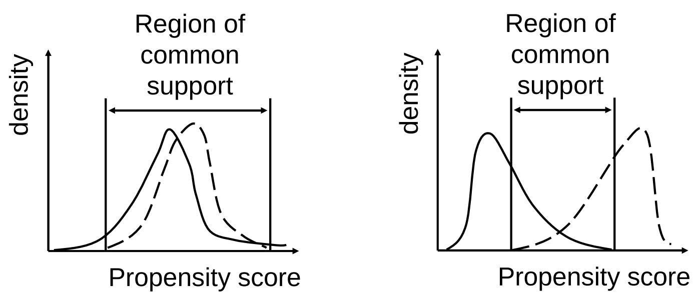
The self-selection bias is eliminated, and differences in outcomes can be ascribed only to the treatment.

[10] For the estimation of ATT it is sufficient to assume that $Y(0) \perp D | X, \ \forall X$ (Smith and Todd, 2005a).
[11] I will use these terms interchangeably in this paper.

Common support: The "common support condition" ensures that the propensity scores of both groups overlap and all participants have a counterpart in the control group. This ensures that only individuals which are sufficiently similar to each other will be matched. The assumption is stated as $0 < P(D = 1|X) < 1$. It implies that an individual has a positive probability of being both in the treated and control group (Heckman, LaLonde, and Smith, 1999), and it also implies that for every treated a control can be found as a matching partner.[12] The assumptions of unconfoundedness and common support are called "strong ignorability" (Rosenbaum and Rubin, 1983).

Figure 2.2 presents a graphical illustration for a situation with similar distributions of propensity scores in both groups and a corresponding large overlap, and another situation with different distributions and a small overlap.[13]

Figure 2.2: **Different distributions of propensity scores and region of common support**



Note: the figure present the distributions of propensity scores and the regions of common support; solid line: distribution for treated; dotted line: distribution for controls; left figure: large overlap of distributions; right figure: small overlap of distributions. Source: Gensler, Skiera, and Böhm (2005).

Further assumptions: Another assumption, the "stable unit treatment value assumption" (SUTVA), states that the participation of a treated does not influence the treatment decision and outcome variable of other treated (Rubin, 1990). Finally, individuals should not change their behavior because of an anticipation of a treatment, because matching results may then be biased. This is known as the Ashenfelter's Dip.[14]

[12] A probability of 0 or 1 would imply that given the covariates X the individual either never or always participates in the treatment and there are no counterparts in the other group (Heckman, Ichimura, and Todd, 1997). For the estimation of ATT it is sufficient to assume that $P(D = 1|X) < 1$. See Smith and Todd (2005a) for a further discussion.

[13] I will discuss consequences of different distributions of propensity scores later.

[14] Ashenfelter (1978) found that the employment situation of individuals worsens shortly before they participate in a labor market program. For more details see Hagen and Steiner (2000), and for a

2.3 How to start and estimate the propensity score

In this section, I first introduce the data I use, the software STATA, and the matching program PSMATCH2. Then, I describe the estimation of the propensity score. I explain how it can be computed from binary treatment models, and how the robustness of the models can be assessed. I also discuss the implications of the estimated scores for the matching process. For a better understanding, I present an example and show the corresponding results.

2.3.1 The data and the software

The dataset used for the following estimations is a combination of the IAB Establishment Panel and the M&A DATABASE St. Gallen. It contains plants that merged between 1996 and 2005, and control plants not involved in any M&A activity since 1980. All plants are located in Germany. The matching process should evaluate the effect of M&A on employment, which is a research question that has been extensively discussed for several decades with ambiguous results.[15] However, the focus of this paper is on the performance of the method, i.e. the empirical results generated in this paper will be of minor interest. I use an observation period of three years: pre-merger characteristics are all measured in $t = 1$, the merger occurs in t=2, and the impact on employment will be measured in t=3. The following variables are of relevance for the matching process: a treatment indicator for M&A (value one if treated and zero otherwise), dummies for different size categories, sector dummies, dummies for different legal forms (partnership, individually-owned, public, etc.; limited; limited by shares), and a dummy for the location of plants in East Germany. The dataset contains 1817 treated observations and 581 controls.[16]

To start the matching procedure, a matching software has to be installed in STATA. In this paper, I use the PSMATCH2 module developed by Leuven and Sianesi (2003) which is a comprehensive and user-friendly program. Most empirical studies applying matching use this software, but there also exist other programs, e.g. `attnd`, `attnw`, `attk`, `attr`, `attrw`, and `atts` from Becker and Ichino (2002), `nnmatch` from Abadie, Drukker, Herr, and Imbens (2004), or `cem` from Blackwell, Iacus, King, and Porro (2009).

The package PSMATCH2 will be installed with the STATA-command

comprehensive description of Ashenfelter's Dip and its impacts on estimations see Hujer, Caliendo, and Radic (2001).

[15] See Mueller (2003a) for a survey.

[16] In contrast to this dataset, the number of controls usually clearly exceeds the number of treated in most datasets.

```
ssc install psmatch2
```

A short description of the method and the relevant syntaxes is given by the command

```
help psmatch2
```

The module PSMATCH2 performs a variety of matching algorithms, and it also incorporates several features like the graphical illustration of the propensity score for the treatment and control group in a histogram, or the storage of the ATT to allow bootstrapping in order to reestimate standard errors of the treatment effect. Furthermore, it enables the researcher to assess the matching quality with respect to the balancing of the covariates. Sections 2.4 and 2.5 will provide more detailed information.

2.3.2 The estimation of the propensity score

The propensity score is the probability of receiving a treatment D in year t conditional on a set of characteristics X for individual i, measured prior to the treatment in $t-1$:

$$P(D_{it}=1)=F(X_{it-1}) \tag{7}$$

With respect to this dataset, the propensity score describes the probability that a plant merges conditioning on a set of pre-merger characteristics X. For the estimation of the propensity score any standard probability model like probit or logit models can be used, and the dependent variable presents the participation decision (Dehejia and Wahba, 2002). That is, for each observation, the models estimate the probability of merging.[17] The results from logit and probit models are usually similar, and thus, the choice between both models is of minor importance.[18] In this paper I apply a probit model with a dummy variable for merger activity as dependent variable (with value one if plants merged and zero otherwise), and a set of explaining variables.

The choice of explanatory variables has to be consistent with economic theory and has to ensure that the CIA holds (Smith and Todd, 2005a). As noted above, all relevant characteristics X which influence the treatment decision (here: M&A) as well as the outcome variable (here:

[17] See Gensler, Skiera, and Böhm (2005) or Dehejia and Wahba (2002) for a formal presentation of logit and probit models in order to estimate the propensity score for each observation.

[18] Logit models are based on a logistic distribution function while probit models are based on a standard normal distribution function. The main difference is that the logit distribution has more density mass on the bounds (Caliendo, 2006). See Verbeek (2005) for a further discussion of logit and probit models.

number of employees) have to be included in the model. Rubin and Thomas (1996) suggested that variables should only be excluded if they are not related to the outcome or if they have no relevance. In contrast, Heckman, Ichimura, Smith, and Todd (1998), Augurzky and Schmidt (2001), and Bryson, Dorsett, and Purdon (2002) recommended a smaller set of variables. The latter argue that the inclusion of insignificant variables leads to a less exact estimation of the propensity score. However, it should be clear that the objective of the matching process is to balance the covariates and not to obtain an exact estimation of the propensity score (Caliendo, 2006). The chosen variables have to be fixed over time, or they have to be measured before the treatment.

With respect to the number of observations Hosmer and Lemeshow (2000) recommended more than 100 observations to obtain meaningful results from the regression. If there are categories due to dummy variables for nominal or ordinal scaled independent variables (e.g. employment size categories), the number of observations in each category should be larger than 25. Moreover, no multicollinearity (Menard, 2001) and autocorrelation (Aldrich and Nelson, 1984) should exist in logistic regression models.

The following probit regression model includes only a small set of explaining variables as suggested by Rubin and Thomas (1996). The STATA-command is:

```
probit DUMMY_MA SIZE_CAT_T1_2-SIZE_CAT_T1_6 SECTOR_T1_2-SECTOR_T1_16
LEGAL_FORM_T1_2 LEGAL_FORM_T1_3 EAST_T1
```

`probit` is the STATA-command to start a probit regression.[19] It has to be followed by the dependent variable, here a dummy for M&A, `DUMMY_MA`. The following variables are all explaining variables, and they have to be consistent with economic theory. For this model, the choice of the variables is influenced by the literature about determinants of M&A,[20] and they are all measured prior to the merger in $t = 1$: `SIZE_CAT_T1_2` to `SIZE_CAT_T1_6` are dummies for different size categories,[21] `SECTOR_T1_2` to `SECTOR_T1_16` are dummies for sectors, `LEGAL_FORM_T1_2` and `LEGAL_FORM_T1_3` are dummies for legal forms, and `EAST_T1` is a dummy for location in East Germany.

Table 2.1 reports the results from the probit regression. Regression parameters are generated

[19] To perform a logit instead of a probit regression, simply exchange the command `probit` with `logit`.

[20] See Conyon, Girma, Thompson, and Wright (2002a, 2002b), Harris and Robinson (2002), Margolis (2006a), Girma and Görg (2007), and others for a discussion about determinants of M&A.

[21] `SIZE_CAT_T1_1` is the reference category and contains the smallest plants. Coefficients of the other size dummies refer to this group. This applies also to other dummy variables, and thus, avoids that exact multicollinearity arises (Verbeek, 2005).

Table 2.1: **Probit regression**

```
. probit DUMMY_MA SIZE_CAT_T1_2-SIZE_CAT_T1_6 SECTOR_T1_2-SECTOR_T1_11
LEGAL_FORM_T1_2-LEGAL_FORM_T1_3 EAST_T1

Iteration 0:   log likelihood = -1328.0495
Iteration 1:   log likelihood = -820.82145
Iteration 2:   log likelihood = -757.98407
Iteration 3:   log likelihood = -754.50846
Iteration 4:   log likelihood = -754.48562
Iteration 5:   log likelihood = -754.48562

Probit regression                                 Number of obs   =       2398
                                                  LR chi2(18)     =    1147.13
                                                  Prob > chi2     =     0.0000
Log likelihood = -754.48562                       Pseudo R2       =     0.4319

------------------------------------------------------------------------------
    DUMMY_MA |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
SIZE_CAT_T~2 |  -.0987562   .1369382    -0.72   0.471      -.36715    .1696377
SIZE_CAT_T~3 |  -.2097664   .1412184    -1.49   0.137    -.4865494    .0670165
SIZE_CAT_T~4 |   .2507293   .1252634     2.00   0.045     .0052176     .496241
SIZE_CAT_T~5 |   .6840995   .1524464     4.49   0.000     .3853101     .982889
SIZE_CAT_T~6 |   2.088776    .157313    13.28   0.000     1.780448    2.397103
 SECTOR_T1_2 |   .0822164   .4442876     0.19   0.853    -.7885714    .9530042
 SECTOR_T1_3 |  -.9571824   .2741334    -3.49   0.000    -1.494474   -.4198909
 SECTOR_T1_4 |  -1.299743   .2659028    -4.89   0.000    -1.820903   -.7785834
 SECTOR_T1_5 |  -1.153739   .1859073    -6.21   0.000    -1.518111   -.7893676
 SECTOR_T1_6 |  -1.378838    .180157    -7.65   0.000     -1.73194   -1.025737
 SECTOR_T1_7 |  -1.070766   .2834113    -3.78   0.000    -1.626242   -.5152899
 SECTOR_T1_8 |  -1.088695   .1872647    -5.81   0.000    -1.455727    -.721663
 SECTOR_T1_9 |  -.3572387    .220214    -1.62   0.105    -.7888502    .0743729
SECTOR_T1_10 |  -1.458521   .8776002    -1.66   0.097    -3.178586    .2615435
SECTOR_T1_11 |  -1.172693   .2044493    -5.74   0.000    -1.573406   -.7719794
LEGAL_FOR~_2 |  -.1516684   .1119252    -1.36   0.175    -.3710377    .0677008
LEGAL_FORM~3 |   1.595019   .1689148     9.44   0.000     1.263952    1.926086
     EAST_T1 |   .2989708    .092083     3.25   0.001     .1184915    .4794501
       _cons |   .8438046   .2017243     4.18   0.000     .4484321    1.239177
------------------------------------------------------------------------------
```

Source: IAB Establishment Panel, M&A DATABASE St. Gallen.

with the maximum likelihood method (ML).[22] One test to evaluate the goodness-of-fit of the binary choice model is the likelihood-ratio test (LR chi2). The principle of this test is to compare a model in which all coefficients are set to zero except the constant, with a model including all variables. For both models a so-called loglikelihood value will be calculated. The larger the difference of these values, the higher the explanation of the independent variables of the model. The difference should exceed the value of the χ^2-distribution for the number of degrees of freedom (here: 18). The respective prob-value describes the probability of coefficients to be zero (Backhaus, Erichson, Plinke, and Weiber, 2010; Rohrlack, 2009). In this regression, the difference of the coefficients to zero is statistically significant.

Another important measure of fit is the McFadden's R^2. It measures how much of the variation of the dependent variable is explained by the regression.[23] Higher values indicate a better fit: values above 0.2 are considered as acceptable, values of 0.4 and higher can be considered as good (Backhaus, Erichson, Plinke, and Weiber, 2010; Rohrlack, 2009; Krafft, 1997). Here, a pseudo-R^2 of 0.432 is sufficiently high.

Finally, the coefficients need some attention. The second column in table 2.1 reports standard errors as a measure of the accuracy of the estimated coefficients. High standard errors

[22] For a detailed description see Verbeek (2005), or Greene (2011).

[23] Cox and Snell, and Nagelkerke provided further pseudo-R^2 measures. For more information see Backhaus, Erichson, Plinke, and Weiber (2010) and Rohrlack (2009).

indicate that estimated coefficients are not precise and reliable. The third and fourth columns report z-values and respective p-values. The z-values are the ratio of the coefficients to the standard errors. Based on z-values, the associated p-values indicate if the regression coefficients are statistically significantly different from zero, given the other variables. In particular, if p-values are equal or smaller than 0.1, 0.05, or 0.01, the coefficients are statistically significant at the 10%, 5%, or 1%-level. The last column presents the 95%-confidence interval: 95% of several confidence intervals, calculated from different samples of the same population, will include the true value of the coefficient. If explaining variables influence the dependent variable, the respective confidence intervals should neither change signs nor include the value zero.

Now, for each observation a propensity score is estimated from the probit model, and it expresses the probability to merge. The variable `pscore`, that will be generated next, simply includes the propensity score for each observation:

```
predict pscore, p
```

The distribution of the propensity scores can be analyzed by the STATA-command

```
bys treated:  summarize pscore
```

and results are shown in table 2.2. It lists the range of propensity scores for both groups, which is from 0.185 to 0.962 for controls, and from 0.185 to 0.999 for treated.[24]

Table 2.2: **Distribution of propensity scores**

```
. bys DUMMY_MA: sum pscore

-> DUMMY_MA = Controls

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
      pscore |       581    .4319791    .1788804   .1850013   .9623671

----------------------------------------------------------------------
-> DUMMY_MA = Treated

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
      pscore |      1817    .8622568    .2234217   .1850013   .9999995
```

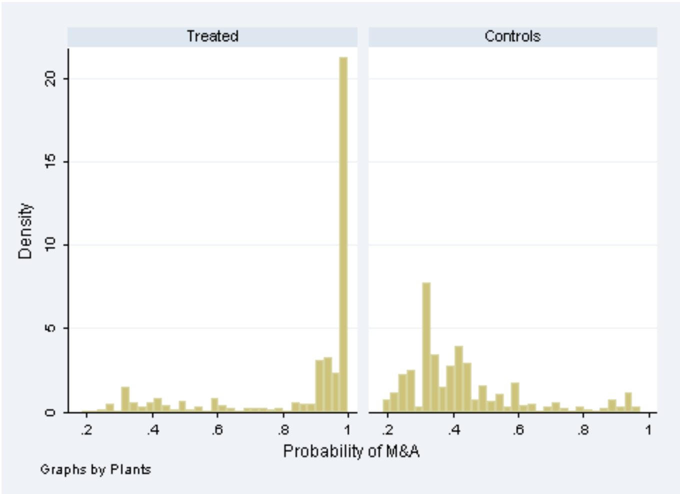
In addition to table 2.2 it is helpful to analyze the distribution of propensity scores among treated and controls with a graphical illustration. Figure 2.3 is generated with command

[24] The ranges of propensity scores start with exactly the same value in both groups. That is, there are obviously plants in both groups which have identical characteristics, and this leads to the same probability for M&A.

```
hist pscore, by(treated)
```

Ideally, the distribution of propensity scores is similar in both groups. If there are many observations for certain ranges of the propensity score, a sufficiently high number of matching partners with similar scores can be found. Instead, if the distribution strongly differs, matching partners with similar propensity scores are scarce for certain ranges. The matching quality suffers if treated with high propensity scores are matched to controls with low propensity scores and vice versa. Figure 2.3 shows an asymmetric distribution of propensity scores for this paper's data.[25]

Figure 2.3: **Graphical distribution of propensity scores**



As already discussed in section 2.2, the estimation of the ATT is only defined in the region of common support (Heckman, LaLonde, and Smith, 1999), i.e. there has to be a potential matching partner for each observation. For this, observations with propensity scores higher than the maximum or lower than the minimum of the other group should be deleted. For example, if the propensity scores in one group range from 0.1 to 0.7, and propensity scores in the other group range from 0.3 to 0.9, the region of common support is from 0.3 to 0.7. However, if propensity scores in both groups range from 0.01 to 0.99, but in one group there are no observations with propensity scores between 0.5 and 0.7, the maximum and minimum comparison fails. A way to deal with this situation is to impose a common support condition by trimming as suggested by Smith and Todd (2005a). According to this approach, a density for values of propensity scores is estimated, and the region of common support consists of those values for propensity scores with a positive density within the distribution of both groups. If the

[25]There are many treated observations with propensity scores close to one. One explanation may be that these observations all belong to sectors that were strongly affected by M&A.

density is exactly zero, the propensity scores will be excluded. Additionally, to ensure that the densities are strictly positive, a further percentage of propensity score values - determined by the researcher - with a very low density are also excluded. It is highly recommended to visually analyze the distribution of the propensity scores, as done in figure 2.3.

The propensity score is now estimated and its distribution among both groups is known. Now, there are several ways how to assign controls to treated in order to create matching pairs. Therefore, the focus of the next section is on the description of different matching algorithms.

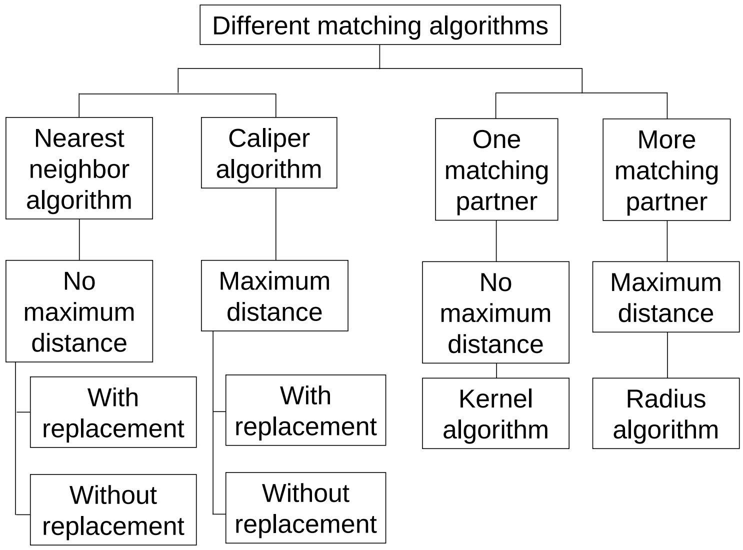
2.4 Choosing a matching algorithm

Since the propensity score is a continuous variable, it seems not very likely to find enough observations out of both groups with exactly the same score. Thus, different algorithms were developed in order to match pairs. For the choice of a matching algorithm, a measure of proximity has to be defined. That is, for each treated a neighborhood will be defined which consists of observations of the control group which are sufficiently close to the treated in terms of propensity scores. Then, controls are assigned with a specific weight to a participant. Thus, different neighborhoods and different weights lead to different matching algorithms (Smith and Todd, 2005a). In this section, nearest neighbor matching, caliper matching, radius caliper matching, and kernel matching will be described in detail. These algorithms are used most often (e.g. Girma and Görg, 2007; Barba-Navaretti, Castellani, and Disdier, 2010, and others) in empirical studies which apply matching. Algorithms can be classified according to the number of matching partners, and according to a maximum distance between treated and controls. Figure 2.4 presents a graphical classification of these algorithms.

In general, the choice for specific matching algorithms should be of minor importance and different algorithms should lead to similar results if the sample size is large enough (Smith and Todd, 2005a). However, if the sample is smaller the choice of the algorithm becomes more important (Heckman, Ichimura, and Todd, 1997) because of the trade-off between bias and variance of the estimators. Therefore, there is no superior algorithm which is recommended in all situations. The structure of the data, especially the sample size, plays a crucial role for the correct choice of algorithms.

Heckman, Ichimura, Smith, and Todd (1998), Smith and Todd (2005a), Imbens (2004), and Caliendo (2006) presented a comprehensive overview of matching algorithms. Here, I present a brief formal description of the algorithms, using a notation similar to Caliendo (2006). The general notation will be introduced: I_1 and I_0 are the samples of treated and controls. The

Figure 2.4: **Classification of different matching algorithms**



Source: Gensler, Skiera and Böhm (2005).

effect of a treatment of each observation $i \in I_1$ is estimated by comparing its outcome with a weighted average outcome of control observations $j \in I_0$ in the following way:

$$\Delta = \frac{1}{N_1} \sum_{i \in I_1} \left[Y_i^1 - \sum_{j \in I_0} W_{N_0}(i,j) Y_j^0 \right]. \tag{8}$$

N_1 and N_0 are the numbers of observations in the treatment group I_1 and control group I_0. The matching estimators differ with respect to the weights attached to controls. Hence, $W_{N_0}(i,j)$ is the weight of a control j from the control group in order to construct the counterfactual for treated observation i of the treatment group. For each treated individual i the sum of the weights of all controls j must equal one: $\sum_j W_{N_0}(i,j) = 1, \ \forall\, i$.

In addition, matching estimators also differ with respect to the neighborhood $C(P_i)$ they define for each treated i. The neighborhood of treated i with its propensity score P_i includes all those controls $j \in I_0$ which have propensity scores P_j that lie within this neighborhood: $P_j \in C(P_i)$. Controls $j \in I_0$ that were matched to a treated i are in the set A_i with $A_i = \{j \in I_0 | P_j \in C(P_i)\}$. Finally, Y_i^1 is the outcome of a treated individual i and Y_j^0 is the outcome of a control j.

2.4.1 Nearest neighbor matching

The first algorithm presented in this paper is a nearest neighbor algorithm. This algorithm assigns the control j to a treated i which is closest in terms of propensity scores P_i and P_j. Formally, the neighborhood for treated i is defined as

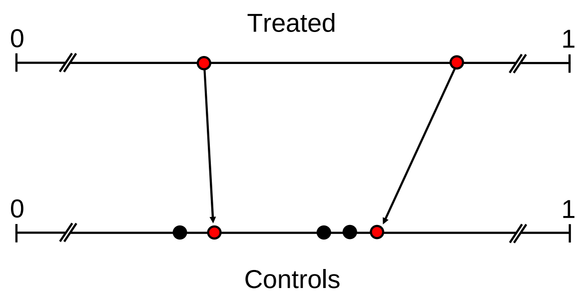
$$C^{NN}(P_i) = \min_j \|P_i(X) - P_j(X)\|, j \in N_0 \tag{9}$$

with $"\|"$ as a norm or length.[26] Control j with a propensity score of $P_j(X)$ which is closest to treated i has the weight

$$W_{N_0}^{NN}(i,j) = \begin{cases} 1 & \text{if } \|P_i - P_j\| = min_j \|P_i - P_j\| \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

If only one single control j is matched to a treated i, the matching is also called "one-to-one-matching". Table 2.5 presents a graphical illustration about the matching process in nearest neighbor matching.

Figure 2.5: **Graphical illustration of nearest neighbor matching**



Note: the two horizontal lines represent the range of propensity scores between zero and one, the upper line for treated, the lower line for controls. The circles represent single observations, each with a certain propensity score. Red circles indicate that observations are matched. Source: own illustration.

The respective STATA-command is

```
psmatch2 DUMMY_MA, outcome(SIZE_T3) pscore(pscore)
```

with `SIZE_T3` as the outcome variable. This command also generates several additional useful

[26] This simply means the distance between treated i and control j in terms of propensity scores.

variables.[27]

The results for this nearest neighbor algorithm are shown in table 2.3. Mean values of the outcome variable for both treated and control group, the difference, and the corresponding standard error as well as the t-statistics with regard to the difference are reported for the unmatched sample. In the row below, calculations for the matched sample are reported. The ATT estimates the effect of M&A on the outcome variable. Here, the number of employees after M&A is 1,151 higher compared to plants that did no merge, and the difference is statistically significant.

Table 2.3: **Nearest neighbor matching**

```
. psmatch2 DUMMY_MA, outcome(SIZE_T3) pscore(pscore)
There are observations with identical propensity score values.
The sort order of the data could affect your results.
Make sure that the sort order is random before calling psmatch2.
----------------------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference         S.E.   T-stat
----------------------------+-----------------------------------------------------------
         SIZE_T3  Unmatched | 1612.47111   135.500861   1476.97025   178.450092     8.28
                        ATT | 1612.47111   460.669785   1151.80132   149.018271     7.73
----------------------------+-----------------------------------------------------------
Note: S.E. for ATT does not take into account that the propensity score is estimated.

           | psmatch2:
 psmatch2: |   Common
 Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |       581 |       581
   Treated |     1,817 |     1,817
-----------+-----------+----------
     Total |     2,398 |     2,398
```

However, at this point some questions about the reliability of this estimation are necessary. This algorithm does not impose a common support condition, and this may lead to bad matches. Moreover, the algorithm allows replacing controls, i.e. a single control j can be used more than once as a matching partner. In general, this may be useful if the distribution of propensity scores strongly differs between both groups which is mostly the case in smaller samples (Caliendo, 2006). This increases the matching quality on average and decreases the bias because matching pairs are better on average. However, replacement increases the variance of the estimator, because less information is used when constructing the counterfactual for each treated individual (Smith and Todd, 2005a). Moreover, a replacement option may also have the

[27] I briefly describe these generated variables: `_treated` is a dummy variable with value one for treated and zero for control observations, `_support` is also a dummy with value one if the observation is within the region of common support, and zero if the observation is off the region of common support. Variable `_pscore` is the estimated propensity score (if the propensity score is estimated with `pscore()`, variable `_pscore` is only a copy), `_outcome_variable` is a variable for the outcome variable, here `_SIZE_T3`. The following generated variables all apply to nearest neighbor matching: `_weight` shows the frequency of how often the observation is used as a match, `_id` creates a new identification number for matched observations. For every treated `_n1` shows the observation number of the matched control if the data is sorted by `_id`. `_nn` stores the number of matched controls, and `_pdif` shows the absolute distance from treated to controls in terms of propensity scores. See `help psmatch2` for further explanation.

negative consequence that a small number of controls is used very often. This can be checked with the command

```
tabulate _weight DUMMY_MA
```

Table 2.4 enables the researcher to see how often a single control is used as a match. Here, one single control was matched 1,008 times with a treated, and only 78 out of 581 controls were actually used as a matching partner. These numbers are not a good basis for a reliable estimation of the ATT.

Table 2.4: **Nearest neighbor matching (weights of controls)**

. tab _weight DUMMY_MA

psmatch2: weight of matched controls	Plants Controls	Treated	Total
1	7	1,817	1,824
2	13	0	13
3	9	0	9
4	5	0	5
5	5	0	5
6	2	0	2
7	3	0	3
8	7	0	7
9	4	0	4
10	3	0	3
11	3	0	3
12	3	0	3
14	1	0	1
15	1	0	1
17	1	0	1
18	1	0	1
19	1	0	1
22	1	0	1
26	1	0	1
27	1	0	1
28	1	0	1
29	1	0	1
41	1	0	1
81	1	0	1
143	1	0	1
1008	1	0	1
Total	78	1,817	1,895

Hence, the following modification of the algorithm does not allow replacing controls and imposes a common support condition. The commands in STATA are

```
psmatch2 DUMMY_MA, outcome(SIZE_T3) pscore(pscore) noreplacement common
```

```
tabulate _weight DUMMY_MA,
```

The results for this algorithm are shown in table 2.5. The ATT now clearly differs from the ATT estimated in table 2.3 with a difference in employment between treated and controls of only 254. Moreover, due to the implementation of a region of common support a large fraction of treated is not matched.

The histogram in figure 2.3 above showed that there are almost no observations in several regions of the propensity score. In this case, a trimming condition may be an alternative to

Table 2.5: **Nearest neighbor matching without replacement and with a common support condition**

```
. psmatch2 DUMMY_MA, outcome(SIZE_T3) pscore(pscore) noreplacement common
There are observations with identical propensity score values.
The sort order of the data could affect your results.
Make sure that the sort order is random before calling psmatch2.
----------------------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference         S.E.   T-stat
----------------------------+-----------------------------------------------------------
         SIZE_T3  Unmatched | 1612.47111   135.500861   1476.97025   178.450092     8.28
                        ATT | 389.681583   135.500861   254.180723    32.199705     7.89
----------------------------+-----------------------------------------------------------
Note: S.E. for ATT does not take into account that the propensity score is estimated.

 psmatch2: |   psmatch2: Common
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        581 |       581
   Treated |     1,236        581 |     1,817
-----------+----------------------+----------
     Total |     1,236      1,162 |     2,398

. tab _weight DUMMY_MA

 psmatch2: |
 weight of |
   matched |        Plants
  controls |  Controls    Treated |     Total
-----------+----------------------+----------
         1 |       581        581 |     1,162
-----------+----------------------+----------
     Total |       581        581 |     1,162
```

impose a common support condition (see subsection 2.3.2). The STATA-command for trimming the observations by those 10% observations which exhibit the lowest estimated density is

```
psmatch2 DUMMY_MA, outcome(SIZE_T3) pscore(pscore) trim(10)
```

The results are shown in table 2.6. Even if the ATT of 318 differs from the ATT of 254 from table 2.5 by approximately one fourth, they are both similar in their magnitude, compared to the ATT of 1,151 from table 2.3.

Table 2.6: **Nearest neighbor matching without replacement and with trimming**

```
. psmatch2 DUMMY_MA, outcome(SIZE_T3) pscore(pscore) noreplacement trim(10)
There are observations with identical propensity score values.
The sort order of the data could affect your results.
Make sure that the sort order is random before calling psmatch2.
----------------------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference         S.E.   T-stat
----------------------------+-----------------------------------------------------------
         SIZE_T3  Unmatched | 1612.47111   135.500861   1476.97025   178.450092     8.28
                        ATT | 454.335628   135.500861   318.834768    35.876856     8.89
----------------------------+-----------------------------------------------------------
Note: S.E. for ATT does not take into account that the propensity score is estimated.

 psmatch2: |   psmatch2: Common
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        581 |       581
   Treated |     1,236        581 |     1,817
-----------+----------------------+----------
     Total |     1,236      1,162 |     2,398
```

A further modification within the nearest neighbor algorithm is to use more than one matching partner. This algorithm is mostly meaningful in large samples, and allows generating more information. It reduces the variance of the estimator, but the bias increases due to worse matching pairs (Caliendo, 2006). The modification is called "oversampling". m controls will be chosen as neighbors for each treated i, and each matched control within the set A_i will be assigned with the same weight $1/m$, and all others receive weight zero:[28]

$$W^{NNO}(i,j) = \begin{cases} \frac{1}{m} & \text{if } j \in \mathrm{A}_i \\ 0 & \text{otherwise.} \end{cases} \tag{11}$$

The command (here for $m = 5$ neighbors) including a common support condition is

```
psmatch2 DUMMY_MA, outcome(SIZE_T3) pscore(pscore) neighbor(3) common
```

and table 2.7 reports the results.

Table 2.7: **Nearest neighbor matching with five neighbors**

```
. psmatch2 DUMMY_MA, outcome(SIZE_T3) pscore(pscore) neighbor(5) common
There are observations with identical propensity score values.
The sort order of the data could affect your results.
Make sure that the sort order is random before calling psmatch2.
----------------------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference         S.E.   T-stat
----------------------------+-----------------------------------------------------------
         SIZE_T3  Unmatched | 1612.47111   135.500861   1476.97025   178.450092     8.28
                        ATT | 653.891727   270.539173   383.352555   42.2143558     9.08
----------------------------+-----------------------------------------------------------
Note: S.E. for ATT does not take into account that the propensity score is estimated.

 psmatch2: |   psmatch2: Common
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        581 |       581
   Treated |       995        822 |     1,817
-----------+----------------------+----------
     Total |       995      1,403 |     2,398
```

As a summary from the different modifications of the nearest neighbor matching algorithm, one can conclude that the ATT displayed in table 2.3 is different in its magnitude from the ATTs shown in tables 2.5, 2.6, and 2.7. Hence, an asymmetric distribution of propensity scores among both groups as shown in figure 2.3 is a possible explanation for a bad matching quality if controls are replaced and used several times as matching partners, and if no common support or trimming condition is imposed. Nevertheless, since the ATTs in tables 2.5, 2.6, and 2.7 also differ from each other between a wide range of 254 and 383, further algorithms should be applied.

[28] Davies and Kim (2004) suggested an algorithm which assigns different weights to the neighbors. Neighbors which are closer to the treated get higher weights.

2.4.2 Caliper matching

If samples are small or distributions of propensity scores differ, nearest neighbors may be far away from each other, and the matching quality is low. To ensure that matching partners are sufficiently close to each other, a maximum distance ϵ, called "caliper", between neighbors can be imposed. The neighborhood for this caliper matching (Cochran and Rubin, 1973) is defined as

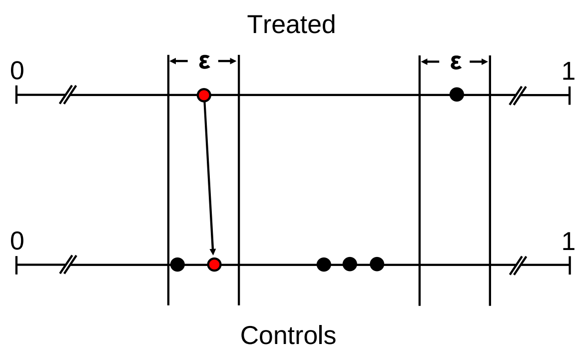
$$C^{CM}(P_i) = \{P_j | \; \|P_i - P_j\| < \epsilon\}, j \in N_0. \tag{12}$$

The control j has the weight

$$W^{CM}(i,j) = \begin{cases} 1 & \text{if } \|P_i(X) - P_j(X)\| = \min_j \|P_i(X) - P_j(X)\| \wedge \|P_i(X) - P_j(X)\| < \epsilon \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

Of course, there may be treated which do not find a matching partner within the neighborhood $C(P_i)$. These treated observations will not be considered for the analysis. The caliper ensures that a control is only assigned to a treated if the control lies within the neighborhood $C(P_i)$. It should be clear that a caliper restriction is an alternative to impose a common support condition (Caliendo, 2006).[29] Figure 2.6 displays how treated and controls are matched within the caliper algorithm.

Figure 2.6: **Graphical illustration of caliper matching**



Note: the two horizontal lines represent the range of propensity scores between zero and one, the upper line for treated, the lower line for controls. The circles represent single observations, each with a certain propensity score. Red circles indicate that observations are matched. The vertical lines indicate the caliper restriction: treated are matched to the nearest control within this caliper range. Source: own illustration.

[29] If `common` is also included in the STATA-command, the number of treated observations should not differ.

The lower the caliper, the more precise the matching, the higher the quality, and the lower the bias. However, the lower the caliper, the higher the variance. Moreover, low caliper values also lead to a higher number of treated which do not find a matching partner, and this reduces the number of matching pairs (Caliendo, 2006). The STATA-command for a caliper algorithm with $\epsilon = 0.01$ is

```
psmatch2 DUMMY_MA, outcome(SIZE_T3) pscore(pscore) caliper(0.01)
```

Table 2.8 reports the results for caliper matching with a maximum distance of 0.01, and the ATT is 354.

Table 2.8: **Caliper matching with a maximum distance of 0.01**

```
. psmatch2 DUMMY_MA, outcome(SIZE_T3) pscore(pscore) caliper(0.01)
There are observations with identical propensity score values.
The sort order of the data could affect your results.
Make sure that the sort order is random before calling psmatch2.
----------------------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference         S.E.   T-stat
----------------------------+-----------------------------------------------------------
         SIZE_T3  Unmatched | 1612.47111   135.500861   1476.97025   178.450092     8.28
                        ATT | 667.982843   314.139706   353.843137   59.2845514     5.97
----------------------------+-----------------------------------------------------------
Note: S.E. for ATT does not take into account that the propensity score is estimated.

 psmatch2: |   psmatch2: Common
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        581 |       581
   Treated |     1,001        816 |     1,817
-----------+----------------------+----------
     Total |     1,001      1,397 |     2,398
```

In table 2.9, the ATT is estimated for a maximum distance between treated and controls of only 0.001. As expected, the number of observations being matched is smaller, but the ATT does not substantially differ from the ATT in table 2.8.

Table 2.9: **Caliper matching with a maximum distance of 0.001**

```
. psmatch2 DUMMY_MA, outcome(SIZE_T3) pscore(pscore) caliper(0.001)
There are observations with identical propensity score values.
The sort order of the data could affect your results.
Make sure that the sort order is random before calling psmatch2.
----------------------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference         S.E.   T-stat
----------------------------+-----------------------------------------------------------
         SIZE_T3  Unmatched | 1612.47111   135.500861   1476.97025   178.450092     8.28
                        ATT |  712.03653   323.554033   388.482496   67.2079517     5.78
----------------------------+-----------------------------------------------------------
Note: S.E. for ATT does not take into account that the propensity score is estimated.

 psmatch2: |   psmatch2: Common
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        581 |       581
   Treated |     1,160        657 |     1,817
-----------+----------------------+----------
     Total |     1,160      1,238 |     2,398
```

2.4.3 Radius caliper matching

Radius caliper matching from Dehejia and Wahba (2002) is a further matching algorithm similar to caliper matching. Instead of using only the closest control within a maximum distance, all neighbors within the radius α will be used. Radius caliper matching leads to the following neighborhood:

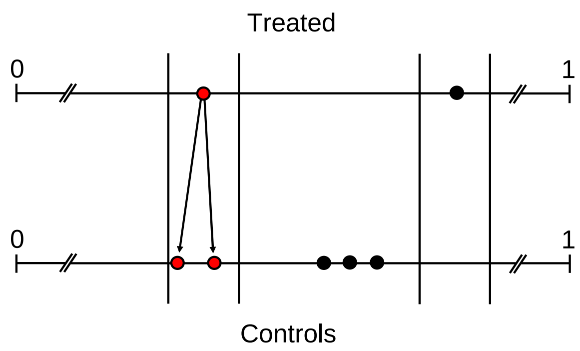
$$C^{RM}(P_i) = \{j| \ \|P_i(X) - P_j(X)\| < \alpha\}, j \in N_0. \tag{14}$$

All controls with a propensity score of $P_j(X)$ which lie within the radius α have the same weight:

$$W^{RM}(i,j) = \begin{cases} \frac{1}{C(P_i)}, & \text{if } j \in C^{RM}(P_i), \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

Table 2.7 presents a graphical illustration for the radius algorithm.

Figure 2.7: **Graphical illustration of radius caliper matching**



Note: the two horizontal lines represent the range of propensity scores between zero and one, the upper line for treated, the lower line for controls. The circles represent single observations, each with a certain propensity score. Red circles indicate that observations are matched. The vertical lines indicate the caliper restriction: treated are matched to all controls lying within the caliper radius. Source: own illustration.

The corresponding STATA-command for a radius of $\alpha = 0.01$ is

```
psmatch2 DUMMY_MA, outcome(SIZE_T3) pscore(pscore) radius caliper(0.01)
```

Radius caliper matching is similar to nearest neighbor matching with more neighbors, and to caliper matching by imposing a maximum distance. Radius caliper matching shares the attractive feature of oversampling, and also avoids the risk of bad matches. Radius caliper matching increases bias and decreases variance (Caliendo, 2006). Tables 2.10 and 2.11 report the results for radius calipers of 0.01 and 0.001. The estimated ATTs of 374 and 377 are very similar to each other.

Table 2.10: **Radius caliper matching with a maximum distance of 0.01**

```
. psmatch2 DUMMY_MA, outcome(SIZE_T3) pscore(pscore) radius caliper(0.01)
----------------------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference         S.E.   T-stat
----------------------------+-----------------------------------------------------------
         SIZE_T3  Unmatched | 1612.47111   135.500861   1476.97025   178.450092     8.28
                        ATT | 667.982843   294.482566   373.500277   42.8193644     8.72
----------------------------+-----------------------------------------------------------
Note: S.E. for ATT does not take into account that the propensity score is estimated.

 psmatch2: |   psmatch2: Common
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        581 |       581
   Treated |     1,001        816 |     1,817
-----------+----------------------+----------
     Total |     1,001      1,397 |     2,398
```

Table 2.11: **Radius caliper matching with a maximum distance of 0.001**

```
. psmatch2 DUMMY_MA, outcome(SIZE_T3) pscore(pscore) radius caliper(0.001)
----------------------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference         S.E.   T-stat
----------------------------+-----------------------------------------------------------
         SIZE_T3  Unmatched | 1612.47111   135.500861   1476.97025   178.450092     8.28
                        ATT |  712.03653   335.290726   376.745803   49.4253376     7.62
----------------------------+-----------------------------------------------------------
Note: S.E. for ATT does not take into account that the propensity score is estimated.

 psmatch2: |   psmatch2: Common
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        581 |       581
   Treated |     1,160        657 |     1,817
-----------+----------------------+----------
     Total |     1,160      1,238 |     2,398
```

2.4.4 Kernel matching

In comparison to the preceding algorithms, kernel matching uses all individuals j from the control group as neighbors for each single treated i. Thus, the neighborhood in kernel matching contains all observations in the control group I_0, which is equal to the set of matched controls A_i:

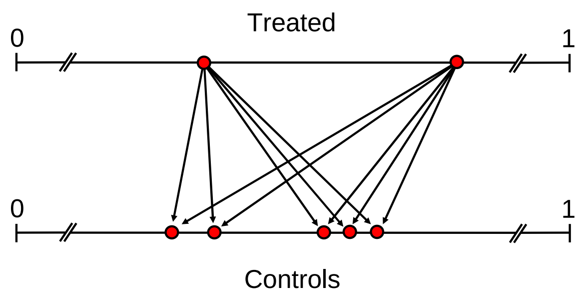
$$C(P_i) = \{I_0\} \tag{16}$$

Controls j are weighted according to their distance to the treated i, i.e. controls which are closer receive a higher weight than others:

$$W_{N_0}^{KM}(i,j) = \frac{G_{ij}}{\sum_{k \in I_0} G_{ik}}, \tag{17}$$

where $G_{ik} = G[(P_i - P_k)/a_{N_0}]$ is a kernel function which downweighs observations j which are distant from the treated i, and a_{N_0} is a bandwidth parameter which impacts the form of the kernel function.[30] Table 2.8 shows a graphical illustration of the kernel algorithm.

Figure 2.8: **Graphical illustration of kernel matching**



Note: the two horizontal lines represent the range of propensity scores between zero and one, the upper line for treated, the lower line for controls. The circles represent single observations, each with a certain propensity score. Red circles indicate that observations are matched. In kernel matching, all treated are matched with all controls at different weights. Source: own illustration.

The researcher can choose between Gaussian (normal), biweight, epanechnikov, uniform, and tricube kernel. They differ with respect to the underlying kernel function that approximates density kurves. However, the choice of the kernel function is less important for the results, as DiNardo and Tobias (2001) showed. In contrast, the choice of the bandwidth parameter is more relevant (Silverman, 1986; Pagan and Ullah, 1999): higher values of the bandwidth parameter a_{N_0} smooth the density function. This leads to a better fit and decreased variance between the estimated and the true density function. However, if the bandwidth parameter is too high the underlying structure may be smoothed away, and the estimate is biased. Thus, the researcher faces a trade-off between a small variance and an unbiased estimate of the underlying true density function (Caliendo, 2006).

The default in STATA is epanechnikov kernel and a bandwidth of 0.06. Because all controls are used as matching partners, applying a common support is likely to improve the matching quality. The respective STATA-command is

```
psmatch2 DUMMY_MA, outcome(SIZE_T3) pscore(pscore) kernel common
```

The results are displayed in table 2.12. The ATT of 388 is close to the estimations from other algorithms like caliper or radius caliper matching.

[30] See Fahrmeir, Künstler, Pigeot, and Tutz (2009) for an introduction to kernel density estimation, and Heckman, Ichimura, Smith, and Todd (1998) for a discussion about kernel matching estimators.

Table 2.12: **Epanechnikov kernel matching with a bandwidth of 0.06**

```
. psmatch2 DUMMY_MA, outcome(SIZE_T3) pscore(pscore) kernel common
----------------------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference         S.E.   T-stat
----------------------------+-----------------------------------------------------------
         SIZE_T3  Unmatched | 1612.47111   135.500861   1476.97025   178.450092     8.28
                        ATT | 653.891727   265.799443   388.092285   39.4509803     9.84
----------------------------+-----------------------------------------------------------
Note: S.E. for ATT does not take into account that the propensity score is estimated.

 psmatch2: |   psmatch2: Common
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        581 |       581
   Treated |       995        822 |     1,817
-----------+----------------------+----------
     Total |       995      1,403 |     2,398
```

The alternative command for kernel matching with a Gaussian kernel[31] and a bandwidth of 0.01 is

```
psmatch2 DUMMY_MA, outcome(SIZE_T3) pscore(pscore) kernel
kerneltype(normal) bandwidth(0.01) common
```

The results of this kernel algorithm are shown in table 2.13. The ATT of 383 is similar to the ATT displayed in table 2.12.[32]

Table 2.13: **Gaussian kernel matching with a bandwidth of 0.01**

```
. psmatch2 DUMMY_MA, outcome(SIZE_T3) pscore(pscore) kernel kerneltype(normal) bwidth(0.01)
common
----------------------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference         S.E.   T-stat
----------------------------+-----------------------------------------------------------
         SIZE_T3  Unmatched | 1612.47111   135.500861   1476.97025   178.450092     8.28
                        ATT | 653.891727   270.643997    383.24773   40.1674143     9.54
----------------------------+-----------------------------------------------------------
Note: S.E. for ATT does not take into account that the propensity score is estimated.

 psmatch2: |   psmatch2: Common
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        581 |       581
   Treated |       995        822 |     1,817
-----------+----------------------+----------
     Total |       995      1,403 |     2,398
```

The STATA-command for a modified kernel matching algorithm with a bandwidth of 1.0 is similar to the one above:

```
psmatch2 DUMMY_MA, outcome(SIZE_T3) pscore(pscore) kernel
kerneltype(normal) bandwidth(1.0) common
```

[31] Commands for other kernel functions are `kerneltype(epanechnikow)` (this is the default), `kerneltype(biweight)`, `kerneltype(uniform)`, `kerneltype(tricube)`.

[32] I estimated several ATTs based on different kernel functions, holding the bandwidth parameter constant. Results were all similar, confirming the findings from DiNardo and Tobias (2001).

In table 2.14 the results from the kernel matching algorithm with a bandwidth of 1.0 are presented. The ATT is now clearly different from the other ATTs based on kernel matching.

Table 2.14: **Gaussian kernel matching with a bandwidth of 1.0**

```
. psmatch2 DUMMY_MA, outcome(SIZE_T3) pscore(pscore) kernel kerneltype(normal) bwidth(1.0)
common
----------------------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference         S.E.   T-stat
----------------------------+-----------------------------------------------------------
         SIZE_T3  Unmatched | 1612.47111   135.500861   1476.97025   178.450092     8.28
                        ATT | 653.891727   137.289282   516.602446   37.9295991    13.62
----------------------------+-----------------------------------------------------------
Note: S.E. for ATT does not take into account that the propensity score is estimated.

 psmatch2: |   psmatch2: Common
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        581 |       581
   Treated |       995        822 |     1,817
-----------+----------------------+----------
     Total |       995      1,403 |     2,398
```

Several conclusions from different algorithms are possible. Eight out of eleven different algorithms estimate comparable ATTs which range between 318 and 388, and all are highly statistically significant. This is in line with the prediction of Smith and Todd (2005a) who stated that the choice of the algorithm should be of minor interest if samples are larger. However, three algorithms estimate ATTs which are different in their magnitude: the very first nearest neighbor algorithm estimates an ATT of 1,151 (table 2.3). This algorithm allows replacement, and it does not impose a common support condition, which is obviously not a good idea given the asymmetric distribution of propensity scores in both groups. Second, the existence of several gaps for some ranges of propensity scores recommends the imposition of a trimming condition rather than a common support condition in nearest neighbor algorithms. This is a likely explanation for an estimated ATT of 254 (table 2.5). Third, a bandwidth parameter of 1.0 in a kernel algorithm is obviously too high, resulting in an ATT of 516 (table 2.14).

Finally, table 2.15 shows a summary of the different algorithms presented in this section with respect to the trade-off between bias and variance.

2.4.5 Bootstrapping

Bootstrapping is a resampling method developed by Efron (1979), and it is used to estimate standard errors if analytical estimates are biased or unavailable (Brownstone and Valletta, 2001). With bootstrapping, repeated samples can be drawn from an original sample, and due to replacement the new bootstrap samples are of the same size as the original sample. Therefore, the population is to the sample as the sample to the bootstrap samples (Fox, 1997).

In the context of matching, n bootstrap-replications calculate n new estimates for propensity

Table 2.15: **Summary of algorithms**

Algorithm	Bias	Variance	ATT
NN matching with replacement	↓	↑	1151.97
NN matching without replacement and common support	↑	↓	254.18
NN matching without replacement and trimming	↑	↓	318.83
NN matching with five neighbors	↑	↓	383.35
Caliper matching, distance of 0.01	↓	↑	353.84
Caliper matching, distance of 0.001	↓	↑	388.48
Radius caliper matching, distance of 0.01	↑	↓	373.50
Radius caliper matching, distance of 0.001	↑	↓	376.75
Kernel matching (Epanechnikov), bandwidth of 0.06	↑	↓	388.09
Kernel matching (Gaussian), bandwidth of 0.01	↓	↑	383.25
Kernel matching (Gaussian), bandwidth of 1.0	↑	↓	516.60

Note: NN = nearest neighbor; ↑ = increase; ↓ = decrease. With respect to kernel matching, bias decreases with low and increases with high bandwidth parameters. There is no clear benchmark, as may be incorrectly concluded from the changing arrows. This is also true for the variance. Source: Caliendo, 2006; own presentation.

scores, the corresponding regions of common support, ATTs, standard errors, and t-statistics. The default for the number of bootstrap-replications in PSMATCH2 is 50, but Efron (1990) suggested 50 to 200 replications in order to yield proper results for standard errors.

Since the program PSMATCH2 stores the estimated ATT in `r(att)`, bootstrapping of the standard error of the ATT is possible. However, Abadie and Imbens (2008) criticized the use of bootstrapping in the context of matching because no formal justification has been provided. Nevertheless, many empirical studies that apply matching perform bootstrapping (e.g. Girma, Görg, and Wagner, 2009). The corresponding STATA-command to apply bootstrapping within the matching process - here for a caliper matching with $\epsilon = 0.001$ and 200 replications - is

```
bootstrap r(att) reps(200):  psmatch2 DUMMY_MA, outcome(SIZE_T3)
pscore(pscore) caliper(0.001)
```

The results are reported in table 2.16. The ATT is still statistically significant.

2.4.6 Difference-in-differences propensity score matching

There are several ways to combine propensity score matching with other methods (Caliendo, 2006): first, difference-in-differences (DiD) propensity score matching as suggested by Heckman, Imichura, and Todd (1997), and Heckman, Imichura, Smith, and Todd (1998); second, regression-adjusted matching estimator from Heckman, Imichura, and Todd (1997), and Heckman, Imichura, Smith, and Todd (1998); and third, bias-corrected matching estimator according to Abadie and Imbens (2011), and Imbens (2004). I only present DiD in combination with

Table 2.16: **Caliper matching (0.001) and bootstrapped standard errors**

```
. bootstrap r(att), reps(200) : psmatch2 DUMMY_MA, outcome(SIZE_T3) pscore(pscore)
 caliper(0.001)
(running psmatch2 on estimation sample)
Note: S.E. for ATT does not take into account that the propensity score is
estimated.

Bootstrap replications (200)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
..................................................    50
..................................................   100
..................................................   150
..................................................   200

Bootstrap results                               Number of obs      =      2398
                                                Replications       =       200

      command:  psmatch2 DUMMY_MA, outcome(SIZE_T3) pscore(pscore)
                    caliper(0.001)
        _bs_1:  r(att)

------------------------------------------------------------------------------
             |   Observed   Bootstrap                         Normal-based
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |   388.4825    70.27417     5.53   0.000     250.7477    526.2173
------------------------------------------------------------------------------
```

propensity score matching. This approach is widely used in empirical studies, whereas there are only a few studies applying the other modifications.

In general, a DiD estimator compares changes in a variable over time between groups. The use of a DiD method in combination with propensity score matching may improve the results compared to the standard matching estimator: after conditioning on observables there may still be differences between the outcomes of participants and non-participants which are due to systematic differences in both groups, i.e. because of selection into the treatment based on unmeasured characteristics (Smith and Todd, 2005a). Due to the comparison of changes instead of levels, DiD propensity score matching helps to eliminate unobserved time-invariant differences between both groups, and relaxes the strong assumption of selection on observables. Smith and Todd (2005a) found that DiD matching performs substantially better than the corresponding cross-sectional matching estimator. For this reason, several empirical studies combine DiD with propensity score matching, e.g. in the fields of international economics Girma and Görg (2006), Hijzen, Inui, and Todo (2007), Görg, Henry, and Strobl (2007), and Barba-Navaretti, Castellani, and Disdier (2010), and others.

The DiD propensity score matching estimator is based on the identifying assumption:

$$E(Y_t^0 - Y_{t'}^0 | P(X), D = 1) = E(Y_t^0 - Y_{t'}^0 | P(X), D = 0). \tag{18}$$

Under the consideration of the common support condition, the estimator can be implemented as

$$\Delta_{ATT}^{DiD} = \frac{1}{N_1} \sum_{i \in I_1 \cap S_P} \left[(Y_{it}^1 - Y_{it'}^0) - \sum_{j \in I_0 \cap S_P} W(i,j)(Y_{jt}^0 - Y_{jt'}^0) \right], \tag{19}$$

with S_P denoting the region of common support, and t' as an indicator for pre-treatment periods. In the respective STATA-command, only the outcome variable is modified, which is now SIZE_31, the change in the number of employees between $t = 1$ and $t = 3$. I use the caliper algorithm with a caliper of 0.01:

```
psmatch2 DUMMY_MA, outcome(SIZE_31) pscore(pscore) caliper(0.001)
```

Table 2.17 shows the estimated ATT from a DiD propensity score matching. Now, the results from the combined estimator are different from the results above: plants that merged decreased their employment over time, while comparable plants that had not merged increased the number of employees. However, differences are not statistically significant, i.e. there is no effect of M&A on the merging plants' employment.

Table 2.17: **Caliper matching (0.001) with a difference-in-differences estimator**

```
. psmatch2 DUMMY_MA, outcome(SIZE_31) pscore(pscore) caliper(0.001)
----------------------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference         S.E.   T-stat
----------------------------+-----------------------------------------------------------
         SIZE_31  Unmatched |  -48.70776  -1.29604131  -47.4117187    28.687518    -1.65
                        ATT |-4.37290715   4.66210046  -9.03500761   16.9785876    -0.53
----------------------------+-----------------------------------------------------------
Note: S.E. for ATT does not take into account that the propensity score is estimated.

 psmatch2: |   psmatch2: Common
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        581 |       581
   Treated |     1,160        657 |     1,817
-----------+----------------------+----------
     Total |     1,160      1,238 |     2,398
```

How can these differences in the results be explained? As argued above, the combination of the standard matching estimator with a DiD estimator relaxes the strong assumption of selection on observables, i.e. with a combined estimator a selection of unobservables is also possible, as long as they are time-invariant. Thus, the different results between the standard and the combined matching estimator imply that unobserved time-invariant variables play a crucial role for the selection. The standard estimator obviously fails to eliminate differences in employment between merged and non-merged plants caused by time-invariant unobservables. To conclude, a combination of DiD and the standard estimator makes a causal interpretation of the results more reliable, and it improves the quality of non-experimental evaluation results significantly (Blundell and Costa Dias, 2000).

2.5 Was the matching procedure successful?

After performing different matching algorithms it is reasonable to ask if the results are useful. There are several indicators which assess the quality of the matching process, and I will briefly explain them.

Common support: The results from matching may be biased if treated and control observations are not similar enough in terms of propensity scores. The implementation of a common support condition eliminates observations lying outside a common region of propensity scores and avoids bad matches. However, excluding a certain number of observations from the analysis is not harmless per se (Lechner, 2008). The estimated treatment effect does no longer correspond to the original parameter of interest, if treatment effects are heterogeneous inside and outside the common support. Moreover, throwing away all observations outside the common support ignores useful information, because treatment effects can still be estimated outside the region of common support.

The relevance of these concerns clearly depends on the fraction of observations lying outside the region of common support. That is, the higher the number of observations excluded from the analysis, the lower the explanatory power of the matching process and the generality of estimations (see tables 2.5, 2.6, etc.). In general, causal inference is restricted only to those matched observations lying within the region of common support. A graphical visualization of the distribution of propensity scores among both groups may be helpful (see figure 2.3).

Weights of controls: A further important indicator is how often a single control is used as a match, as already shown in table 2.4. Again, the command is

```
tabulate _weight treated
```

and lists the frequency a single control is used as a matching partner. The matching quality suffers if the same controls are matched too often with different treated observations. If this problem is severe (e.g. table 2.3), the `noreplacement` option may improve the results and allows each control to be matched only once.

Balancing property: The intension of matching is to balance the distribution of all covariates X between treated and untreated after matching. This balancing property is a striking challenge in order to obtain good matching quality. There are several ways to check if the balancing property is fulfilled. Rosenbaum and Rubin (1985) calculate a standardized bias, i.e. it calculates the balancing of each variable before and after matching in terms of standardized bias, and t-tests for the equality of means. The standardized bias for each covariate is the difference in the sample means in both groups as a percentage of the square root of the average of the sample variances in both groups (Rosenbaum and Rubin, 1985). Before matching, the standardized bias SB is calculated as

$$SB = 100\frac{(\bar{X}_1 - \bar{X}_0)}{\sqrt{0.5[V_1(X) + V_0(X)]}}. \tag{20}$$

After matching, the standardized bias SB^M is given by

$$SB^M = 100\frac{(\bar{X}_1^M - \bar{X}_0^M)}{\sqrt{0.5[V_1^M(X) + V_0^M(X)]}}. \tag{21}$$

For each covariate, $\bar{X}_1$ and $\bar{X}_0$ are the sample means in the treated and control group, and $V_1(X)$ and $V_0(X)$ are the corresponding variances. $\bar{X}_{1M}$, $\bar{X}_{0M}$, $V_{1M}(X)$ and $V_{0M}(X)$ are values after matching. The lower the standardized bias after matching, the more both groups are balanced with respect to a specific variable. However, there is no benchmark for the level of the standardized bias, but Rosenbaum and Rubin (1985) assume that values above 20 imply serious bias, and Caliendo (2006) pointed out that in most empirical studies a bias reduction below 3 or 5% is seen as sufficient. The respective STATA-command for the standardized bias is

```
pstest SIZE_CAT_T1_2-SIZE_CAT_T1_6 SECTOR_T1_2-SECTOR_T1_16
LEGAL_FORM_T1_2 LEGAL_FORM_T1_3 EAST_T1
```

The results, which are based on the caliper algorithm with $\epsilon = 0.001$, are shown in table 2.18. Columns one and two report mean values of all covariates X for the matched and unmatched sample, and the third column reports the calculated standardized bias also for both samples. The table displays the percentage reduction of the bias in the fourth column. Finally, two-sample t-tests check the balancing property, i.e. if differences in the means of the covariates for both groups are insignificant after matching. Therefore, the last two columns provide the results of the t-test and the corresponding p-values. As an example, consider variable `SIZE_CAT_T1_2`, the dummy variable for the second size category measured in $t = 1$: the mean value for the unmatched sample is 0.075 for treated, and 0.198 for controls, and the difference in means is highly statistically significant. For the matched sample, the mean is 0.094 both for treated and controls. That is, the standardized bias of 36.4 in the unmatched sample[33] was reduced to a bias of 0.0, and this corresponds to a bias reduction of 100%. That means, matching completely balanced this variable.

Table 2.18: **Quality checks: standardized bias**

```
. pstest SIZE_CAT_T1_2-SIZE_CAT_T1_6 SECTOR_T1_2-SECTOR_T1_11 LEGAL_FORM_T1_2-
LEGAL_FORM_T1_3 EAST_T1

-----------------------------------------------------------------------------
                         |       Mean               %reduct |     t-test
    Variable     Sample  | Treated Control    %bias  |bias| |    t    p>|t|
-------------------------+----------------------------------+----------------
SIZE_CAT_T~2  Unmatched  | .07485   .19793    -36.4          |  -8.56  0.000
                Matched  | .09437   .09437      0.0   100.0  |   0.00  1.000
                         |                                  |
SIZE_CAT_T~3  Unmatched  | .06329     .179    -36.0          |  -8.55  0.000
                Matched  | .09741    .0898      2.4    93.4  |   0.47  0.636
                         |                                  |
SIZE_CAT_T~4  Unmatched  | .13704    .3167    -43.9          | -10.00  0.000
                Matched  | .21157   .19482      4.1    90.7  |   0.75  0.451
                         |                                  |
SIZE_CAT_T~5  Unmatched  | .10017   .09983      0.1          |   0.02  0.981
                Matched  | .12177   .12329     -0.5  -351.4  |  -0.08  0.933
                         |                                  |
SIZE_CAT_T~6  Unmatched  |  .5366   .03442    133.7          |  23.76  0.000
                Matched  | .37139   .38813     -4.5    96.7  |  -0.62  0.532
                         |                                  |
 SECTOR_T1_2  Unmatched  | .04898   .00344     28.8          |   5.02  0.000
                Matched  |      0        0      0.0   100.0  |      .      .
                         |                                  |
 SECTOR_T1_3  Unmatched  | .02091   .02065      0.2          |   0.04  0.970
                Matched  | .01218   .01218      0.0   100.0  |   0.00  1.000
                         |                                  |
 SECTOR_T1_4  Unmatched  | .02367    .0327     -5.5          |  -1.19  0.232
                Matched  | .01979   .01979      0.0   100.0  |  -0.00  1.000
                         |                                  |
 SECTOR_T1_5  Unmatched  | .18162   .18589     -1.1          |  -0.23  0.817
                Matched  | .27702   .28615     -2.4  -114.0  |  -0.37  0.713
                         |                                  |
 SECTOR_T1_6  Unmatched  | .25427   .34768    -20.5          |  -4.40  0.000
                Matched  | .38813   .35616      7.0    65.8  |   1.20  0.231
                         |                                  |
 SECTOR_T1_7  Unmatched  | .03137   .03098      0.2          |   0.05  0.963
                Matched  | .01979   .01065      5.3 -2245.7  |   1.35  0.177
                         |                                  |
 SECTOR_T1_8  Unmatched  |  .0809   .20138    -35.1          |  -8.18  0.000
                Matched  | .13546   .16134     -7.5    78.5  |  -1.32  0.187
                         |                                  |
 SECTOR_T1_9  Unmatched  | .14034   .04131     35.0          |   6.54  0.000
                Matched  | .05479    .0624     -2.7    92.3  |  -0.59  0.557
                         |                                  |
SECTOR_T1_10  Unmatched  | .00055   .00172     -3.5          |  -0.85  0.395
                Matched  |      0        0      0.0   100.0  |      .      .
                         |                                  |
SECTOR_T1_11  Unmatched  | .03687   .11015    -28.3          |  -6.83  0.000
                Matched  | .07002   .07763     -2.9    89.6  |  -0.53  0.598
                         |                                  |
LEGAL_FOR~_2  Unmatched  | .37259   .83821   -108.3          | -21.31  0.000
                Matched  | .82953   .85693     -6.4    94.1  |  -1.37  0.172
                         |                                  |
LEGAL_FORM~3  Unmatched  | .52119    .0241    134.5          |  23.62  0.000
                Matched  | .07458   .05936      4.1    96.9  |   1.10  0.270
                         |                                  |
     EAST_T1  Unmatched  |  .2284   .27367    -10.4          |  -2.23  0.026
                Matched  | .19635   .16743      6.7    36.1  |   1.36  0.174
                         |                                  |
```

For further robustness checks, the `pstest`-command can be extended by the option `summary`:

```
pstest SIZE_CAT_T1_2-SIZE_CAT_T1_6 SECTOR_T1_2-SECTOR_T1_16
LEGAL_FORM_T1_2 LEGAL_FORM_T1_3 EAST_T1, summary
```

Table 2.19 displays the results: the upper part presents the distribution of the mean standardized

[33] Because the mean value of controls is larger than the mean value of treated, the bias has a negative sign. See equations 20 and 16.

bias (MSB) across all variables, that is, the sum of the bias of all variables divided by the number of variables, both for the unmatched and matched sample. Here, the MSB decreased from 36.7 before matching to 3.1 after matching, which can be seen as a sufficient bias reduction.

Table 2.19: **Quality checks: mean standardized bias and other indicators**

```
. pstest SIZE_CAT_T1_2-SIZE_CAT_T1_6 SECTOR_T1_2-SECTOR_T1_11
LEGAL_FORM_T1_2-LEGAL_FORM_T1_3 EAST_T1, quietly summary
-------------------------------------------------------------
         Summary of the distribution of the abs(bias)
-------------------------------------------------------------

                       BEFORE MATCHING
-------------------------------------------------------------
      Percentiles      Smallest
 1%     .1123461       .1123461
 5%     .1123461       .1818326
10%     .1818326       .2238897       Obs                  18
25%     3.474108       1.101518       Sum of Wgt.          18

50%     28.56449                      Mean           36.74933
                        Largest       Std. Dev.      43.78558
75%     36.42978       43.89066
90%     133.7132       108.2914       Variance       1917.177
95%     134.4857       133.7132       Skewness       1.390628
99%     134.4857       134.4857       Kurtosis       3.652939
-------------------------------------------------------------

                       AFTER MATCHING
-------------------------------------------------------------
      Percentiles      Smallest
 1%            0              0
 5%            0              0
10%            0              0       Obs                  18
25%            0              0       Sum of Wgt.          18

50%     2.815315                      Mean           3.131599
                        Largest       Std. Dev.      2.689283
75%     5.251813        6.37194
90%     7.000823       6.674671       Variance       7.232245
95%     7.540554       7.000823       Skewness       .1966783
99%     7.540554       7.540554       Kurtosis       1.695646
-------------------------------------------------------------

--------------------------------------------------------------
     Sample |    Pseudo R2      LR chi2        p>chi2
------------+-------------------------------------------------
  Unmatched |        0.432      1146.57         0.000
    Matched |        0.006        11.05         0.806
--------------------------------------------------------------
```

A further test in order to check the quality of the matching process is to repeat the probit estimations for the matched sample, and to observe what happens to the values of the pseudo-R^2 and LR-test. As explained earlier in this paper, the McFadden's pseudo-R^2 measures how much of the variation of the dependent variable is explained by the covariates. Thus, if matching was successful and covariates are balanced, a variation of the dependent variable is only explained by the treatment, but not by the covariates any more. For this reason, the respective pseudo-R^2 for the estimation of the matched sample should be close to zero. In table 2.19, it is 0.006 which is low enough. A similar rationale applies to the calculation of the LR-test and the corresponding prob-values. If matching was successful, the coefficients of the regression should be zero or at least close to zero. Hence, the difference between the calculated loglikelihood value for this model, and the loglikelihood value of an alternative model in which all coefficients are set to zero, except the constant, should be small. This implies a low explanation of the independent variables of the model. The corresponding prob-value, which describes the probability that coefficients are zero, should be high. Here, the prob-value of 0.806 can be regarded as high enough.

There are further possibilities to test the balancing and sensitivity of the matching results. However, they are not automatically generated by PSMATCH2, and thus, I do not report them. For example, Girma and Görg (2007) also tested the robustness with a Hotelling's t-squared test, with a test cast within a regression framework developed by Smith and Todd (2005b), and finally, with a reestimation of the propensity score model with minor changes in the regression model as suggested by Dehejia (2005).

2.6 Conclusion - is matching better than the rest?

The objective of matching is to identify the causal effects of a treatment. For this, the method compares the outcome of individuals that participated in a treatment with a comparison group of individuals with the same characteristics that did not participate in the treatment. Thus, matching enables the researcher to observe what would have happened to treated individuals if they had not participated in the treatment.

This paper is an introduction to the implementation of propensity score matching in the software STATA with the program PSMATCH2 from Leuven and Sianesi (2003). The target group of this paper are researchers that do not want to be held up too long with details of the underlying theoretical framework, but wish to get a quick basic understanding of the method and how it can be performed in STATA. For this, I present a theoretical introduction, describe the stepwise implementation of the method, and explain and interpret the results. I use a combined dataset from the IAB Establishment Panel and the M&A DATABASE St. Gallen.

Matching is a helpful tool within microeconometric evaluation methods, because it enables the researcher to better analyze causal effects. It exhibits several useful advantages, because its logic is simple, it is intuitively plausible, and it is easy to handle because of only little statistical and mathematical assumptions. Moreover, the method is based on a solid theoretical framework from Roy (1951) and Rubin (1974), and it directly estimates the relevant causal parameter of the model. For this, a matching analysis is technically easy to perform, and its rational is also easy to communicate (DiPrete and Gangl, 2004).

However, there are also open questions. For example, matching is based on the strong assumption of conditional independence which states that selection of individuals to treatment is only based on observables. Because the assumption can't be tested statistically, it is the researcher's task to present convincing arguments that it holds. As described in this paper, the

CIA can be relaxed by combining the matching estimator with a DiD estimator. This combined estimator eliminates time-invariant unobserved heterogeneity between treated and controls. In addition, details of propensity score matching are not yet standardized, e.g. how to model the propensity score, or how to do inference. This may lead to different conclusions from different researchers, even if they use the same data (Angrist and Pischke, 2009).

Hence, it is important to be aware of these weaknesses when performing matching. The researcher is well advised to make sure that the data is sufficiently good to ensure the assumptions to be fulfilled. Moreover, the results can be expected to be more reliable the larger the sample, and the larger the overlap of treated and control groups. Nevertheless, making the estimation results from empirical investigation even more robust, it seems plausible not to rely only on a single method, but also to apply other methods, e.g. to start with a regression analysis, as also suggested by Angrist and Pischke (2009).

Chapter 3

M&A and labor productivity: new evidence from micro-data for German plants

3.1 Introduction

Mergers and acquisitions (M&A)[1] are an elementary component in altering the structure, scope, or size of firms.[2] Although M&A are, in principle, substitutable mechanisms to internal growth (McGauckin and Nguyen, 1995), they lead to vivid discussions in public and among policy makers about their impacts on the merging firms' performance and the economy as a whole. Unfortunately, the existing empirical literature about the effects of M&A on several firm performance parameters[3] fails to provide clear results.[4] For example, and with respect to productivity, existing studies - see below - found either positive, negative, or no effects.

There are several reasons for this ambiguity. One explanation may be that we are comparing the incomparable because studies differ from each other in important dimensions. For example, there are different types of mergers, like horizontal M&A (firms which compete in the same market combine), vertical M&A (a firm combines with its supplier), or conglomerate M&A (firms of unrelated lines of businesses combine) (Carlton and Perloff, 2005), and studies often focus only on one type.[5] Moreover, studies may focus on different types of firms involved in the merger, either on "acquirers" or "buyer firms" (firms which acquire other firms), on "targets", "objects", or "acquired" (firms which are acquired by firms), or on "sellers" (firms which sell parts of the overall entity). Additionally, studies differ in further aspects, e.g. with respect to countries and industries in which the merger occurs, underlying observation periods, domestic or cross-border M&A, estimations on plant-level or firm-level, or definitions of performance parameters, e.g. labor vs. total factor productivity. Therefore, the ambiguous results are not too surprising if the observation units differ in relevant characteristics. Morck, Shleifer, and Summers (1988) confirm this concern by stating that "research results on friendly bids may have little to say about hostile bids, and vice versa". A second explanation for differing research results may be that earlier studies suffer from measurement errors or bias (McGuckin and Nguyen, 1995) due to the use of random samples of non-merging firms as control groups.

[1] There is no consistent definition about the difference between "mergers" and "acquisitions" in the literature. A way to distinguish between mergers and acquisitions is their legal entity: in an acquisition or takeover, firms preserve their legal entity. In a merger, firms lose their legal entity and combine into a new firm (Jansen, 2008). However, these differences are not accounted for in this paper, and therefore, I use the terms M&A, mergers, acquisitions, and takeovers as synonyms.

[2] I use the term "firm" if no greater precision is needed. However, the empirical part of the paper is about plants.

[3] See Jansen (2008), Mueller (2003a), or Tichy (2001) for a survey of empirical studies.

[4] Tichy (2001) presented 18 stylized facts about the effects of mergers on different performance parameters. However, I do not see a sufficient empirical support for most of these statements. This is also critisized by Lyons (2001).

[5] The distinction between different types of mergers is not always clear and highly dependent on industry classifications (Pesendorfer, 2003).

A comparison of both groups can lead to misleading conclusions if merging firms and non-merging firms differ in pre-merger characteristics that influence performance. In other words, firms with specific characteristics self-select in M&A activity, and changes in performance are then incorrectly attributed to the impact of the merger (Girma, Thompson, and Wright, 2006). Recent studies applied different estimation strategies with advanced econometric methods in order to improve research about causality and self-selection.[6]

The objective of this paper is to study the effects on plants' productivity due to M&A, and to overcome the problem of self-selection of plants in M&A activity. I analyze productivity because it is an applicable measure of a plant's efficiency (McGuckin and Nguyen, 1995), and I believe that the analysis of efficiency is crucial to assess the success of mergers. I take into account that there is no typical merger (Tichy, 2001), and thus, I distinguish between acquirers and targets, and between horizontal and non-horizontal (vertical and conglomerate) mergers.[7] In order to control for selection bias and identify the causal average effect of M&A on plants' productivity, I apply a matching approach, a newer microeconometric evaluation method.

To preview my results, I find that merging plants are more productive than non-merging plants. This difference is due to a pre-merger heterogeneity between plants. In particular, I find that more productive plants self-select in merger activities. Furthermore, I find a weak evidence for a causal effect of M&A on productivity changes for acquirers.

There are several empirical studies about mergers and productivity performance of firms related to my paper. For example, positive effects were found in Pesendorfer (2003) or Maksimovic, Phillips, and Prabhala (2011). No effects were found in studies from Bandick (2011), or Mattes (2010), and negative productivity effects were found in Gioia and Thomsen (2004).

The present paper sets itself apart from existing studies in two dimensions. First, I use a new dataset combined from the IAB (Institute of Employment Research) Establishment Panel and the M&A DATABASE, St. Gallen. To the best of my knowledge this dataset has not been used by others so far. The dataset includes German plants that were involved in M&A activity between 1996 and 2005, and a control group with plants not involved in any M&A activity since 1980. The dataset allows distinguishing between acquirers, targets, horizontal, and non-horizontal merging plants, and thus, addresses the critics about reduced comparability of different studies mentioned above. Moreover, other datasets are often biased towards larger

[6] Angrist and Pischke (2010) talk about a "credibility revolution in empirical economics" because of better data availability and better research designs.

[7] Firms involved in either vertical or conglomerate mergers operate in different markets in comparison to firms involved in horizontal mergers. Therefore, vertical and conglomerate mergers are often considered together as non-horizontal mergers (Church, 2004). In this study, I am not able to distinguish between vertical and conglomerate mergers because of too few observations.

companies, but this dataset is at plant-level and also enables an analysis of small and medium-size plants. Additionally, since the number of observations of foreign M&A is small in this dataset, my paper places the focus on domestic mergers. This should not be a drawback, because most of the recent studies only analyze the effects from foreign M&A. Hence, the effects from domestic M&A are ignored, although they account for at least 50% of all mergers in Germany (Spanninger, 2011a). Second, I apply a three-step evaluation strategy. As usual, I start with an analysis of descriptive statistics. Then, I will perform a regression analysis to control for other variables that may influence productivity. Finally, I perform a propensity score matching technique based on Rosenbaum and Rubin (1983, 1985). Matching has increasingly been applied in recent studies to identify causalities. It has helped to overcome biased results from many earlier studies which simply compared merging and non-merging firms, but did not control for self-selection effects. I will combine the matching estimator with a difference-in-differences estimator as suggested by Blundell and Costa Dias (2000) and Smith and Todd (2005a) in order to eliminate time-invariant heterogeneity between groups. I will also apply a number of robustness checks in order to test the credibility of the matching results.

The remainder of the paper is organized as follows: Section 3.2 describes the economic theory of mergers and acquisitions and presents an overview of related literature. Section 3.3 describes the data, and section 3.4 reports on the results from the empirical investigation. Section 3.5 summarizes the results of the paper.

3.2 Theoretical background and related literature

In this section, I present several reasons for mergers that have been identified in literature (e.g. Scherer, 2002; Mueller, 2003a), and I ask about the implications for merging firms' pre- and post-merger performance, in particular productivity: do firms with specific characteristics self-select in merger activity? Do firms acquire over-performing firms, so-called "cherries", or underperforming firms, so-called "lemons"? What is the effect of M&A on the merging firms' performance? And do the results differ with the type of merger?

A motivation for M&A widely discussed in the literature is synergy gains. These synergies can occur in different ways: the merger allows for the reorganization of business structures, which supports business growth, allowing for changes in the mixes of goods and services, and may improve firms' technical and organizational systems (Seth, 1990; Capron and Mitchell, 1998). I expect the effect of synergies on the merging firms' productivities to be positive. Since both firms have to participate in realizing the gains, these gains can be expected to be shared between the acquirer and the target (Mueller and Sirower, 2003). However, no clear prediction of the pre-merger performance of merging firms can be made. On the one hand, acquiring firms face several costs if they take over another firm: they have to finance the merger itself (with stocks or cash), and they also have to bear fixed costs for the integration of the acquired firm. Hence, it seems plausible that only better performing firms are able to incur these fixed costs.[8] In contrast, it could also be the case that poorly performing acquirers self-select in merger activity to improve their own performance through efficiency gains from synergy effects. With respect to targets, it seems more plausible to assume that gains from synergies are extracted from good rather than bad performing targets.

A further possible motivation for takeovers is to replace inefficient management which does not maximize shareholder wealth (Scharfstein, 1988). According to Jensen (1988), the corporate takeover market acts as a "court of last resort", i.e. takeovers are an external source of discipline if internal control mechanisms are weak or ineffective. Brealey, Myers, and Allen (2008) state that firms, which have unexploited opportunities to cut costs and increase sales and earnings, are candidates for a takeover by firms which have better management. Hence, if these poorly performing targets are taken over, the merger has a disciplining effect on the acquired firms' management, and improved efficiency in terms of productivity can be expected.

The matching theory[9] of ownership changes, developed by Lichtenberg and Siegel (1992a), states that firms permanently evaluate the fit between the owner and its plants. The quality of the match, which is reflected in productivity levels, is a key determinant in the firm-level decision to maintain or to relinquish the ownership of a plant. The implications of the model are straightforward: plants with a low productivity will be subject to ownership changes, and the ownership change improves the match, resulting in a productivity growth.

Another strand of literature focuses on managers' opportunistic behavior as a reason for mergers. One explanation for mergers is that managers suffer from hubris, e.g. they overestimate their abilities to improve the target's performance (Roll, 1986). In addition, empire-building motives may also play a role: managers have personal interests, e.g. higher expected financial rewards when they widen the firm's size and scope (Baumol, 1959). A further motive is that managers merge to entrench themselves and make it costly to shareholders to replace them (Shleifer and Vishny, 1989). If M&A occurs for these reasons, it is difficult to make

[8] This argument is related to arguments discussed in the literature about heterogeneous firms, their productivity, and the role of fixed costs for export-activities and engagements in FDI (e.g. Melitz, 2003; Helpman, Melitz, and Yeaple, 2004): only more productive firms are able to bear the fixed costs of exporting and FDI.

[9] This theory was developed by Jovanovic (1979) and describes job turnover, and it is applied in labor market studies.

assumptions about the pre-merger performance of both acquiring and target firms. However, I believe that mergers that are not primarily motivated by efficiency reasons are likely to decrease productivity, or at least leave productivity unchanged.

Up to now, I have not discussed the differences between foreign and domestic mergers. There are arguments which expect stronger productivity effects from foreign M&A. First, it has become a stylized fact in international trade literature that multinational firms outperform domestic firms (e.g. Helpman, Melitz, and Yeaple, 2004; Greenaway and Kneller, 2007) because of a higher stock of knowledge capital like brands, patents, technologies, etc. (Markusen, 2004), and a better ability to exploit firm-specific assets (Caves, 1996). In foreign takeovers, targets realize additional efficiency gains from their multinational parents, for example, because of a costless transfer of the acquiring firms' assets (Bellak, Pfaffermayr, and Wild, 2006), or they take advantage from the multinationals' network (Bellak and Pfaffermayr, 2002). Second, there are also good arguments to assume that resources flow in the other direction, i.e. from acquired foreign affiliates to their parent companies (Dunning, 1998). If multinational firms are asset-seeking, they may want to acquire local firms with better technology and know-how than their own, leaving limited or even no scope for a knowledge transfer from parents to affiliates (Salis, 2008). This implies that targets are cherries, and post-merger changes mostly occur in the acquiring firm. In contrast, productivity effects may be stronger from domestic than foreign M&A. Information about target firms decreases as the distance between acquirers and targets increases. For this, acquirers choose targets which are close to them. A better touch of the local market enables intensive rationalization and radical reforms (Lehto, 2006). This may positively affect the firms' productivity, more than after foreign takeovers.

In addition to the theoretical discussions so far, there are further aspects specific about horizontal and non-horizontal mergers worth mentioning. Horizontal mergers can lead to increased efficiency through economies of scale, i.e. if the average costs decline as the firms' size increase, mergers reduce the combined firms' average costs (Mueller, 2003a). In this case, mostly smaller and high cost firms should horizontally merge, because they can expect the largest gains.[10] Horizontal mergers can also generate efficiencies by eliminating costly duplications, for example if they combine the firms' sales or distribution forces (Pesendorfer, 2003). A vertical merger can create efficiencies if it enables firms to buy inputs at lower prices. This was discussed by Spengler (1950) for the case of an upstream and downstream monopolist, both setting their prices above marginal costs. A vertical integration of the upstream monopolist

[10] This is not empirically confirmed (Mueller, 2003a).

eliminates double marginalization, and input will be transferred within the firm at marginal costs. Further efficiency gains from vertical mergers are possible, either from a reduction of transaction costs, if goods or services are transferred between firms (Williamson, 1975), or in the presence of hold-up problems (Church, 2004). Hold-up describes the situation in which firms behave opportunistically and underinvest if there are relation-specific investments but incomplete contracts. A vertical integration may instead promote investments and reduce the hold-up problem.[11] Conglomerate mergers enable firms to produce a range of goods, or to bundle several products together, and the firms may realize efficiencies due to economies of scale, scope, and learning in production and distribution (Church, 2004; Church, 2008a). Based on these theoretical arguments, no clear prediction is possible for pre- or post-merger productivity of horizontally and non-horizontally merging firms.

The literature also discusses increased market power as a motivation for mergers. All horizontal mergers are at least slightly anticompetitive (Farrell and Shapiro, 2001), because they reduce the number of firms in the same market, and they involve a loss of direct competition between firms. Vertical mergers increase market power, if they establish market entry barriers at one or more links in the vertical production chain (Comanor, 1967). For example, input and customer foreclosure are ways to establish entry barriers.[12] Anticompetitive effects emerge from conglomerate mergers due to tacit collusion (Mueller, 2003a), i.e. firms compete against each other over time and cooperate with their rivals in order to maintain higher prices. Similarly, the same may apply if firms meet in different markets at the same time. No clear prediction about pre-merger performance of plants is plausible, if increased market power is the motivation for mergers. However, economic intuition allows for some expectations about post-merger effects: horizontal mergers reduce the number of firms in the market, leading to lower competitive pressure on the merging firms. Thus, an increase in market power seems more likely to follow this type of merger compared to non-horizontal mergers (Gugler, Mueller, Yurtoglu, and Zulehner, 2003). Since there is evidence that competition improves productivity (e.g. Nickell, 1996), I expect productivity to decrease if the competitive pressure decreases after horizontal M&A.

As a conclusion, the theory does not give clear predictions about pre- and post-merger productivity performance of firms, and hence, the question is passed to empirical research. But as I will briefly show below, empirical studies also do not give a clear answer.

[11] Grossman and Hart (1986) pointed out that integration does not entirely solve the hold-up problem.

[12] Input foreclosure occurs if integrated upstream firms have market power and do not sell to downstream rivals in the post-merger period any more, or they sell at higher prices, or offer lower quality to downstream rivals. Consumer foreclosure occurs if a downstream firm no longer sources supply from independent upstream firms but only from the integrated upstream firm after the merger (Church, 2008b).

Related literature: My paper is integrated into the empirical literature about research on pre- and post-merger productivity performance of firms. There is some empirical support that acquiring firms prefer cherries rather than lemons. For the US, McGuckin and Nguyen (1995) observed higher pre-merger productivity for smaller plants but lower for larger plants, and Ollinger, Nguyen, Blayney, Chambers, and Nelson (2006) found above-industry productivity of acquired plants prior to the merger. The empirical support is even stronger for cross-border M&A: Girma and Görg (2004) found evidence for firms in the UK, Bellak, Pfaffermayr, and Wild (2006) for Austrian firms, and Salis (2008) for Slovenian firms. In a study for Norway, Balsvik and Haller (2011) estimated that foreign firms acquire cherries, while domestic firms acquire lemons. For Germany, Mattes (2010) found that foreign firms acquire both cherries and lemons. In contrast to these results, Lichtenberg and Siegel (1992a) observed low pre-merger productivity levels of acquired plants for the US, Gioia and Thomsen (2004) found support that targets are lemons in a study for Denmark, and Castellani and Zanfei (2004) also provided evidence that domestic Italian firms acquired by foreign investors are not the most productive.

A majority of recent studies that mostly focused on cross-border M&A, examined positive post-merger productivity effects. This is in contrast to earlier studies, e.g. from Ravenscraft and Scherer (1987), and possible explanations for different research results have already been mentioned at the beginning of this paper. For example, Girma, Thompson, and Wright (2006) estimated positive productivity effects for acquired firms in the UK after acquisitions of European or US multinationals. Similar results for the UK were found in Conyon, Girma, Thompson, and Wright (2002b), and in Griffith and Simpson (2004). Studies from Arndt and Mattes (2010) for Germany, Petkova (2009) for India, Arnold and Javorcik (2009) for Indonesia, Piscitello and Rabbiosi (2005) for Italy, Lichtenberg and Siegel (1992a), Maksimovic and Philips (2001), and Maksimovic, Phillips, and Prabhala (2011) for the US also estimated positive effects after cross-border M&A. Again for the US, Pesendorfer (2003) examined only horizontal mergers and found positive effects, too. For France, Bertrand and Zitouna (2008) also examined horizontal mergers and estimated positive productivity effects, but effects were larger for cross-border M&A compared to domestic M&A. Similarly, Balsvik and Haller (2011) found that Norwegian target plants increase productivity after a foreign acquisition, but decrease after domestic mergers. In a study about foreign acquisitions of Swedish firms, Bandick (2011) estimated a productivity increase after vertical mergers, but no effects after horizontal acquisitions. Further studies that did not find changes in productivity after cross-border M&A are from Bellak, Pfaffermayr, and Wild (2006) for Austria, Mattes (2010) for Germany, and Salis (2008) for Slovenia. For the US, McGuckin and Nguyen (1995) also observed post-merger productivity improvements for acquired plants, but productivity losses for the acquirer's existing plants. Schoar (2002) found similar results. Negative productivity effects for acquired plants after international takeovers were estimated in a study from Gioia and Thomsen (2004) for Denmark.

3.3 The data

The panel dataset used in the following empirical analysis is a combination of two datasets: the Establishment Panel (Betriebspanel) of the Institute for Employment Research, Nuremberg (Institut für Arbeitsmarkt- und Berufsforschung, IAB), and the M&A DATABASE of the University of St. Gallen. The dataset was created by TNS Infratest Sozialforschung GmbH München.[13]

The IAB Establishment Panel is a representative employer survey for Germany and contains a wide range of questions on topics related to employment policy. The survey, which has existed since 1993 is carried out annually, and currently contains around 16,000 plants of all sizes and sectors of the economy. The data does not report about M&A activities of plants, and thus, the data was linked to the M&A DATABASE of the University of St. Gallen. The M&A DATABASE contains information about 65,000 transactions since 1985 for Germany, Austria, and Switzerland. For every deal there is information about the acquirer, target, and seller, as well as about sales, profits, employees, location, and sector. Moreover, the data also includes information if the merger was horizontal, vertical forward, vertical backward, conglomerate, or concentric.[14] Between 1996 and 2005 - the observation period in which the plants in this dataset merged - there is information of about 23,717 transactions with 40,736 German firms involved.[15]

Based on these two independent datasets, TNS Infratest constructed a new dataset which consists of two groups of plants. These groups exhibit a similar structure with regard to sector, size, location, and legal form. The first group, which is the treatment group, consists of plants which merged between January 1996 and December 2005. To create this group of merging plants,[16] all plants which appeared both in the M&A DATABASE and in the IAB Establishment

[13] The creation of this dataset preceded a pilot study from Bellmann and Kirchhof (2006). They used Thomson ONE Banker instead of M&A DATABASE. The latter's advantage is that it also includes smaller firms.

[14] In a concentric merger, firms from different but neighboring industries merge. It is comparable to conglomerate mergers with complementary or neighboring products (Church, 2004).

[15] The number of companies is higher than the number of transactions because up to three companies can be involved in a merger - as acquirer, object and seller.

[16] I use the terms "merged plants" and "treated" interchangeably.

Panel had to be identified. Plants were only assigned to the treatment group if they were surveyed at least once prior and once after the merger in order to have enough information. If plants merged more often between 1996 and 2005, they may also appear more often in the dataset. The treated group consists of 7,801 observations from 958 different plants.[17]

The second group, the control group, consists of plants that had not merged between 1980 and 2005. This group was created in a way that control plants are as similar as possible to merged plants in the treatment group, that is, controls should act as statistical twins to those treated plants. Hence, treated plants were categorized according to sector, size category, location in West or East Germany, and legal form. Next, controls had to be identified with the same combination of sector, size, location, and legal form. However, not for every treated could a control be found with identical characteristics. In addition, control plants had also to be surveyed at least twice between the years 1993 and 2006,[18] and they were only kept in the dataset if they had not been involved in any merger activities since 1980. This was controlled through other datasets or the plants' websites. Several plants exist that appear more often in the control group because plants can serve several times as a control in the referred observation period. The control group consists of 1,009 observations from 291 different plants.[19]

The dataset has some useful features and differs from datasets used in other studies in a number of ways:

- The data allow for differentiation between acquirers, targets, and sellers, and between horizontal and non-horizontal mergers.
- The dataset includes plants of all sizes, i.e. small and medium-size plants can also be analyzed.
- Plants are from different sectors, from either West or East Germany, and have different legal forms.
- Since the data is at plant-level, each plant can be assigned to a specific sector. With firm-level data, this is often not possible if firms are multi-plant firms.
- All plants in the control group had not merged since 1980.

[17]This means that the treatment group consists of 958 different plants which all have different identification numbers in the IAB Establishment Panel.

[18]The first wave of the IAB Establishment Panel is from 1993, and the survey of 2006 provides information about the year 2005.

[19]See appendix A for a more detailed description of the construction of both treatment and control groups.

- The dataset allows for a differentiated econometric analysis because of the availability of a control group and a rich set of variables.

The empirical investigation in the next section requires some modifications of the dataset. First, if a plant was involved in several mergers within a single year, the mergers were considered as only one merger, because several information is on a yearly basis. Second, I analyze labor productivity over an observation period of four years from $t = 1$ to $t = 4$. Thus, treated and control observations were only kept if data existed for all four years. Third, treated plants were dropped if a plant merged more than once within three years. This allows assigning effects to one specific merger. Fourth, labor productivity is defined as sales per employee, and therefore I excluded plants which reported balance sheet totals instead of sales. Fifth, I also dropped plants which were surveyed without reporting about their sales, because productivity is measured as sales per employee. Sixth, observations which exhibit abnormal values for labor productivity growth rates were deleted. I define abnormal values if growth rates between $t = 3$ and $t = 4$ or $t = 1$ and $t = 4$ deviate two standard deviations from the respective industry average.[20] Seventh, seller plants were also excluded from the dataset because the number of observations is too small. And eighth, I only keep treated plants that merge domestically because the number of foreign M&A is also too small for an analysis.

Table 3.1: **Classification of treated plants**

Plants	Horizontal	Non-horizontal	Unkown	Total
Acquirers	46	8	22	76
Targets	22	12	15	49
Total	68	20	37	125

Note: there are plants which merged, but there is no information about the type of merger. These are labeled as "Unknown".

However, these modifications reduce the number of observations significantly:[21] from 7,801 to 125 treated, and from 1,009 to 520 untreated observations.[22] Table 3.1 presents an overview

[20] These extreme values may be due to errors or rare events. For example, consider a firm that produces a certain machine in a year, and reports only low sales in the same year. If the firm sells the machine in the next year, it will report high sales. These extreme numbers may have a high impact on the empirical results (Wagner, 2007b).

[21] Gugler and Yurtoglu (2004) used the Global Mergers and Acquisitions database for their study. They also analyzed only a small fraction of the original dataset, i.e. from a large sample of 140,289 mergers there was only sufficient data for 646 mergers.

[22] As stated above, I only kept a treated plant more than once in the dataset if there are at least three year between the mergers in order to avoid overlapping effects. There are also untreated plants that appear more often in the control group, because they are surveyed for several times in different waves. Therefore, one might think that it is not a good idea that a single plant is used more often as a control

of the number of treated plants and distinguishes between acquirers and targets, and plants involved in horizontal and non-horizontal mergers. Unfortunately, the type of merger is not always known for every treated plant. Moreover, due to a small number of non-horizontally merging plants, the robustness of the results for this subgroup is limited in the following empirical analysis.

3.4 Empirical investigation

3.4.1 Empirical strategy

Productivity measurements: Productivity is a measure of firms' efficiency performance. In general, productivity is defined as the relation of a firm's output to its input, and there are different ways to measure it. For example, total factor productivity (TFP) measures changes in output not explained by changes in inputs such as labor or capital. TFP is a theoretically superior measure of productivity because it takes all input factors into account. However, the calculation of TFP requires more data about inputs, e.g. capital etc., which is often not available in datasets. For this reason, several empirical studies alternatively measure the ratio of a firm's output to only one single input, e.g. labor or capital. This study is based on labor productivity because the dataset includes information about labor and output, measured in sales, but there is not sufficient information about other inputs to properly calculate TFP. More precise, for each firm i at time t, I define labor productivity LP_{it} as sales S_{it} (measured in Euro) per employees N_{it}. The number of employees includes all workers, i.e. independent of being liable to social security or not:

$$LP_{it} = S_{it}/N_{it}. \tag{1}$$

Plant- vs. firm-level analysis: As argued at the beginning of my paper, studies exist that analyze effects at plant-level, while others use firm-level data. Headquarter activities like

because it was surveyed several times, while another single plant will be used only once because it was surveyed less often. Consequently, the results for controls may be biased towards the plants that were surveyed more often. Nevertheless, this should be no great problem for the following reasons: first, I do not believe that there is a systematic bias in relevant variables in those plants that were surveyed more often. In particular, why should a plant that was surveyed in the IAB Establishment Panel more often than another plant exhibit systematically higher or lower productivity changes over time? To be sure, I tested for a correlation and did not find any evidence. Second, if I would allow keeping an untreated observation to appear only once in the control group, the control group would shrink to approximately one fourth of its size, and useful information would be lost. Nevertheless, I also performed the empirical investigation with a control group with each individual plant appearing only once in the dataset as a robustness check. As expected, the estimations based on this smaller control group were similar.

marketing, R&D, finance operations, etc. are at firm-level, whereas plant-level includes activities like production and assembling (Barba Navaretti and Venables, 2004). Thus, a merger may have a different effect on firm-level compared to plant-level. For example, if M&A generates synergy effects in R&D or marketing, they should primarily be measured at firm-level. If, instead, M&A leads to improved production processes, effects should be measured at plant-level.

The dataset used in this paper is at plant-level. Hence, estimations of performance changes of the overall firm are not possible. However, this must not be a drawback if the research focus is on smaller and medium-size firms which are often single-plant firms. Moreover, a plant-level analysis has the advantage that plants can be assigned easier to a specific industry sector compared to firms (Bellmann and Kirchhof, 2006). Finally, the plant is the unit which is fully involved in the merger and captures the whole productivity effect of M&A, while the effect may disperse at firm-level which measures the average productivity of all plants.

Observation period: The observation period in this paper covers four years, i.e. from $t = 1$ to $t = 4$. All mergers occur in $t = 2$. Due to this construction plants' pre-merger and post-merger productivity performance can both be analyzed. Of course, longer observation periods prior to the merger would be desirable, e.g. because a decreasing performance could be the trigger for a merger, or because of the existence of Ashenfelter's Dip:[23] if plants prepare themselves for the merger, performance can already be affected prior to the merger, and estimations about post-merger effects may then be biased. However, I am not able to lengthen the observation period to more than one year prior to the merger because of data limitations. For this reason, I start the observation period one year prior to the merger, similar to many other existing studies (e.g. Girma and Görg, 2007).

I calculate post-merger performance in years $t = 3$ and $t = 4$, because it seems plausible to assume that it takes some time for the effects to arise after mergers (e.g. for the reorganization of the merging plant's production).[24] This also allows calculating post-merger growth rates instead of levels, and it reduces selection bias that arises due to unobserved time-invariant differences between both groups.[25] In addition to changes between years $t = 3$ and $t = 4$ ("post-merger period"), I also analyze growth rates over the whole observation period between

[23] The Ashelfelter Dip describes that the unemployed people's attempts at job seeking decrease shortly before they participate in a labor market program.

[24] Data from $t = 4$ corresponds to at most three years after the merger, if the merger occurred at the beginning of year $t = 2$ (see appendix B). Maksimovic, Phillips, and Prabhala (2011) found that most of the restructuring occurs within a three-year period. If they changed the time window to five years, the results did not change.

[25] I will discuss this issue later more precisely.

$t = 1$ and $t = 4$ ("total period"). This allows capturing changes that occur within several weeks or months around the merger.[26]

For the empirical setting, I follow Schank, Schnabel, and Wagner (2010) and use a "rolling observation window".[27] All mergers in the dataset occur between 1996 and 2005, and these years correspond to $t = 2$. As a consequence, year $t = 1$ corresponds to a year between 1995 and 2004, and year $t = 4$ corresponds to a year between 1998 and 2007. Hence, this leads to ten cohorts with a respective four-year window (1995 - 1998, 1996 - 1999, ..., 2004 - 2007). For controls, the first possible year for $t = 1$ is 1993 (the starting year of the IAB Establishment Panel), and the last possible year for $t = 4$ is 2005, leading to ten cohorts with a four-year window (1993 - 1996, 1994 - 1997, ..., 2002 - 2005).[28] Finally, sales are deflated by the aggregated consumer price index over the whole observation period.

The selection problem and the estimation methods: The objective of this paper is to analyze productivity effects from M&A. A simple comparison of merged plants with non-merged plants may lead to biased results if plants are not selected randomly to the group of merged plants. For example, if mostly better performing plants, or better performing targets are subject to M&A, a simple comparison to a group of non-merged plants will lead to the conclusion that merging plants are more productive than non-merging plants. However, this simple comparison will not reveal the true effects of M&A, because it does not take into account that plants with certain pre-merger characteristics self-select in M&A activity, i.e. the comparison suffers from a selection bias.

The empirical analysis in this paper focuses on the selection problem and tries to identify the causal effect. As usual in empirical studies, the investigation starts with an analysis of descriptive statistics. I analyze different pre-merger characteristics between both groups, compare productivity levels over time as well as growth rates between both groups.

Nevertheless, descriptive statistics are not a convincing test for self-selection and causal effects. Productivity differences between merging and non-merging plants may be the result

[26] Appendix B provides more information on the construction of the observation period and some remarks on growth rates.

[27] For a better understanding, figure B2 in the appendix shows a graphical illustration of the different cohorts.

[28] The dataset includes plants which merged between 1996 and 2005. However, some of the four-year cohorts for merged plants also cover the years 1995, 2006, and 2007, but there is no information if plants also merged in these years, or only between 1996 and 2005. For example, a plant that merged in 2005 could also merge in 2007 again. If this was true, my results would be biased by overlapping effects due to more mergers. Nevertheless, I choose to keep these observations because I already excluded those multi-mergers from the dataset (see section 3.3), and thus, it is not too likely that the remaining plants also merged in 1995, 2006, or 2007.

of differences in other variables than M&A that determine productivity. For this, I apply a regression analysis to control for variables that influence productivity. The construction of the framework is similar to Schank, Schnabel, and Wagner (2010): the regression allows looking at differences in the average plant productivity between both groups over time, and it controls for plant characteristics that are expected to be related to a plant's average productivity performance. However, the literature does not consider regression analysis as a reliable method to clearly solve selection problems and identify causalities (e.g. Backhaus, Erichson, Plinke, and Weiber, 2010).[29]

Newer econometric evaluation methods like propensity score matching from Rosenbaum and Rubin (1983, 1985) present a way to solve the selection problem and identify causalities. The problem is that a faster productivity growth of plants that merged does not necessarily reflect a causal effect of M&A on the plants' productivity. Instead, it could also be that plants with higher productivity self-select in the merged group, but would have experienced higher growth even without merging. However, both states are never observable at the same time, which leads to the problem of the missing counterfactual situation (Schank, Schnabel, and Wagner, 2010). Propensity score matching allows replacing this missing counterfactual by the construction of an appropriate control group, i.e. the method pairs merged with non-merged plants that are similar in their pre-merger characteristics, and therefore, exhibit a similar probability of merging. Performance differences can then be attributed to the merger. As suggested by Blundell and Costa Dias (2000) and Smith and Todd (2005a), I combine matching with a difference-in-differences estimator to compare changes instead of levels. This improves the matching results because it eliminates time-invariant heterogeneity between both groups.

In the following empirical investigation, I first present the results for the treatment group including all treated plants. Then, I consider the results for all subgroups of treated plants: acquirers, targets, plants in horizontal mergers, and plants in non-horizontal mergers.

3.4.2 Descriptive statistics

I start with a first look at the data. Table 3.2 presents summary statistics, and all numbers are measured in $t = 1$.[30] I differentiate between treatment and control groups, and the respective numbers in the tables are absolute numbers, while the numbers in parentheses are percentages.

The table presents the distribution of treated and controls to different size categories. The

[29] See Caliendo (2006), Gelman and Hill (2007), or Angrist and Pischke (2009) for a discussion about regression analysis compared to matching methods.

[30] All calculations in this study are performed with STATA 11.1.

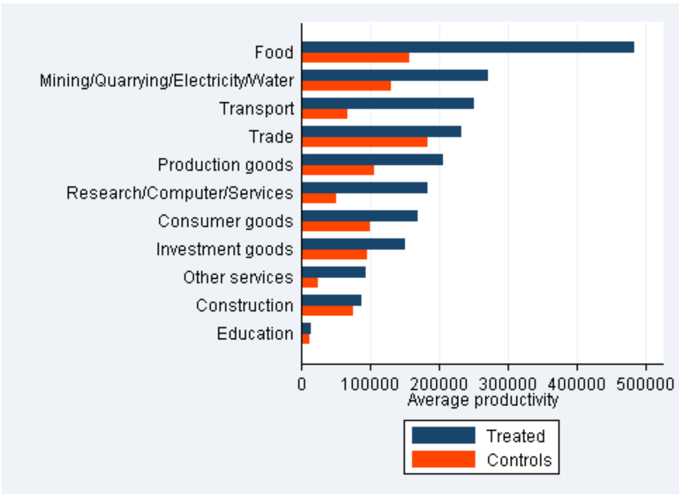
Table 3.2: **Summary statistics: different variables**

Variable	All	Acquirers	Targets	Horizontal	Non-horizontal	Control
Size category						
1-19 employees	10 (8.00)	4 (5.26)	6 (12.24)	4 (5.88)	4 (20.00)	112 (21.54)
20-49 employees	8 (6.40)	6 (7.89)	2 (4.08)	4 (5.88)	2 (10.00)	91 (17.50)
50-99 employees	8 (6.40)	4 (5.26)	4 (8.16)	4 (5.88	1 (5.00)	90 (17.31)
100-299 employees	29 (23.20)	20 (26.32)	9 (18.37)	19 (27.94)	3 (15.00)	171 (32.88)
300-499 employees	12 (9.60)	6 (7.89)	6 (12.24)	7 (10.29)	2 (10.00)	34 (6.54)
>=500 employees	58 (46.40)	36 (47.37)	22 (44.90)	30 (44.12)	8 (40.00)	22 (4.23)
Total	125 (100.00)	76 (100.00)	49 (100.00)	68 (100.00)	20 (100.00)	520 (100.00)
Sector						
Mining/Quarrying/Electricity	10 (8.00)	8 (10.53)	2 (4.08)	7 (10.29)	2 (10.00)	3 (0.58)
Food	1 (0.80)	1 (1.32)	0 (0.00)	1 (1.47)	0 (0.00)	24 (4.62)
Consumer goods	10 (8.00)	6 (7.89)	4 (8.16)	5 (7.35)	1 (5.00)	29 (5.58)
Production goods	29 (23.30)	13 (17.11)	16 (32.65)	14 (20.59)	4 (20.00)	78 (15.00)
Investment goods	40 (32.00)	21 (27.63)	19 (38.78)	18 (26.47)	9 (45.00)	144 (27.69)
Construction	3 (2.40)	2 (2.63)	1 (2.04)	0 (0.00)	0 (0.00)	26 (5.00)
Trade	13 (10.40)	10 (13.16)	3 (6.12)	9 (13.24)	2 (10.00)	112 (21.54)
Transport	10 (8.00)	9 (11.84)	1 (2.04)	8 (11.76)	0 (0.00)	24 (4.62)
Education	1 (0.80)	0 (0.00)	1 (2.04)	1 (1.47)	0 (0.00)	7 (1.35)
Research/Computer/Ser	2 (1.60)	1 (1.32)	1 (2.04)	1 (1.47)	1 (5.00)	59 (11.35)
Other services	6 (4.80)	5 (6.58)	1 (2.04)	4 (5.88)	1 (5.00)	14 (2.69)
Total	125 (100.00)	76 (100.00)	49 (100.00)	95 (100.00)	20 (100.00	520 (100.00)
Legal form						
Partnership, individually-owned, etc.	12 (9.76)	8 (10.67)	4 (8.33)	7 (10.45)	2 (10.00)	105 (21.00)
Limited	74 (60.16)	41 (54.67)	33 (68.75)	36 (53.73)	14 (70.00)	376 (75.00)
Limited by shares	37 (30.08)	26 (34.67)	11 (22.92)	24 (35.82)	4 (20.00)	19 (3.80)
Total	123 (100.00)	75 (100.00)	48 (100.00)	67 (100.00)	20 (100.00)	500 (100.00)
Location in East Germany	38 (30.40)	19 (25.00)	19 (38.78)	24 (35.29)	6 (30.00)	136 (26.20)
Wage per month (mean)	2215.4	2330.6	2031.1	2249.8	2126.1	1715.9
Export rate (% of sales, mean)	26.83	25.17	29.59	24.74	31.00	11.46
Plant is in foreign ownership	13 (15.85)	5 (10.64)	8 (22.86)	7 (11.48)	1 (8.33)	5 (1.43)
Further training	117 (94.35)	70 (93.33)	47 (95.92)	65 (95.59)	18 (90.00)	340 (66.67)
Single-plant firm	37 (29.84)	21 (27.63)	16 (33.33)	21 (30.88)	6 (31.58)	397 (79.72)

Note: numbers refer to $t = 1$. Numbers in parentheses are percentage of total number of plants. Reduced number of observations is due to missing data for several variables.

numbers show that treated plants are larger than control plants. For example, around 46% of treated plants have 500 and more employees, while only 4% of controls are of this size. Next, treated plants are mostly concentrated in the production goods and investment goods sector, while the largest fraction of controls belongs to the investment goods and trade sector.[31] Considering the legal form of plants, most are "Limited". However, treated plants have more often the legal form of "Limited by shares" compared to controls. In addition, the table also shows how many plants are located in West and East Germany within both groups of plants, and the distribution is similar. Furthermore, treated plants pay higher wages, and have higher export rates. This is consistent with the findings of a higher productivity in treated plants because more productive plants exhibit a higher export activity (e.g. Melitz, 2003; Greenaway and Kneller, 2007), and exporting firms also pay higher wages (e.g. Schank, Schnabel, and Wagner, 2010). Moreover, treated plants are more often in foreign ownership. They also offer more further training to their employees prior to the merger, which is plausible for plants that are more productive. Finally, approximately 30% of treated plants are single-plant firms compared to approximately 80% of control plants. In general, the numbers and distributions do not vary substantially between subgroups of treated.

Figure 3.1: **Labor productivity in different sectors**



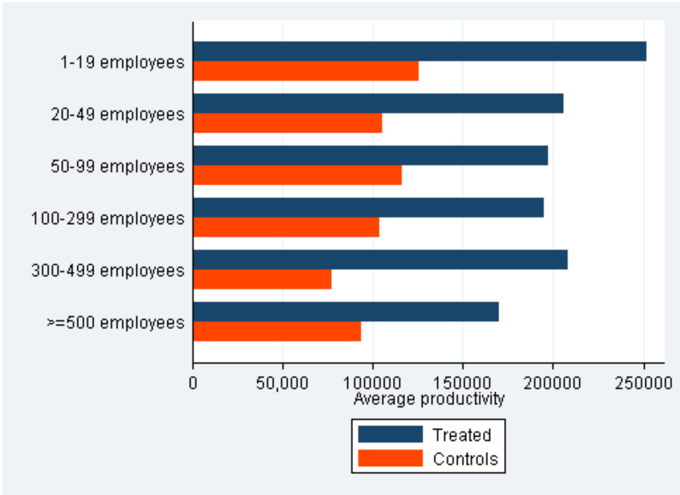
Note: numbers refer to year $t = 1$.

In the next step, I compare the productivity of the treatment group including all treated plants to the control groups in year $t = 1$, depending on the sector they belong to and their size. Figure 3.1 shows that for almost all sectors the average productivity level is higher for treated plants prior to the merger. Productivity levels differ substantially between sectors, but differences have to be analyzed with care due to a small number of underlying observations in several sectors (see table 3.2). Figure 3.2 reveals a similar picture: treated plants have a higher average productivity before they merge, independent of their size in terms of employees.[32]

[31] See appendix A for a description of sector classification.

[32] The corresponding figures for subgroups look similar but are not reported.

Figure 3.2: **Labor productivity in different size categories**



Note: numbers refer to year $t = 1$.

Table 3.3 presents the average productivity levels in the treatment group, in subgroups of treated, and in the control group over the whole observation period between $t = 1$ to $t = 4$. Merging plants have an average pre-merger productivity of approximately 185,600 (sales in Euro per employee), which reflects a pre-merger differential of around 94% compared to plants that had not merged. This productivity gap persists over the whole observation period. The picture is similar for all subgroups with a productivity differential between 82% (targets) and 102% (horizontal M&A). I also apply a t-test to test the statistical significance of productivity differences between merged plants and control plants. The test does not assume equal variances in the respective comparing groups.[33] For the treatment group, and for each subgroup of treated, I test the null hypothesis H_0 in the years $t = 1$ to $t = 4$: mean of labor productivity of treated = mean of labor productivity of controls, against the alternative hypothesis H_1: mean of labor productivity of treated $\neq$ mean of labor productivity of controls. P-values of at most 0.01, 0.05, or 0.1 indicate that the null hypothesis H_0 can be rejected at the 1%-, 5%- or 10%-error level in favor of the alternative hypothesis H_1. The results show that the differences in productivity are statistically significant for the treatment groups and all subgroups in all years, mostly at the 1%-significance-level. I also apply t-tests to compare productivity means between acquiring and target plants, and between plants in horizontal and non-horizontal mergers for all years, but differences are not statistically significant at any level (the results are not reported).

Table 3.3 shows small changes in the productivity levels for controls over time, but changes

[33] I apply Levene's statistic for a test of the equality of variances.

Table 3.3: **Summary statistics: labor productivity**

All	Mean	(Std. Dev.)	N	P-value
Productivity in t=1	185686	(133530)	125	0.000
Productivity in t=2	193291	(139130)	125	0.000
Productivity in t=3	192222	(136985)	125	0.000
Productivity in t=4	194691	(139854)	125	0.000
Growth from t=3 to t=4	0.015	(0.160)	125	0.322
Growth from t=1 to t=4	0.043	(0.367)	125	0.113
Acquirers	**Mean**	**(Std. Dev.)**	**N**	**P-value**
Productivity in t=1	192909	(137048)	76	0.000
Productivity in t=2	208563	(152415)	76	0.000
Productivity in t=3	202710	(144060)	76	0.000
Productivity in t=4	205274	(146611)	76	0.000
Growth from t=3 to t=4	0.021	(0.173)	76	0.283
Growth from t=1 to t=4	0.060	(0.435)	76	0.160
Targets	**Mean**	**(Std. Dev.)**	**N**	**P-value**
Productivity in t=1	174483	(128467)	49	0.000
Productivity in t=2	169605	(112991)	49	0.000
Productivity in t=3	175955	(124918)	49	0.000
Productivity in t=4	178277	(128412)	49	0.000
Growth from t=3 to t=4	0.005	(0.140)	49	0.705
Growth from t=1 to t=4	0.016	(0.223)	49	0.391
Horizontal	**Mean**	**(Std. Dev.)**	**N**	**P-value**
Productivity in t=1	192977	(153094)	68	0.000
Productivity in t=2	202562	(164727)	68	0.000
Productivity in t=3	198637	(157976)	68	0.000
Productivity in t=4	201408	(157998)	68	0.000
Growth from t=3 to t=4	0.021	(0.187)	68	0.328
Growth from t=1 to t=4	0.047	(0.465)	68	0.302
Non-horizontal	**Mean**	**(Std. Dev.)**	**N**	**P-value**
Productivity in t=1	176996	(137674)	20	0.016
Productivity in t=2	169666	(96990)	20	0.003
Productivity in t=3	175183	(108436)	20	0.004
Productivity in t=4	179437	(126317)	20	0.007
Growth from t=3 to t=4	-0.001	(0.115)	20	0.911
Growth from t=1 to t=4	0.022	(0.158)	20	0.353

Control	Mean	(Std. Dev.)	N
Productivity in t=1	95699	(68071)	520
Productivity in t=2	95015	(66262)	520
Productivity in t=3	94076	(62966)	520
Productivity in t=4	94050	(63984)	520
Growth from t=3 to t=4	-0.004	(0.284)	520
Growth from t=1 to t=4	-0.014	(0.294)	520

Note: p-values refer to the t-test about statistical significance of difference of means between treated and control groups.

are larger for treated and subgroups of treated. For this, I calculate percentage post-merger productivity changes between $t = 3$ and $t = 4$, and also changes for the total period between $t = 1$ and $t = 4$, approximated by logarithms:

$$lnLP_{(3-4)i} = lnLP_{4i} - lnLP_{3i}, \tag{2}$$

and

$$lnLP_{(1-4)i} = lnLP_{4i} - lnLP_{1i}. \tag{3}$$

$lnLP_{(3-4)i}$ and $lnLP_{(1-4)i}$ are the variables for percentage changes between $t = 3$ and $t = 4$, and between $t = 1$ and $t = 4$ for each plant i, and LP_{ti} with $t = \{1, ..., 4\}$ is the variable for labor productivity in year t.

The results are also shown in table 3.3. For the treatment group including all merging plants, the post-merger growth rate is 1.5%, and 4.3% for the whole observation period. The subgroups also exhibit positive post-merger changes, except the subgroup of non-horizontal mergers, and changes are between 0.5% (targets) and 2.1% (acquirers, horizontal M&A). With respect to changes for the total periods, all subgroups exhibit a productivity increase between 1.6% (targets) and 6.0% (acquirers) on average. In contrast, the control group faces a decrease in the post-merger period (-0.4%), and also in the total period (-1.4%). Again, I apply a t-test to test the statistical significance of productivity changes. In particular, I test the null hypothesis H_0: mean of percentage change of labor productivity of treated = mean of percentage change of labor productivity growth of controls. The alternative hypothesis H_1 is: mean of percentage change of labor productivity growth of treated $\neq$ mean of percentage change of labor productivity growth of controls. As a result, changes are not statistically significant at any usual significance level. That is, plants that merge do not exhibit productivity changes that are statistically significantly different from plants that do not merge.

The results of the descriptive statistics displayed in tables 3.2 and 3.3 show that merging plants are different from plants that do not merge even before the merger. The heterogeneity refers to several characteristics. The pre-merger labor productivity of merging plants is almost twice as high as of control plants, and this is true for all subgroups. The findings support the view that more productive plants self-select in merger activity. However, according to descriptive statistics, there is no evidence that M&A affects productivity, i.e. mergers do neither statistically significantly change the plants' productivity in the post-merger period between $t = 3$ and $t = 4$ nor in the whole observation period between $t = 1$ and $t = 4$.

3.4.3 Regression analysis

In descriptive statistics I found statistically significant differences in productivity levels, but not for growth rates between treated and control groups. However, I do not consider descriptive statistics as a reliable estimation strategy to analyze questions about self-selection and causality. In this subsection, I perform an OLS-regression analysis, and I analyze the effect of some interaction variables of interest. I also control for several plant characteristics that can be expected to be correlated with a plant's productivity. The logarithmized regression model[34] is specified as follows:

$$lnLP_{it} = \beta_0 + \beta_1 MA_i + \sum_{t=2}^{4} \beta_t (MA_i * PERIOD_t) + \beta_5 CONTROL_{it} + \epsilon_i \qquad (4)$$

$lnLP_{it}$ is the logarithm of labor productivity, i is an index for a plant, and t is the index for the years $t = 1$ to $t = 4$. MA_i represents a dummy with value one if a plant i merged, and zero if the plants is a control. The coefficient β_1 measures the percentage difference of the average productivity between the treated and control groups in $t = 1$. The interaction terms $MA_i * PERIOD_t$ control for changes in productivity over time. The term is a product of the dummy variable for M&A (MA_i) and a dummy variable for years $t = 2$ to $t = 4$ ($PERIOD_t$). The coefficients β_t measure if the difference in average productivity between both groups changes over the years $t = 2$ to $t = 4$. The vector $CONTROL_{it}$ includes different variables - see below - that can be expected to impact a plant's productivity. Finally, ϵ_i is an error term. I perform five different regression specifications: the first for the group of treated plants, and the others for each subgroup of treated.

The results of the regressions are shown in table 3.4. In the first regression specification that included all merging plants (first column), the coefficient of the M&A dummy is 0.488. This corresponds to a productivity differential of approximately 63%.[35] This magnitude is high from an economic point of view, and the difference is statistically significant at the 1%-level. The coefficients of the interaction dummies that control for productivity changes over time are of low magnitude, but they are not statistically significant at any usual significance level. This implies that the pre-merger productivity gap between treated and controls does not statistically significantly change over the years $t = 2$ to $t = 4$.

The regression framework also includes control variables which are expected to be related to a plant's productivity: first, larger plants may exhibit higher productivity, e.g. due to economies of scale. Thus, I include a variable for plant size, measured as the logarithm of the number of employees, and the squared logarithm of employees. Both variables are not statistically significant. In addition, a plant's productivity may also be affected by the employment and qualification structure of its workforce. For this, I include variables for different employment

---

[34] This specification is similar to Schank, Schnabel, and Wagner (2010).

[35] In loglinear regression models the coefficients of explaining variables can be transformed into a percentage change. Here, $exp\{0.488\} \approx 1.629$, and this corresponds to a productivity differential of approximately 63% (Verbeek, 2005).

Table 3.4: **OLS-regression (dependent variable: log. labor productivity)**

Variables	All (1)	Acquirers (2)	Targets (3)	Horizontal (4)	Non-horiz. (5)
M&A (D)	0.488*** (0.091)				
M&A*Period=2 (D)	0.013 (0.041)				
M&A*Period=3 (D)	0.021 (0.044)				
M&A*Period=4 (D)	-0.031 (0.052)				
Acquirer (D)		0.542*** (0.113)			
Acquirer*Period=2 (D)		0.070 (0.059)			
Acquirer*Period=3 (D)		0.073 (0.063)			
Acquirer*Period=4 (D)		0.043 (0.070)			
Target (D)			0.411*** (0.123)		
Target*Period=2 (D)			-0.069 (0.052)		
Target*Period=3 (D)			-0.027 (0.052)		
Target*Period=4 (D)			-0.139** (0.066)		
Horizontal M&A (D)				0.548*** (0.117)	
Horizontal*Period=2 (D)				0.010 (0.063)	
Horizontal*Period=3 (D)				-0.015 (0.068)	
Horizontal*Period=4 (D)				-0.108 (0.084)	
Non-horizontal M&A (D)					0.274 (0.206)
Non-horizontal*Period=2 (D)					-0.027 (0.074)
Non-horizontal*Period=3 (D)					-0.013 (0.071)
Non-horizontal*Period=4 (D)					-0.074 (0.084)
log. Employment	0.105 (0.080)	0.148 (0.092)	0.094 (0.096)	0.151 (0.102)	0.059 (0.121)
Squared log. Employment	-0.012 (0.009)	-0.017* (0.010)	-0.009 (0.011)	-0.018* (0.011)	-0.004 (0.014)
Proportion of skilled employees	0.334*** (0.085)	0.321*** (0.087)	0.343*** (0.085)	0.365*** (0.089)	0.329*** (0.088)
Proportion of management	0.695 (0.468)	0.540 (0.435)	0.774 (0.502)	0.566 (0.460)	0.411 (0.421)
Proportion of apprentice	-0.326 (0.417)	0.004 (0.382)	-0.117 (0.408)	-0.057 (0.393)	-0.022 (0.390)
Proportion of female employees	-0.185* (0.105)	-0.181* (0.104)	-0.160 (0.110)	-0.162 (0.106)	-0.173 (0.110)
log. Investment p. employee	0.024*** (0.005)	0.021*** (0.006)	0.023*** (0.005)	0.019*** (0.005)	0.021*** (0.006)
Further training (D)	0.102** (0.048)	0.073 (0.047)	0.111** (0.050)	0.097** (0.048)	0.083* (0.049)
Legal form "Limited" (D)	-0.104* (0.055)	-0.126** (0.054)	-0.150*** (0.058)	-0.123** (0.056)	-0.184*** (0.056)
Legal form "Limited by shares" (D)	-0.228*** (0.083)	-0.273*** (0.090)	-0.165* (0.085)	-0.235** (0.091)	-0.079 (0.094)
Location in East Germany (D)	-0.277*** (0.057)	-0.265*** (0.060)	-0.265*** (0.060)	-0.280*** (0.060)	-0.253*** (0.064)
Foreign owned plant (D)	0.208** (0.102)	0.233* (0.119)	0.199* (0.108)	0.226** (0.111)	0.324** (0.126)
Single-plant firm (D)	-0.112** (0.054)	-0.080 (0.056)	-0.125** (0.056)	-0.119** (0.058)	-0.079 (0.056)
Constant	11.940*** (0.386)	10.046*** (0.467)	9.961*** (0.501)	11.258*** (0.346)	9.482*** (0.414)
Observations	2369	2186	2096	2168	1982
R^2	0.568	0.569	0.593	0.574	0.590

Notes: robust standard errors in parentheses (adjusted for intragroup correlation). *** p<0.01, ** p<0.05, * p<0.1. (D) means variable is a dummy. Reference categories for legal form is "Partnership, individually-owned, public, and others". Regressions include dummies for sectors and years. The reduced number of observations is due to missing data for several variables. Data source: IAB Establishment Panel, M&A DATABASE St. Gallen.

groups, measured as a proportion within the plant's total workforce:[36] skilled employees[37], management[38], and apprentices. According to the regression, the higher the proportion of skilled employees, the higher the plant's labor productivity, and the coefficient is statistically significant. Contrary, the coefficients for the proportion of management and apprentices are not significant. However, there is a statistically significantly negative correlation between the proportion of female employees and the plant's productivity, and this may be explained by lower productive female part time workers. Additionally, the more a plant invests, measured as logarithmized investments per employee, the higher the plant's labor productivity. The coefficient is highly significant. I also control for the impact of further training which is statistically significantly positive. The coefficients of the dummies for the legal forms "Limited" and "Limited by shares" are both negative and significant, implying that plants that have these legal forms are less productive compared to plants of the reference group "Partnership, individually-owned, etc.". Plants located in East Germany obviously have a statistically significantly lower productivity, but if plants are in foreign property, they exhibit significantly higher productivity. Finally, if plants are single plants, they have a significantly lower productivity.

The regression specifications 2 to 5 in table 3.4 refer to subsections of treated. The respective M&A dummy is one if the plant is involved in M&A activity either as acquirer, or as target, or as a plant involved in horizontal or non-horizontal mergers. The M&A dummy is zero if the plant is a control. The results provide evidence that subgroups of merged plants also exhibit higher pre-merger productivity, except for the subgroup of non-horizontal M&A: the productivity differential is 72% for acquirers, 51% for targets, and 73% for plants involved in horizontal mergers, and they are all statistically significant at the 1%-level. Almost all coefficients of the interaction terms are not statistically significant at any level, i.e. there is no evidence for a change in the productivity difference between merging and non-merging plants over time. Only targets exhibit a statistically significant productivity decrease of approximately 14% in $t = 4$ compared to $t = 1$. The magnitude of coefficients of control variables and their statistical significance are mostly similar to the first regression.

The regression estimates seem to be robust: the values of R^2, which describe how much

[36]Similar to Schank, Schnabel, and Wagner (2010) I do not consider the logarithm of the proportional variables because there are several observations which exhibit a value of 0 which makes a transformation into logarithmized values impossible. Thus, a direct transformation of the value of the coefficient into percentage changes is not possible. However, the sign of the coefficient and its statistical significance provide sufficient evidence about the correlation to the dependent variable.

[37]Skilled employees are employees doing qualified jobs that require vocational training or the equivalent, training on the job or relevant professional experience, a university degree or higher education.

[38]This group includes working proprietors, directors, and managers.

of the variation of the dependent variable is explained by the variations of the independent variables, are sufficiently high: for example, the first regression specification exhibits a R^2-value of 0.568. The R^2-values of the other specifications are similar. I performed further robustness tests (but did not report the results): the F-test tests the joint hypothesis that all coefficients, except the intercept, are equal to zero; the variance-inflation-factor (VIF) controls for multicollinearity, and the Durbin-Watson-test controls for autocorrelation. The test results provide evidence for a proper model specification. This is also supported by the fact that the estimated coefficients have the expected signs from an economic point of view. Furthermore, I also eliminated intragroup correlation and corrected biased standard errors.[39]

How do the results from regression analysis compare to findings from descriptive statistics? Despite a few exceptions, the results are in line with each other: plants that merge exhibit a statistically significantly higher pre-merger productivity compared to non-merging plants. This points to a self-selection of more productive plants into merger activity, even if I control for other plant characteristics. However, regression does not find a statistically significant productivity differential for the subgroup of non-horizontal mergers, which may be due to the small number of observations. A further difference between descriptive statistics and regression analysis are the magnitudes of the productivity differential in $t = 1$: in regression analysis, they are clearly smaller compared to descriptive statistics, and this is obviously due to several control variables which are correlated to a plant's productivity. The results from descriptive statistics and regression analysis do not point to a statistically significant productivity change over time, except for targets: regression estimates a statistically significant productivity decrease in $t = 4$. The coefficients from control variables are mostly in line with findings from descriptive statistics.

3.4.4 Difference-in-differences propensity score matching

The objective of this paper is to analyze self-selection of plants into merger activity, and the causality between M&A and a plant's productivity. The regression analysis from above makes it possible to prove correlations between variables, but it is not able to unambiguously detect causality. Even if correlation is a necessary condition for causality, it is not a sufficient one (Backhaus, Erichson, Plinke, and Weiber, 2010). For this, I apply a different estimation approach in order to analyze causal effects.

[39] In a panel context, observations on the same plant may be correlated in different time periods, but observations on different plants are not correlated (Baum, Schaffer, and Stillman, 2003). This intragroup correlation can bias standard errors and provide false information about the statistical significance of coefficients. Therefore, I eliminated intragroup correlation by clustering observations of the same plants in order to yield adjusted standard errors.

The fundamental evaluation problem: The problem for the empirical modeling is this: even if a plant exhibits a higher productivity growth after M&A, there must not necessarily be a causality between M&A and productivity growth. The reason may also be that plants with certain characteristics, e.g. higher productivity, self-select in M&A activity, and would have experienced a higher productivity growth even in the absence of M&A. For this reason, it would be desirable to compare both outcomes for the same plant, but the latter scenario cannot be observed. Formally,[40] this is:

$$\Delta = Y_i(1) - Y_i(0). \tag{5}$$

Y represents the productivity outcome. Hence, $Y_i(1)$ is the post-merger productivity if plant i merged, and $Y_i(0)$ if the same plant had not merged. D is a treatment variable with $D = 1$ if the plant merged, and $D = 0$ if the plant had not merged. However, there is no data for $Y_i(0)$ because it is the missing counterfactual.

The observation of the individual treatment effect is not possible. Thus, the micro-econometric evaluation literature (e.g. Caliendo, 2006; Dehejia and Wahba, 2002; Heckman, Ichimura, and Todd, 1997) defines a (population) average treatment effect (ATT) which is

$$ATT = E[Y(1)|D = 1] - E[Y(0)|D = 1]. \tag{6}$$

Causal inference depends on the second term $E[Y(0)|D = 1]$ which cannot be observed, because it describes the expected productivity of the group of merged plants if they had not merged. Taking $E[Y(0)|D = 0]$, the expected productivity of control plants, as alternative is possible as long as plants randomly assign to the group of non-merging plants. However, in non-experimental data, it is most likely that there is some sort of selection, i.e. that components that determine the decision to merge, also determine the productivity outcome (Caliendo, 2006). Estimations based on a simple comparison of both groups would then be seriously biased.

For this, matching techniques as developed by Roy (1951), Rubin (1974), and Heckman, LaLonde, and Smith (1999), are able to construct a valid control group in order to eliminate the endogeneity bias.[41] With this approach every merged plant is matched with a "statistical twin" that had not merged, i.e. the matching partners are as similar as possible in relevant characteristics prior to the merger. Remaining differences in the productivity outcome are then caused by the merger.

The method requires that selection is only on observables, i.e. conditioning on the values

[40] The notations are similar to Caliendo (2006).

[41] See Caliendo (2006) for a comprehensive introduction to the method of propensity score matching.

of a set of observable characteristics X which are not affected by the merger decision, the productivity outcome of both groups would be the same in the absence of M&A (Lechner, 1999). This is also known as the "conditional independence assumption" (CIA). Moreover, the "common support condition" ensures that propensity scores of both groups overlap, and all treated have a counterpart in the control group. Hence, only individuals which are sufficiently similar to each other will be matched (Caliendo, 2006). Additionally, the Stable Unit-Treatment Value Assumption (SUTVA) states that the behavior of one individual has no impact on the behavior of another individual.

Since the matching partners are compared with respect to several observable characteristics X, a dimensionality problem arises. Rosenbaum and Rubin (1983) showed that the use of a propensity score $P(X)$ as a single index is also sufficient: it is a measure of the plant's probability to merge conditional on observed characteristics X. The ATT can now be estimated as

$$ATT = E[Y(1)|D = 1, P(X)] - E[Y(0)|D = 0, P(X)]. \tag{7}$$

Estimation of the propensity score: The propensity score is estimated with a probit or logit model.[42] I use a probit model which consists of a dummy variable for merger activity as a dependent variable (with value one if plants merged and zero otherwise), and a set of explaining variables which have to fulfill the CIA. The explanatory variables are expected to determine the plant's choice for M&A and the productivity outcome simultaneously, and they are all measured in $t = 1$. There are different opinions about the correct number of variables: Rubin and Thomas (1996) suggest that variables should only be excluded if they are not related to the outcome or if they have no relevance. In contrast, Heckman, Ichimura, Smith, and Todd (1998), Augurzky and Schmidt (2001), and Bryson, Dorsett, and Purdon (2002) yield better estimation results with a smaller set of variables. The latter argue that including insignificant variables leads to a less exact estimation of the propensity score. Here, I follow the arguments of Bryson, Dorsett, and Purdon (2002) and use less variables.[43] Nevertheless, the objective of the matching process is to balance the covariates and not to obtain an exact estimation of the propensity score (Caliendo, 2006).

I perform a probit regression including all treated observations, and also for all four sub-

[42] Both models usually yield similar results (Caliendo, 2006).

[43] I performed several probit models with different numbers of variables to address the arguments about the proper number of explaining variables. The results that were most robust are based on this model specification.

groups:

$$P(MA_{it=2} = 1) = F(labor\,productivity_{it=1},\ size\,dummies_{it=1},$$
$$legal\,form_{it=1},\ location_{it=1},\ industry\,dummies_{it=1},\ year\,dummies_{it=1}). \quad (8)$$

The choice of the explanatory variables is determined by the theoretical and empirical literature about acquisitions, e.g. from Girma and Görg (2007), Margolis (2006a), Harris and Robinson (2002), and Conyon, Girma, Thompson, and Wright (2002a). I include a variable for the logarithm of labor productivity in order to address the argument that only better performing plants are able to acquire, as well as the cherry-picking-argument which states that acquirers only buy the best performing targets. Moreover, empirical evidence shows that size is an important determinant for M&A (e.g. Girma and Görg, 2007). In addition, I include dummies for legal forms, for the location in East Germany, for different industries, and for different years.

Table 3.5: **Probit regression (dependent variable: M&A dummy)**

Variables	All (1)	Acquirers (2)	Targets (3)	Horizontal (4)	Non-horiz. (5)
log. Productivity in t=1	1.257*** (8.105)	1.753*** (6.837)	1.018*** (5.037)	1.468*** (6.344)	1.019*** (3.734)
Employees 20-49 (D)	-0.458 (-1.462)	-0.156 (-0.362)	-0.819* (-1.872)	-0.201 (-0.431)	-0.783* (-1.683)
Employees 50-99 (D)	-0.631* (-1.926)	-0.450 (-0.959)	-1.019** (-2.367)	-0.603 (-1.227)	-1.913** (-2.400)
Employees 100-299 (D)	0.029 (0.107)	0.741* (1.829)	-0.601* (-1.751)	0.660 (1.616)	-0.916** (-2.104)
Employees 300-499 (D)	0.409 (1.093)	1.252** (2.286)	-0.006 (-0.013)	1.271** (2.364)	-0.132 (-0.222)
Employees >=500 (D)	1.766*** (5.410)	2.585*** (5.130)	1.253*** (3.279)	2.413*** (4.841)	0.958* (1.891)
Legal form "Limited" (D)	-0.049 (-0.212)	0.072 (0.246)	-0.207 (-0.622)	-0.128 (-0.404)	-0.073 (-0.157)
Legal form "Limited by shares" (D)	1.442*** (4.144)	1.869*** (4.404)	0.978* (1.905)	1.712*** (3.962)	0.865 (1.163)
Location in East Germany (D)	0.757*** (3.910)	0.764*** (2.868)	0.871*** (3.347)	0.985*** (3.584)	0.869** (2.275)
Constant	-14.134*** (-8.057)	-20.159*** (-6.778)	-11.970*** (-5.348)	-16.852*** (-6.452)	-11.771*** (-3.726)
Observations	622	567	499	540	441
Pseudo-R^2	0.496	0.584	0.423	0.552	0.386

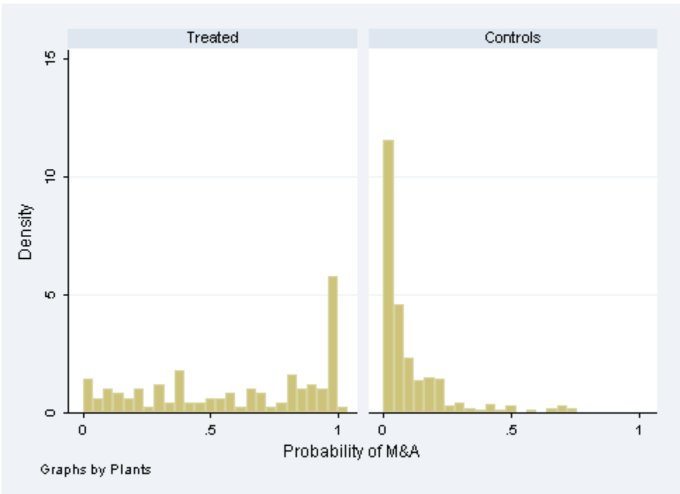
Notes: t-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1. (D) means variable is a dummy. The reference category for employees is "Employees 1-19", and for legal form it is "Partnership, individually-owned, public, and others". Regressions also include dummies for sectors and years. The reduced number of observations is due to missing data for several variables. Data source: IAB Establishment Panel, M&A DATABASE St. Gallen.

The results are shown in table 3.5. The first column presents estimations for the whole group of treated. The coefficient[44] of the productivity variable is positive, of high magnitude, and statistically significant at the 1%-level. This means, the higher a plant's productivity, the more likely it merges. The coefficients of the dummies for different size categories have to be

[44] The interpretation of coefficients in binary treatment models is more difficult than in linear regression models (see Backhaus, Erichson, Plinke, and Weiber, 2010). However, the magnitude and sign of the coefficient provide sufficient information about the impact of the dependent variable.

interpreted with respect to the reference category (number of employees <20): the negative coefficient of the dummy for employment size of 50-99 is statistically significant. Plants with employees of 500 and more have a statistically significantly higher probability to merge. If plants have legal form "Limited by shares", their probability to merge statistically significantly increases. Finally, if plants are located in East Germany, they are more likely to be involved in M&A. A McFadden's pseudo-R^2 of 0.496 is sufficiently high, and implies that around half of the variation of the dependent variable is explained by the regression.[45] The results of the regression specifications for subgroups are similar with some exceptions, in particular for the subgroup of non-horizontal mergers. Moreover, they confirm that plants that acquire others are better performers, and targets are cherries.

Figure 3.3: **Distribution of propensity scores for treated and control group**



Finally, these probit estimations generate a propensity score $P(X)$ for each observation, expressing the pre-merger probability of being involved in M&A activity in $t = 2$. Figure 3.3 presents a graphical illustration of the distribution of propensity scores for the groups of treated and controls:[46] the distribution is different for both groups, and matching partners with similar propensity scores are rare for certain ranges. The matching quality suffers if treated with high propensity scores are matched with controls having low propensity scores, and vice versa.[47] Thus, when pairing treated with control plants, it has to be taken into account that the ATT is only defined for the region of common support as mentioned above. However, this does also

[45] Values above 0.2 are considered as acceptable, and values of 0.4 and above can be considered as good (Backhaus, Erichson, Plinke, and Weiber, 2010).

[46] The distributions of propensity scores for subgroups of treated are not reported, but they look similar.

[47] Several treated observations exist with a propensity score close to one. This may be because plants with a certain combination of characteristics exhibit a strong merger activity. The same is true for several controls with a propensity score close to zero.

mean that causal inference is restricted to these observations.

Matching algorithm: It is not very likely to find a matching partner with exactly the same score because the propensity score is a continuous variable. For this, a neighborhood has to be defined, and each control has to be assigned with a specific weight. A general form of the treatment effect is

$$\Delta = \frac{1}{N_1} \sum_{i \in I_1} \left[Y_i^1 - \sum_{j \in I_0} W_{N_0}(i,j) Y_j^0 \right]. \tag{9}$$

I_1 and I_0 are the respective groups of treated and control plants. N_1 and N_0 are the number of plants in the treatment group I_1 and control group I_0. $W_{N_0}(i,j)$ is the weight of a control j from the control group which is assigned to a treated plant i. For each treated plant i the sum of the weights of all controls j is equal to one: $\sum_j W_{N_0}(i,j) = 1, \; \forall \; i$. For every treated plant i with propensity score P_i, a neighborhood $C(P_i)$ is defined, and neighbors of i are controls $j \in I_0$ with $P_i \in C(P_i)$. Y_i^1 is the outcome of a treated plant i, and Y_j^0 is the outcome of a control j.

There are several matching algorithms which differ with respect to the definition of the neighborhood, and the weights assigned to the controls. In larger samples, the results from different algorithms should be similar (Smith and Todd, 2005a). However, if samples are smaller, the choice of the algorithm is important. For this, I tested different algorithms[48] and achieved the most robust results - see below - with kernel matching. Whereas other algorithms only use one or few controls as matching partners for each single treated observation, the kernel algorithm uses all individuals j from the control group as neighbors for each single treated i. Thus, the neighborhood in kernel matching contains all observations in the control group I_0:

$$C(P_i) = \{I_0\} \tag{10}$$

The weights of controls j depend on their distance to the treated i, i.e. controls which are closer receive a higher weight than others:

$$W_{N_0}^{KM}(i,j) = \frac{G_{ij}}{\sum_{k \in I_0} G_{ik}}, \tag{11}$$

where $G_{ik} = G[(P_i - P_j)/a_{N_0}]$ is a kernel function[49] that downweighs observations j which are

[48] I also applied nearest neighbor matching with different numbers of neighbors, caliper and radius caliper matching with different maximum distances of propensity scores between treated and controls, and I modified all of them with respect to a replacement option for controls.

[49] For an introduction to kernel density estimation see Fahrmeir, Künstler, Pigeot, and Tutz (2009).

distant from the treated i. a_{N_0} is a bandwidth parameter which impacts the form of the kernel function. I use a kernel based on a Gaussian normal function and a bandwidth of 0.06.[50]

In smaller datasets like here, kernel matching has an important advantage: compared to other algorithms, it uses more information from a lager number of controls that flow into the parameter estimation. Hence, this may reduce the variance of the estimator. However, the algorithm may also lead to bad matches, because all controls are used as matching partners, even those which lie far away. For this reason, it is important to impose a common support condition (Caliendo, 2006). In addition, I also apply a trimming condition as suggested by Smith and Todd (2005a): if there are no controls for some intervals within the region of common support, which is the case according to figure 3.3, the respective treated observations will be excluded.[51]

After conditioning on observables, there may still be differences between the productivity outcomes of treated and control plants. They can be due to systematic differences in both groups because of selection into the treatment based on unmeasured characteristics (Smith and Todd, 2005a). For this reason, I combine propensity score matching with a difference-in-differences (DiD) estimator, i.e. I compare changes over time instead of levels. This eliminates unobserved time-invariant differences between both groups, relaxes the strong assumption of selection on observables, and improves the quality of the results significantly (Blundell and Costa Dias, 2000).

Under the consideration of the common support condition, the estimator can be implemented as

$$\Delta_{ATT}^{DiD} = \frac{1}{N_1} \sum_{i \in I_1 \cap S_P} \left[(Y_{it=4}^1 - Y_{it'}^0) - \sum_{j \in I_0 \cap S_P} W(i,j)(Y_{jt=4}^0 - Y_{jt'}^0) \right], \qquad (12)$$

with S_P denoting the region of common support, and t' being either $t = 1$ or $t = 3$, because I analyze growth rates between $t = 1$ and $t = 4$, and between $t = 3$ and $t = 4$.

Heckman, Ichimura, Smith, and Todd (1998) discussed kernel matching estimators.

[50]There are several types of kernel functions like Gaussian (normal), biweight, epanechnikov, uniform, and tricube kernel. DiNardo and Tobias (2001) stated that the choice of the kernel function is of minor interest. Silverman (1986) and Pagan and Ullah (1999) argued that the choice of the bandwidth parameter affects the results more strongly.

[51]If, for example, the common support is from 0.1 to 0.9, but there are no controls between 0.3 and 0.4, the trimming condition excludes the treated observations within this range. In other words: the region of common support only consists for those values of propensity scores which have a positive density within the distribution of both groups. If the density is exactly zero, the propensity scores will be excluded. To ensure that the densities are strictly positive, a further percentage of propensity score values - here 10% - with a very low density are also excluded (Caliendo, 2006).

Results: For the matching process, I use the STATA-module PSMATCH2 of Leuven and Sianesi (2003). The results of the difference-in-differences propensity score matching estimations are shown in table 3.6. It displays the average percentage productivity change for treated and control, the ATT which describes the average difference between both groups, the standard error of the ATT, as well as t-statistics. In order to test the statistical significance of the ATT, I apply bootstrapping with 150 replications to estimate standard errors again.[52] For this, the table presents the respective p-values, indicating whether the ATT is statistically significantly different from zero or not.

Table 3.6: **ATT for labor productivity changes**

Labor productivity changes between $t = 3$ and $t = 4$:						
Group of treated	Treated	Controls	ATT	S.E.	T-statistic	p-value
All	0.017	0.008	0.009	0.040	0.22	0.779
Acquirers	-0.009	0.014	-0.023	0.047	-0.49	0.502
Targets	-0.005	0.008	-0.013	0.037	-0.36	0.677
Horizontal	0.000	-0.001	0.001	0.044	0.03	0.973
Non-horizontal	0.001	-0.001	0.002	0.049	0.04	0.974

Labor productivity changes between $t = 1$ and $t = 4$:						
Group of treated	Treated	Controls	ATT	S.E.	T-statistic	p-value
All	0.072	-0.064	0.135	0.060	2.26	0.019
Acquirers	0.091	-0.081	0.171	0.093	1.84	0.047
Targets	0.030	-0.044	0.074	0.044	1.67	0.214
Horizontal	0.065	-0.065	0.131	0.106	1.23	0.195
Non-horizontal	0.008	-0.007	0.015	0.058	0.26	0.761

Notes: p-values are estimated for bootstrapped standard errors with 150 replications.

The upper part of the table presents the estimated results for percentage changes for the matched sample in the post-merger period. The first row displays the results for the treatment group including all treated observations ("All"). Average changes for treated and controls are both positive, and the ATT is 0.009, i.e. plants that merge exhibit a productivity growth that is 0.9% higher compared to control plants. However, this difference is not statistically significantly different from zero at any acceptable level. This implies that M&A obviously does not affect the merging plants' post-merger productivity growth. The estimated ATTs are negative for acquirers and targets, and positive for the subgroups of horizontal and non-horizontal M&A, but none of them is statistically significant.

[52] Abadie and Imbens (2008) argue that no formal justification has been provided to use bootstrapping in the context of matching. In contrast, many recent empirical studies applying matching also use bootstrapping (e.g. Girma, Görg, and Wagner, 2009).

The lower part of the table presents the results for percentage changes for the total period. Starting with the treatment group including all treated observations, the average productivity change is positive for treated, and negative for matched controls, leading to a positive ATT of 0.135, which is also statistically significant at the 5%-level. This result implies that M&A positively impacts the merging plant's productivity growth over the whole observation period. For subgroups, productivity changes are all positive, and the respective ATTs are also positive. For the subgroup of acquirers the ATT is statistically significant at the 5%-level, but not for the other subgroups. This implies that the effect measured in the treatment group that included all plants results from the subgroup of acquirers.

Robustness tests: The matching literature has developed several robustness tests in order to assess the quality of results. First, the objective of matching is to balance the covariates, i.e. matched observations should be similar. This means that after matching there should be no significant differences in the mean values of all explanatory variables between the treated and the control group. This can be analyzed by the standardized bias by Rosenbaum and Rubin (1985) which is given by

$$SB = 100\frac{(\bar{X}_1 - \bar{X}_0)}{\sqrt{0.5[V_1(X) + V_0(X)]}}. \tag{13}$$

After matching, the standardized bias is defined as

$$SB^M = 100\frac{(\bar{X}_1^M - \bar{X}_0^M)}{\sqrt{0.5[V_1^M(X) + V_0^M(X)]}}. \tag{14}$$

$\bar{X}_1$ and $\bar{X}_0$ are the mean values for a specific variable of the treatment and control group and V_1 and V_0 are the respective variances. $\bar{X}_1^M$, $\bar{X}_0^M$, V_1^M and V_0^M are the corresponding values after matching.

The upper part of table 3.7 reports the mean standardized bias (MSB) for all variables before and after matching for the kernel algorithm. With respect to the matching procedure that includes all variables the MSB is approximately 8. There is no clear benchmark for the MSB which indicates if a sample is balanced or not. Rosenbaum and Rubin (1985) suggested that values for the standardized bias should be smaller than 20. However, in most empirical studies a bias reduction below 3 or 5% is seen as sufficient (Caliendo, 2006). Hence, one can conclude that the balancing was acceptable, but not ideal, which has to be kept in mind for the interpretation of the results. The values of the MSB for the subgroups are similar with values of 9 for the groups of acquirers and horizontal merging plants, and 10 for the subgroup of targets. The group of non-horizontal M&A exhibits a poor value of 17.

Table 3.7: **Robustness tests**

Mean standardized bias (MSB)			
Group of treated	Mean/Std. Err.	Before matching	After matching
All	Mean	32.168	7.965
	Std. Dev.	29.426	5.066
Acquirers	Mean	34.000	9.418
	Std. Dev.	33.431	6.020
Targets	Mean	34.544	10.846
	Std. Dev.	27.250	7.428
Horizontal	Mean	33.122	9.118
	Std. Dev.	30.093	5.645
Non-horizontal	Mean	33.191	17.361
	Std. Dev.	25.641	12.948

Common Support				
Group of treated	Sample	Off support	On support	Total
All	Untreated	0	499	499
	Treated	46	77	123
Acquirers	Untreated	0	492	492
	Treated	35	40	75
Targets	Untreated	0	451	451
	Treated	22	26	48
Horizontal	Untreated	0	473	473
	Treated	34	33	67
Non-horizontal	Untreated	0	421	421
	Treated	10	10	20

Pseudo-R^2 and log likelihood test				
Group of treated	Sample	Pseudo-R^2	LR chi2	p>chi2
All	Unmatched	0.496	307.02	0.000
	Matched	0.040	8.61	0.979
Acquirers	Unmatched	0.584	258.61	0.000
	Matched	0.054	5.90	0.994
Targets	Unmatched	0.423	133.79	0.000
	Matched	0.056	4.05	0.999
Horizontal	Unmatched	0.552	223.41	0.000
	Matched	0.037	3.34	1.000
Non-horizontal	Unmatched	0.386	62.91	0.000
	Matched	0.272	7.10	0.851

In addition, I also reported the results for the region of common support. The middle part of table 3.7 shows that a large fraction of treated observations - almost 40% - lies outside the region of common support, i.e. these observations are not used for the estimation. In the subgroups, the fraction of treated lying outside the region of common support is even higher, up to approximately 50%. The exclusion of a high fraction of treated from the analysis due to the imposition of a common support condition creates difficulties for the interpretation of the results (Lechner, 2008): first, useful information is ignored because treatment effects could still be estimated outside the region of common support. Moreover, if treatment effects are heterogeneous inside and outside the common support, estimated treatment effects may no longer correspond to the original parameter of interest. Consequently, the expressive power and generality of these matching results is reduced, and inference is only valid for the region of common support. This means that estimations only provide information about what would have happened to a merged plant's productivity growth if the same plant had not merged.

There are further indicators for the quality of the results: pseudo-R^2, likelihood ratio-test (LR), and respective p-values are calculated for the matched sample. If matching was successful, there is no difference in the covariates between subsamples of treated and controls, and a probit estimation of only matched firms has no explanatory power (Sianesi, 2004). This would be reflected by a low pseudo-R^2 value. Moreover, a low value for the LR-test and a p-value close to value one also indicate that the matched sample has no explanatory power. According to the numbers displayed in the lower part of table 3.7, the results are satisfying, except for the subgroup of non-horizontal M&A.

The conclusion from matching is that the results are in line with findings from descriptive statistics and regression analysis with respect to post-merger productivity effects: there is no evidence for a merger-induced change between $t = 3$ and $t = 4$. With respect to changes between $t = 1$ and $t = 4$, there are two differences to the results from descriptive statistics and regression: first, matching estimated a positive causal effect on plants' productivity between $t = 1$ and $t = 4$ for the treatment group including all treated. After differentiating between subgroups, the effect obviously originates from acquirers. This effect was neither estimated in descriptive statistics nor in regression analysis. Second, regression estimated a negative change on targets' productivity in $t = 4$, and this estimation is not confirmed by matching. The differences in the results can be explained as follows: the estimated effects of the matching method only refer to those treated which find comparable controls, i.e. which lie within the region of common support as explained above. In contrast, and due to its functional form, a regression produces estimates even if there are no untreated comparables to treated, i.e. the functional form fills in the missing data and extrapolates for those treated without comparable controls (Caliendo, 2006).

3.5 Conclusion - what do we learn?

This study analyzes the causal effect of M&A on labor productivity of German plants that merged between 1996 and 2005. I differentiate between acquirers, targets, horizontally, and non-horizontally merging plants. The findings provide strong evidence for a self-selection of plants into merger activity, as already found in earlier studies (e.g. McGuckin and Ngyuen, 1995; Salis, 2008). Productivity is significantly higher for merged plants prior to the merger in all subgroups.[53] With respect to acquirers, this implies that only better performing plants are able to acquire other plants, e.g. due to fixed or other costs accompanied by a merger. Based on these estimations, it is less likely that poor performers see the acquisition of other firms as a way to improve their own performance. The findings also show that acquirers buy over-performing targets (cherries). This suggests that the likely motivation behind M&A are potential gains from synergies, but there is no support for the matching theory of ownership from Lichtenberg and Siegel (1992a), and the theory of inefficient management from Scharfstein (1988) and Jensen (1988) which both imply that targets are poor performers prior to the merger.

What about causal effects? First, with respect to post-merger changes between $t = 3$ and $t = 4$, I do not find any effects for all subgroups. This is in line with existing studies, e.g. from Bellak, Pfaffermayr, and Wild (2006), or Salis (2008). Hence, it can be concluded that there is no statistically significant post-merger productivity change in German plants after domestic M&A. Second, matching estimates a weak positive effect for acquirers with respect to the total period between $t = 1$ and $t = 4$. However, the common support condition imposed in matching excludes a large fraction of observations that lie outside the region of common support. Hence, the generality of the results is limited, and inference is restricted to observations within this region. For the interpretation of the matching results, this means that the productivity of acquirers increases after M&A in comparison to a non-merging plant that is identical (or at least similar) in relevant pre-merger characteristics, and the productivity increase is only caused by the merger. In other words: acquirers face a higher productivity growth as if they would have had not merged. The positive productivity effects on acquirers from mergers suggest that firms acquire others in order to increase efficiency and to gain synergies, e.g. in the form of a reorganization of business structures, or from adapting efficiency improving technical or organizational systems from the target. These findings support the arguments from Dunning

[53] The results for the subgroup of non-horizontal M&A are not fully in line with the results for other subgroups, e.g. the results from regression do not confirm a higher pre-merger productivity. However, the results for non-horizontally merging plants are not robust due to a small number of observations.

(1998) who states that resources flow from acquired to parent companies.[54] The fact that I only find effects in acquiring plants over the observation period $t = 1$ to $t = 4$, but not between $t = 3$ and $t = 4$ suggests that productivity changes take place immediately around the merger. There are no effects for targets which imply that they do not obviously gain from the acquirers' advantages like technology, organization, etc. This does also not support the argument from Mueller and Sirower (2003) who state that synergy gains are equally distributed between acquirers and targets.

The estimations for the subgroup of acquirers are mostly in line with newer empirical findings that showed a tendency towards productivity improvements (e.g. Girma, Thompson, and Wright, 2006), or that found no effects (e.g. Bellak, Pfaffermayr, and Wild, 2006). The findings contradict earlier studies which estimated negative effects (e.g. Ravenscraft and Scherer, 1987), and different explanations are possible, e.g. better data availability or improved econometric methods controlling for selection bias. Alternatively, my estimation results may also be a cautious indication that fewer and fewer acquirers can afford to merge for non-profit maximizing reasons in an economic environment that becomes more and more competitive through internalization. Finally, the results legitimate the criticism of "comparing the incomparables" mentioned at the beginning of my paper: only a differentiation into subgroups shows that the causal effect that was estimated from matching for the treatment group including all treated is caused by the subgroup of acquirers.

[54] However, Dunning's (1998) arguments refer to foreign acquisitions.

A Remarks on the dataset

The empirical analysis of this paper is based on a combined dataset of the IAB Establishment Panel and the M&A DATABASE from St. Gallen, by TNS Infratest Sozialforschung GmbH München.[55] In an earlier pilot study, Bellmann and Kirchhof (2006) showed that the IAB Establishment Panel is capable of an analysis of the effects of mergers. The focus of their study was on the effects on employment, and the data about mergers was from Thomson ONE Banker. Since Thomson ONE Banker only includes firms of a larger size, the M&A DATABASE also includes small and medium-size firms, and thus, is more comprehensive. In addition, the M&A DATABASE also lists information about the seller of each deal[56] (Thomson ONE Banker only reports about buyer and target firms), location, number of employees, sales, etc.

The treatment group: The observation period for plants that merged is from January 1996 to December 2005. The creation of the treatment group was carried out in several steps by TNS Infratest Sozialforschung. First, in order to combine both datasets, companies in the M&A DATABASE were compared to plants in the IAB Establishment Panel and classified according to the degree of similarity with respect to name, location, and sector. Next, merged plants were only kept if they were surveyed at least once before and once after the merger. This led to some complications due to the set up of the survey: information about employees refers to June 30th of the respective year, whereas information about sales, investments, etc. refer to the previous year. To take these circumstances into account, the following definition was chosen:

- If M&A was between January 1st and June 30th of year T, the survey in year T was considered to be conducted after M&A, even if some information refers to a point of time before M&A.
- If M&A was between July 1st and December 31th of year T, the survey in year T was considered to be conducted before M&A, even if some information refers to a point of time after M&A.

[55] Two methodology reports exist about the creation of the treatment group and the control group from TNS Infratest: Beschäftigungseffekte von Fusionen und Übernahmen - Methodenbericht Untersuchungsgruppe (März 2007); Beschäftigungseffekte von Fusionen und Übernahmen - Methodenbericht Untersuchungsgruppe (December 2007).

[56] Nevertheless, this study does not analyze the effects on sellers because the number of observations is too small.

This restriction reduced the number of plants which were found in both datasets to 7,801. According to the degree of similarity of plants in both datasets, observations are distributed across four different categories:

- Quality class 1: name, location, and sector match exactly (1,426);
- Quality class 2: name, location, and superior sector match exactly (146);
- Quality class 3: name and sector match exactly; multi-plant firm (5,961);
- Quality class 4: name and location match exactly (268).

These 7,801 merger cases consist of 958 different plants in the IAB Establishment Panel. This is because one plant may be involved in several mergers within the observation period.

The control group: Next, a group of control plants that had not merged between 1980 and 2005 had to be found. These controls must be as similar as possible to plants in the treatment group. Each of the 7,801 treatment observations exhibits an individual combination of sector, size, legal form, and location in West or East Germany. Therefore, TNS Infratest Sozialforschung defined 2,143 categories which differ with respect to these characteristics, and each of the 7,801 treatment observations was assigned to one of these 2,143 categories. Now, the objective was to find controls for each category. This is, within a category, treated and controls are homogeneous with respect to the characteristics. An example: there are three treated plants which all belong to the agriculture sector, have less than 10 employees, have "GmbH" as legal form, and are located in West Germany. The combination of these specific characteristics constitutes one of the 2,143 categories. After that, three control plants should also be identified which exhibit the same characteristics of this specific category.

30,110 plants from the IAB Establishment Panel were identified as statistical twins to treated plants (with respect to sector, size, legal form, and location in West or East Germany), and they may potentially act as a control. Within these 30,110 observations, several plants appear more often if they were surveyed for the IAB Establishment Panel for several years. The challenge was to identify those "true" controls within the 30,110 potential controls, that is, plants for which we can be sure that they have not been involved in any M&A activity. Hence, plants which already appeared in the treatment group were excluded, as well as plants which were similar to plants from the M&A DATABASE, but which were not in the treatment group, because they merged outside the reference period between 1996 and 2005. This step excludes 1,204 from the 30,110 observations. Then, plants which were not surveyed at least twice between 1993, the starting year of the IAB Establishment Panel and 2006,[57] the end of the observation period, were dropped.

The rest of the remaining 27,676 potential controls had to be checked manually. As a first check, plants were eliminated if their name appeared in the M&A DATABASE. In addition, the dataset "Markus" from Bureau van Dijk also provided information about M&A activities of plants. Finally, for most plants websites were used as a source of information about merger activity.

As stated above, the number of treated should equal the number of controls within each of the 2,143 categories. Hence, for each category, potential controls were checked for whether they were "true" controls. This was repeated until the number of true controls equaled the number of treated, and the remaining potential controls for the respective category were no longer considered. However, for several categories no controls could have been found, because potential controls have all merged.[58] Figure A1 presents a graphical illustration of this process.

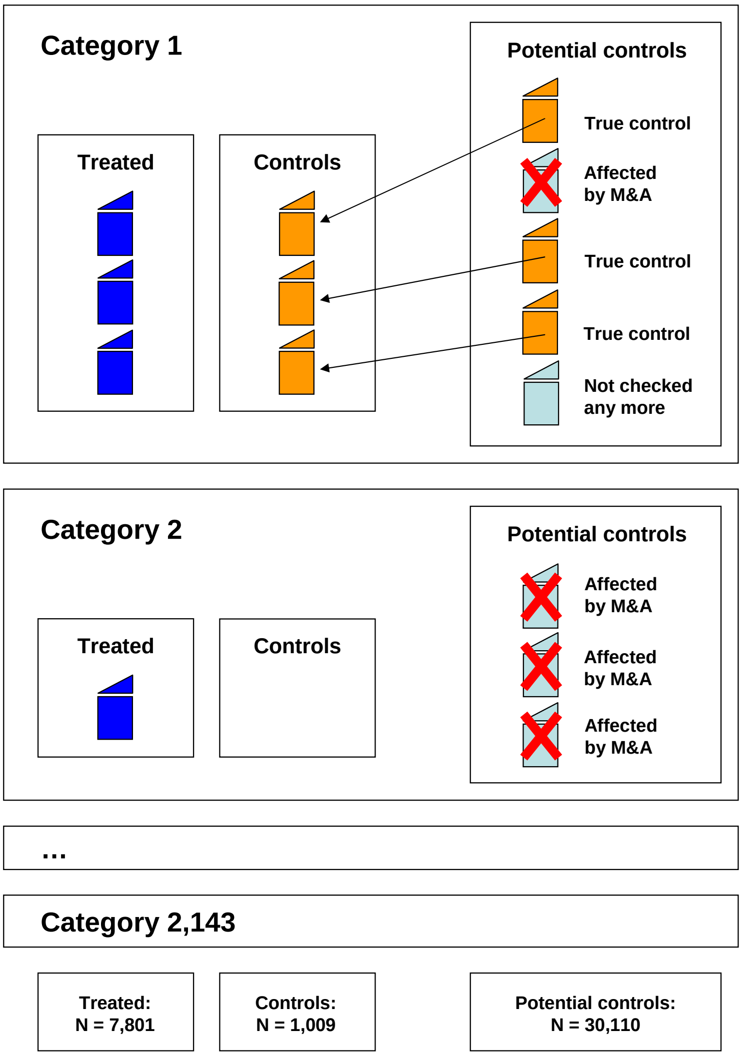
In total, 12,755 plants were checked in 400 hours of research by TNS Infratest. As a result, 1,009 controls from 291 different plants appear in the control group. This is because a plant can act as a control over several years and for different categories as well. The structure of the control group is similar to the structure of the treated group with respect to sector, size, legal form, and location, but controls were not involved in any merger activity during the reference period.

A note on sector classification: The 2-digit sector classification of the IAB Establishment Panel follows the NACE code (Nomenclature Générale des Activités Économiques). The NACE code changed in 1999 and in 2003, leading to different classifications of plants over time. Whereas the classification change in 2003 is not a problem, the break in 1999 is more severe. Due to this, I transfered the sector classification of the year 2000 to the years before in order to achieve a consistent sector classification of plants. However, this leads to a drawback if plants changed sectors due to a merger: they may not be classified correctly. I checked this aspect manually and could not identify incorrect classifications.

[57] The observation period for treated plants ends in December 2005. To gain information about controls for the year 2005, the 2006 survey is relevant, because several questions refer to the year before, e.g. sales.

[58] For example, almost no controls were found in the financial sector.

Figure A1: **Graphical illustration of the creation of the control group by TNS Infratest**



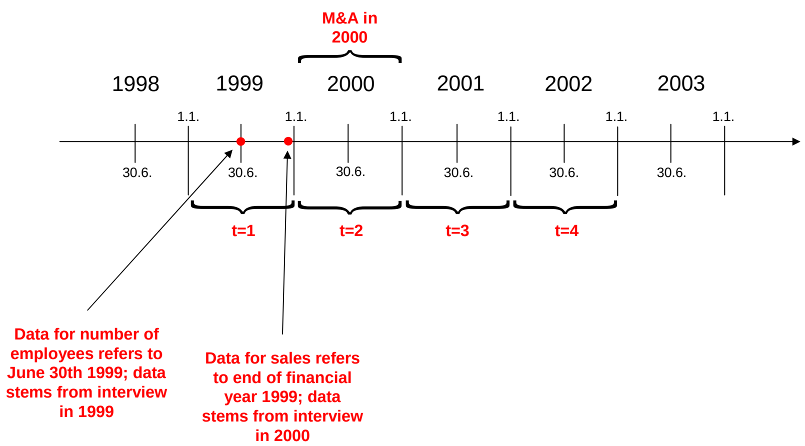
Note: the large boxes represent different categories. In each of the 2,134 categories, plants are homogeneous with respect to sector, size classification, location in West or East Germany and legal form. The number of controls should equal the number of treated in each category. Therefore, 30,110 potential controls from the IAB Establishment Panel, which represent statistical twins to the treated, were assigned to these different categories. Then, each potential control was checked for whether it had merged since 1980. If this was true, the respective plant was identified as a "true" control. If, however, a potential control had merged since 1980, it was discarded. If the required number of controls was found, remaining controls were not considered anymore (category 1). It could also be that no controls were found for a certain category (category 2). Finally, 1,009 controls were identified.

B The observation period

The four-year observation period: The calculation of a plant's productivity from the IAB Establishment Panel creates some difficulties which will be briefly explained here. There are some questions in the IAB Establishment Panel wave, e.g. about numbers of employees, which refer to June 30th of the same year the survey is carried out. For example, the survey of the year 1999 delivers information about the number of employees the plants employed at June 30th 1999. Other questions, e.g. about sales or investments, refer to the year before. These different reference dates create some challenges which are relevant for the empirical setting.

Consider an example shown in figure B1: M&A takes place between January 1st and December 31th 2000 ($t = 2$). The analysis starts with the observation of the data prior to the merger, i.e. in year 1999 ($t = 1$). The 1999 wave asks for the number of employees on June 30th 1999. However, to get information about sales in 1999, the survey of 2000 provides the respective information. The same applies to data for the periods $t = 2$ to $t = 4$.

Figure B1: **Illustration of the observation period**



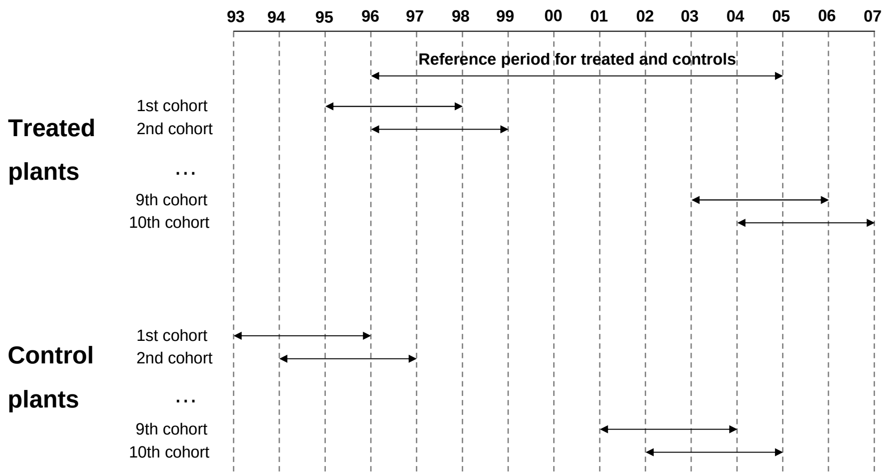
For the calculation of post-merger growth-rates, data from years $t = 3$ and $t = 4$ is needed: imagine, M&A occurs in September 2000. Then, labor productivity (defined as sales per employee) for $t = 2$ is calculated from data before and after the merger, in particular, from employment data prior to the merger (June 30th) and - depending on the end of the financial year - from sales data after the merger (e.g. December 31th). This may lead to biased estimates for changes between $t = 2$ and $t = 3$. Thus, calculations of post-merger growth rates seem to be more reliable for the period $t = 3$ to $t = 4$, because numbers of both years refer to post-merger time. Alternatively, I could change the observation window for M&A

activity to July 1st 1999 and June 30th 2000. But this may lead to a biased calculation of labor productivity in $t = 1$, because employment data refers to a point of time prior to M&A, but sales may refer to a point of time after M&A.

In addition to changes between $t = 3$ and $t = 4$, I also analyze changes between $t = 1$ and $t = 4$. This is because M&A may quite quickly (within several weeks or months) affect labor productivity rather than one or two years later. Imagine M&A occurs in January 2000 ($t = 2$). For an estimation of post-merger changes in productivity, data for year 2001 ($t = 3$) refers to June 30th in 2001 (employment), i.e. almost one and a half years later, and to December 31th in 2001 (sales), i.e. almost two years later. The same applies to data for year 2002 ($t = 4$), that is, data refers to two and a half and three years after the merger. Therefore, the main effects may not be measured any more. An alternative calculation of changes between years $t = 2$ and $t = 4$ is not a good idea: again, data about employment refers to June 30th and sales (probably mostly) to December 31th. If the merger occurred between July 1st and December 31th of year $t = 2$, I would consider pre-merger employment data and post-merger sales data.

The rolling observation window: For my empirical investigation, I use a "rolling observation window" which is illustrated by figure B2. For treated, ten four-year windows exist, the first between 1995 and 1998, and the last between 2004 and 2007. Correspondingly, for controls the first four-year window is from 1993 to 1996, and the last is from 2002 to 2005.

Figure B2: **Graphical illustration of cohorts**



Note: for the first cohort in the treatment group, year $t = 1$ corresponds to 1995, $t = 2$ to 1996, $t = 3$ to 1997, and $t = 4$ to 1998. This applies to all cohorts in both groups analogously.

Chapter 4

Anybody afraid of M&A?
Effects on German plants' employment and skill-intensity

4.1 Introduction

News about mergers and acquisitions (M&A) is often accompanied by announcements about job reductions. For example, unions feared job losses due to a hostile takeover of the German construction company Hochtief by the Spanish construction company ACS in 2011,[1] the internet company AOL signalized a reduction of employment due to a merger with the online magazine Huffington Post in 2011,[2] the media reported about larger layoffs in the German banks Commerzbank and Dresdner Bank due to their merger in 2009,[3] and the same happened when the German bank HypoVereinsbank and the Italian bank Unicredit merged in 2005.[4] These reports create a picture of M&A as employment-reducing events that impact public optinion and political debates.

Despite these examples for job reductions after M&A, the research results do not confirm a negative causality between M&A and employment changes in the merging firms.[5] Instead, effects are not clear. One explanation for these ambiguous results may be that no typical merger exists (Tichy, 2001) because mergers are different with respect to several dimensions. Studies exist - see below - which analyzed different types of mergers, i.e. horizontal (firms which compete in the same market combine), vertical (firms combine with their supplier), or conglomerate mergers (firms of unrelated lines of businesses combine) (Carlton and Perloff, 2005). Moreover, some studies differ with respect to the type of the firm they considered, i.e. between acquirers ("acquirers" or "buyer firms" are firms which acquire other firms or plants[6]), targets ("targets", "objects", or "acquired" are firms which are acquired by firms), and sellers ("sellers" are firms which sell parts of the overall entity). In addition, the results may differ because some studies analyzed cross-border M&A while others analyzed domestic M&A, or because the empirical investigation was based on different observation periods, or due to a differentiation between firm- and plant-level. Hence, I suppose that we may be comparing the incomparable, and thus, should no longer wonder why the results differ. Another explanation of the ambiguous results may be that studies applied different estimation strategies. For example, measurement errors may emerge because earlier studies simply compared samples of merged

[1] N24, June 17th 2011: http://www.n24.de/news/newsitem_6983386.html [June 19th 2011].

[2] Welt Online, March 4th 2011: http://www.welt.de/print/die_welt/kultur/article12705115/Kompakt.html [June 6th 2011].

[3] Süddeutsche Zeitung, September 1st 2008: http://www.sueddeutsche.de/geld/stellenabbau-nach-bankenfusion-bittere-briefe-im-advent-1.691076 [June 18th 2011].

[4] WirtschaftsWoche, July 29th 2005 http://www.wiwo.de/unternehmen-maerkte/hypovereinsbank-kommt-nicht-zur-ruhe-301728/ [June 18th 2011].

[5] I use the terms "firm" and "plant" interchangeably as long as no greater precision is needed.

[6] I will use the terms "acquirer" or "buyer" interchangeably in this paper.

and control firms, not controlling for a potential self-selection of firms with certain pre-merger characteristics into M&A activity. Changes in performance parameters were then incorrectly attributed to the impact of the merger (Girma, Thompson, and Wright, 2006). For this, recent studies took advantage of newer and more advanced econometric methods which improve research about questions of causality and self-selection.

Hence, from existing studies it is difficult to draw any conclusions about causalities for the reasons mentioned above. The objective of my paper is to overcome some of these problems by distinguishing between buyer and target plants,[7] and between horizontal and non-horizontal (including both vertical and conglomerate) mergers.[8] Moreover, I use a matching approach, a newer estimation method to control for self-selection and identify the average effect of M&A on employment. I consider all employees, i.e. independent of their liability to social security or not. In addition, I also analyze changes in the skill-intensity of the merging plant's workforce, defined as the percentage of skilled employees within the workforce.

To preview my results, I found that plants that merge and plants that do not merge differ in their pre-merger characteristics. For example, there is a self-selection of larger plants into M&A activity, and this is true for all subgroups. However, I do not find causal effects of M&A on employment, but there is evidence for an effect of M&A on targets' skill-intensity, creating a U-shaped skill-intensity development path over time.

This paper is integrated in the existing literature about employment and skill-intensity performance of merging firms, taking firm heterogeneity into account. With respect to M&A and employment effects, there are studies that estimated a positive impact (e.g. McGuckin and Nguyen, 2001). In contrast, other studies found negative effects (e.g. Margolis, 2006b), or no effects (e.g. Bellak, Pfaffermayr, and Wild, 2006). The number of studies on the impact of M&A on firms' skill-intensity is scarce. The results are also ambiguous, i.e. studies either found a decrease in high-skilled workers (e.g. Lipsey and Sjöholm, 2003), or a reduction of low-skilled workers (e.g. Grima, 2004).

I extend the existing literature in a number of ways. First, only very few studies - see below - that discussed M&A and employment also analyzed possible changes of the plant's educational mix. Second I use a dataset that has not been used by others as far as I know. The dataset, which is a combination of the IAB Establishment Panel and the M&A DATABASE St. Gallen

[7]The number of seller plants in the dataset is too small for a separate analysis.

[8]Horizontal mergers involve firms that both operate in the same market, while firms that merge vertically or conglomerately operate in different market. For this, these mergers are also considered as non-horizontal (Church, 2004). In my study, a separation between vertical and conglomerate mergers is not possible because of too few observations.

contains plants that merged between 1996 and 2005, and also a control group of plants that have not merged since 1980. It differs from other datasets used for similar research questions because it allows distinguishing between the types of merging firms (acquirers and targets) and the types of mergers (horizontal and non-horizontal). Furthermore, the data includes not only large firms and major M&A, but plants and M&A of practically all sizes, and it also includes plants from different industries. In addition, the dataset focuses on domestic M&A, while most recent studies only analyze foreign acquisitions, even if domestic M&A account for a large fraction.[9] Third, I apply a differentiated estimation strategy. As usual, I start with descriptive statistics. Then, in order to control for other variables than M&A that influence a plant's employment and skill-intensity, I perform a regression analysis. However, neither descriptive statistics nor regression analysis are able to reliably detect causality and deal with the problem of self-selection. For this, I apply a propensity score matching approach developed by Rosenbaum and Rubin (1983, 1985). A matching method compares a group of merging plants with a group of plants that had not merged, but which are as similar as possible in their pre-merger characteristics. Thus, the matching method allows the construction of the counterfactual situation in which the merger did not occur. To improve the robustness of the results, I combine matching with a difference-in-differences estimator as suggested by Blundell and Costa Dias (2000) and Smith and Todd (2005a).

The paper is structured as follows: section 4.2 provides a theoretical background and presents the results from related literature. Section 4.3 describes the data, and section 4.4 performs different estimation strategies: for each strategy, I present the results for employment effects first, and then, describe the results from the empirical investigation of skill-intensity effects. Section 4.5 summarizes the results from this study.

4.2 Theoretical background and related literature

From pre-merger characteristics of merging firms, and from estimated effects on employment and skill-intensity some predictions about the underlying motives for M&A can be inferred.[10] One broad strand of literature focuses on managers that merge in order to maximize profits. In particular, mergers are often motivated by increasing returns to scale. The combination of several firms into one firm can lead to efficiency gains through a reduction in the number of

[9] In Germany, approximately 50% of all mergers are domestic (Spanninger, 2011a).

[10] See Jansen (2008), Margolis (2006b), Scherer (2002), Tichy (2001), and others for a discussion of reasons for mergers.

employees required to maintain the level of production. If efficiency gains are achieved through a reduction of fixed costs in central administration it seems more likely that employees working in central administration are disproportionately laid off. These layoffs can occur both in the acquiring and acquired firms (Margolis, 2006b). If the workers laid off in central administration are skilled, one might expect a decrease of the plants' skill-intensity after M&A.

Another prominent explanation for the existence of mergers is the inefficient management hypothesis (Manne, 1965). It states that firms acquire poorly performing firms with managers that do not maximize shareholder wealth. After the takeover managers will be replaced by better managers that maximize profits. These are able to realize efficiency gains which most likely include cost economies and labor reductions.

However, in a neoclassical framework employment changes due to M&A must not necessarily be negative. The effects will also depend on the complementarity of the merged firms and the post-merger market position (Conyon, Girma, Thompson, and Wright, 2002a). Moreover, there may be no employment effects if the takeover is a capital investment. Effects may even be positive if the takeover brings new capital to the target and improves its financial possibilities, for example because investments and innovations can now be realized which could not have been financed before (Bellmann and Kirchhof, 2006).

Synergies from mergers can also come from reorganization processes in the respective firms. Reorganization processes include measures like incentive wages, job rotation, restructuring of departments, etc. which can be implemented in a firm, and more likely in the target firm. These organizational changes can enhance productivity (e.g. Kölling and Schank (2002) or Bauer (2003) provided empirical evidence) and create new fields of activities and new qualification profiles (Beckmann, 2000) which may have a positive impact on employment. In contrast, employment should fall if organizational change leads to a streamlining of production processes. Thus, there is no clear prediction for employment changes due to a reorganization process (Bellmann and Kohaut, 1999). With respect to the impact of organizational changes on the qualification structure, Bellmann and Pahnke (2006) argued that different measures of reorganizational changes (e.g. shifts of competences, group work, close contact to customers) require skills like social and communicative competences, the ability to judge, taking the initiative, etc. Assuming that skilled rather than unskilled employees are endowed with these skills, organizational changes should lead to a reduction of the number of employees with lower qualification.[11]

[11]Bellmann and Pahnke (2006) confirmed these considerations in their empirical study.

Shleifer and Summers (1988) argued that a new management is less committed to employees, and thus, renegotiate about implicit aspects of employment contracts and conditions.[12] These renegotiations may be seen as a "breach of trust" by employees: a renegotiation discourages employees from making ex ante commitments to the firm, and they will not invest in firm-specific human capital. Hence, if the new management is able to renege on implicit labor contracts, employment is likely to be reduced.

Building on Shleifer and Summers (1988), Gugler and Yurtoglu (2004) discussed that firms merge in order to restore an optimal employment level in rigid labor markets. It seems plausible that the speed of labor adjustment with which firms respond to shocks is lower the higher the costs of labor adjustment. If labor adjustment costs are high, hiring employees is a somewhat irreversible decision, making it likely that some firms in countries with rigid labor markets carry excess labor. If a merger brings new management which is less committed to upholding past contracts with stakeholders, the merger is an effective way to achieve a desired restructuring and to reduce the excess labor. Depending on the level of rigidity in different labor markets, employment effects from M&A should differ: in countries with rigid labor markets, like Germany, labor demand can be expected to decrease more after M&A compared to countries with relatively flexible labor markets, for example the US.[13]

In contrast to the profit maximizing motives, there also exist other reasons for mergers that focus on managers' opportunistic behavior. For example, managers pursue the firm's growth and want to widen its scope and size, but do not maximize profits. For this, managers who follow this empire-building strategy may have a preference for larger targets. Other reasons that do not focus on profit maximizing are hubris, i.e. managers overestimate their abilities to improve the target's performance (Roll, 1986), or they merge in order to entrench themselves and make it costly to shareholders to replace them (Shleifer and Summers, 1988). The impact of M&A on the firms' employment and skill-intensity can hardly be predicted if mergers occur for these reasons, but employment losses are expected to be less likely compared to mergers motivated by profit maximizing.

Employment effects differ depending on the type of mergers. Employment losses are more likely in horizontal mergers than in non-horizontal mergers if the respective industry exhibits increasing returns to scale as argued above. Moreover, if there is a declining industry, that is, firms face a reduction in output due to a declining demand and bear excess capacity, a horizontal

[12]Shleifer and Summers (1988) place their focus on hostile M&A.

[13]Gugler and Yurtoglu (2004) provided empirical support for their hypothesis. Similarly, Abraham and Houseman (1993, 1995) found out that employment levels in the manufacturing sector adjust faster in the US than in Germany and Japan.

merger allows these firms to retire older and surplus capacity, leading to negative employment effects (Dutz, 1989). Vertical mergers are a way to reduce transaction costs (Williamson, 1975). As a consequence, employees in the sales function in the upstream firm may be laid off, and the downstream firm may react with a reduction of employees in the procurement function. However, these cost savings may generate an output expansion which is sufficiently high to offset the job losses associated with transaction cost reductions, but this scenario seems unlikely. The employment effects of conglomerate and unrelated mergers are not clear: no negative employment effects should be expected if managers undertake unrelated mergers in order to diversify firm earnings. If, however, the unrelated merger is a disciplinary merger in order to use the market for corporate control to divert assets to a better management - see the arguments for the inefficient management hypothesis above -, cost savings and employment losses are possible (Conyon, Girma, Thompson, and Wright, 2002a).

Effects on employment can vary between domestic and foreign M&A (Lehto and Böckerman, 2008). Employment losses may be larger after domestic M&A for the following reason: information about target firms become more incomplete with an increase in the geographic distance between both firms (Lehto, 2006). Therefore, acquirers are located close to targets in order to have better knowledge of the local markets. Closer proximity enables firms to undertake profound rationalization and radical structural reforms after a merger, causing negative employment effects. In contrast, a foreign acquirer may be less committed to fulfilling implicit contracts that preserve employment, leading to larger employment losses after cross-border M&A compared to domestic M&A.

Finally, some predictions about the firms' pre-merger characteristics are possible. There are several reasons why larger firms may self-select in M&A activity: acquiring firms have to finance the merger itself, and they have to bear fixed costs for the integration of the acquired firm. For this, it seems plausible that only better performing firms are able to incur these fixed costs. If there is a positive relation between a firm's size and its productivity (see studies from Bernard, Jensen, Redding, and Schott, 2007, or from Mayer and Ottaviano, 2007), one can expect that acquiring firms are larger on average. The same applies to targets if acquirers "cherry-pick": only better performing firms are acquired, which should then be expected to be larger. In contrast, if firms buy poor performing firms ("lemons"), targets should be expected to be smaller. In addition, acquirers are also expected to be larger if a merger is financed by debts: larger firms face lower credit constraints, whereas smaller firms have a smaller equity basis to acquire other firms (Beck and Demirguc-Kunt, 2006; Audretsch and Elston, 2002). Contrary to this, Mueller (2003a) presents arguments why smaller firms self-select in horizontal mergers: assume an industry which exhibits significant scale economies, and the average costs decrease as the firm's scale increases. If the decrease in average costs falls as scale increases, cost reductions are higher for smaller firms and vice versa. Thus, if scale economies are the motivation for horizontal M&A, mostly smaller firms in an industry should merge.[14] Assuming that a firm's size is an indicator for higher market share, one can expect that a firm that merges in order to increase its market power primarily looks for larger targets. With respect to pre-merger skill-intensity, one could theoretically expect that plants with a higher skill-intensity self-select in M&A activity if also more productive plants self-select in M&A, because higher productivity may be a result of higher skill-intensity.

Summarizing the theories, there is no clear prediction for employment effects, but a negative employment effect seems more likely than a positive, as already argued by different authors, e.g. Conyon, Girma, Thompson, and Wright (2002a). Nevertheless, questions about net employment effects and corresponding changes in the merging plants' skill-intensities are passed to empirics, which have not been able to give a clear answer either.

Related literature: My paper relates to several earlier papers that also discussed the effects of M&A on employment. Several studies exist for the US. For example, Brown and Medoff (1988) reported negative effects on employment for smaller firms in the state of Michigan between 1978 and 1984. Similar results were found in Bhagat, Shleifer, and Vishny (1990) who provided evidence for negative effects based on labor data from press reports for the years 1984 to 1986. Lichtenberg and Siegel (1992b) also estimated negative employment effects for larger US manufacturing firms between 1977 and 1987. In contrast to these studies, positive effects were reported in a study from McGuckin, Nguyen, and Reznek (1995) who used US plant-level data. They found an increase in employment in acquired plants. However, they did not estimate significant employment effects at firm-level. In another study from McGuckin, Nguyen, and Reznek (1998), the authors analyzed the US food manufacturing sector for the period 1977 to 1987 and estimated positive employment effects. Again, positive effects from mergers were reported by McGuckin and Nguyen (2001) for the entire US manufacturing sector for the years 1977 to 1987, but they also found job losses after mergers in bigger plants. In a study by Ollinger, Nguyen, Blayney, Chambers, and Nelson (2005) about mergers between 1977 and 1987 in eight US food industries, employment increased after M&A.

The picture for Europe is also ambiguous, but with a tendency to negative effects. There are several studies for the UK. Conyon, Girma, Thompson, and Wright (2001) distinguished

[14] Mueller (1980a) did not empirically confirm this hypothesis.

between hostile and friendly mergers and found negative employment effects for the period from 1983 to 1996. Similar results were found in another study by Conyon, Girma, Thompson, and Wright (2002a). They analyzed employment effects after M&A between 1967 and 1996 and distinguished between related versus unrelated as well as friendly versus hostile takeovers, and they reported significant decreases in employment. Negative effects after a foreign takeover were also found in Girma and Görg (2004) for the electronics industry. Girma (2005) identified negative as well as positive employment effects for targets in the manufacturing sector after foreign takeovers, depending on their size. Also positive effects were found in Amess and Wright (2007), but they only analyzed management buyouts. In addition to these studies for the UK, further studies exist for other European countries. For the Austrian manufacturing sector, Bellak, Pfaffermayr, and Wild (2006) did not identify effects on employment growth after foreign acquisition. For France, Margolis (2006b) provided evidence for negative employment effects of M&A, with more workers laid off in acquired firms compared to their acquirers in the short term. Siegel and Simons (2010) reported negative employment effects for Sweden, but they were not able to differentiate between cross-border and domestic M&A. In a study about employment effects of M&A in Europe and the USA, Gugler and Yurtoglu (2004) estimated negative effects for Europe, but not for the USA. Lehto and Böckerman (2008) identified negative employment effects for cross-border mergers and for domestic mergers for Finnish firms. Arndt and Mattes (2010) did not find employment effects for cross-border M&A in Germany between 1997 and 2003. Mattes (2010) also used data for Germany for the years 2000 to 2007 and did not estimate employment effects for acquired plants after foreign takeovers.

In comparison to a considerable number of studies of M&A and employment effects, evidence for changes in plants' skill-intensity after M&A is rare, and it is also difficult to draw any general conclusions. For example, in a study based on US manufacturing data, Lichtenberg and Siegel (1990) reported employment losses mostly for central office staff. In a study about the UK electronics industry, Girma and Görg (2004) found negative employment effects from foreign takeovers, especially for unskilled workers. Lipsey and Sjöholm (2003) found a decrease in white-collar workers after M&A and an increase in blue-collar-workers in Indonesian manufacturing plants. In a study for Portugal, Almeida (2003) did not identify significant changes in the workforce educational composition after foreign acquisition. Huttunen (2007) showed that the share of highly educated workers declines, although slightly and slowly, after the ownership has changed from domestic to foreign in the acquired plants. In addition to this literature, there exist some empirical studies about effects from hostile takeovers[15] which point to job

[15]See Conyon, Girma, Thompson, and Wright (2001) for an overview of the empirical results on

losses especially for white-collar workers and less for blue-collar workers: for the UK, Franks and Mayer (1996) provided evidence for higher resign rates of directors after a hostile takeover in comparison to a friendly takeover; Hirshleifer and Thakor (1994) found out that mostly board members are subject to high levels of displacement after a hostile takeover for the US and UK; Bhagat, Shleifer, and Vishny (1990) reported about mainly white-collar job losses in the US due to hostile takeovers; Lichtenberg and Siegel (1990) came to similar results, also for the US.

4.3 The data

The dataset used for this investigation is a merged dataset from the Establishment Panel (Betriebspanel) of the Institute of Employment Research Nuremberg (Institut für Arbeitsmarkt- und Berufsforschung, IAB), and the M&A DATABASE of the University St. Gallen. The dataset was created by TNS Infratest Sozialforschung GmbH München.

The IAB Establishment Panel is a representative employer survey for Germany, and questions are on topics related to employment policy. The survey exists since 1993 and currently covers around 16,000 plants of all sizes and from all sectors of the economy.[16] The M&A DATABASE contains information about 65,000 transactions since 1985 for Germany, Austria and Switzerland, and it provides information about acquirer, target, and seller firms, their sales, profits, number of employees, location, and sector. The data allows further distinction between horizontal, vertical forward, vertical backward, conglomerate, or concentric mergers.[17] In this study, plants merged between 1996 and 2005. For this period, the dataset includes information about 23,717 transactions with 40,736 German firms.[18]

Based on the IAB Establishment Panel and the M&A DATABASE, TNS Infratest constructed a new dataset consisting of two groups of plants which exhibit a similar structure with regard to several characteristics: sector, size, location, and legal form. The first group, the treatment group, consists of plants which were involved in merger activity between January 1996 and December 2005. To create this group of merged plants,[19] all plants which appeared

employment effects after hostile mergers.

[16] See Fischer, Janik, Müller, and Schmucker (2009) for a comprehensive description of the IAB Establishment Panel.

[17] If firms from different but neighboring industries combine, the merger is concentric. This kind of merger is comparable to a conglomerate merger with complementary or neighboring products (Church, 2004).

[18] In a pilot study, Bellmann and Kirchof (2006) tested the capability of a similar dataset, combined from the IAB Establishment Panel and Thomson ONE Banker. The advantage of the M&A DATABASE compared to Thomson ONE Banker is that it also includes information about smaller firms.

[19] I use the terms "merged plants" and "treated" interchangeably.

in the M&A DATABASE as well as in the IAB Establishment Panel had to be identified. Plants were only considered as treated if they were surveyed at least once prior and once after the merger in order to have enough information. If plants merged more often between 1996 and 2005, they may also appear more often in the dataset. The original treatment group consists of 7,801 observations from 958 different plants.[20]

The second group, the control group, consists of plants which were not involved in any M&A activity between 1980 and 2005. This group was created in such a way that control plants are as similar as possible to treated plants in the treatment group, i.e. controls should act as statistical twins to the treated. Therefore, treated plants were categorized according to sector, size, legal form, and location (West or East Germany). Then, controls had to be identified with the same combination of sector, size, legal form, and location. Of course, not for every treated plant was it possible to find a control with an identical combination of characteristics. In addition, controls also had to be surveyed at least twice between 1993 and 2006,[21] and they were only kept in the dataset if they had not been involved in any merger activity since 1980. This was checked via other datasets or the plants' websites. Several plants appear more often in the control group because plants can serve several times as a control in the referred observation period. The original control group consists of 1,009 observations from 291 different plants.[22]

This new dataset has some useful features in comparison to datasets used in other studies:

- It allows distinguishing between acquirer, target, and seller plants and between different types of mergers, i.e. between horizontal and non-horizontal mergers.

- The data includes plants of different sizes, i.e. small and medium-size plants are also included in the analysis.

- Plants in the dataset are also not restricted to a certain sector, a location in West or East Germany, or a legal form.

- Due to the analysis at plant-level a plant can be assigned to a specific sector. This is often not possible for (mostly larger and multi-plant) firms.

[20]This means that the treatment group consists of 958 different plants which all have different identification numbers in the IAB Establishment Panel.

[21]The first wave of the IAB Establishment Panel was conducted in 1993, and the survey of 2006 includes questions referring to the preceding year 2005.

[22]See appendix C for a more detailed description about the construction of the treatment and control groups.

- There are only plants in the control group which had not merged for a long time, i.e. at least since 1980.
- The dataset makes a differentiated econometric analysis possible because of the availability of a control group, and a rich set of variables.

For the empirical strategy performed in the next section, some modifications of the dataset are inevitable. First, if a plant is affected by more than one merger activity within the same year, they are taken together as only one merger because much of the information is on a yearly basis (Gugler and Yurtoglu, 2004). Second, the analysis covers an observation period of four years from $t = 1$ to $t = 4$. Thus, treated and controls are only kept if data exists for all four successive years. Third, treated plants were dropped if there are less than three years between the merger activities of a single plant. This avoids overlapping effects, i.e. effects can be assigned to a specific merger. Fourth, observations which exhibit abnormal values for employment or skill-intensity growth rates will be deleted. I define abnormal values if growth rates deviate two standard deviations from the respective industry average.[23] Fifth, seller plants were excluded from the dataset because too few observations exist. Finally, in very few cases plants were involved in foreign M&A. However, the number was too small for a separate analysis, and thus, I also dropped these plants in order to study only domestic M&A.

Table 4.1: **Classification of treated plants**

Plants	Horizontal	Non-horizontal	Unknown	Control
Acquirers	55	7	26	88
Targets	40	21	16	77
Total	95	28	42	165

Note: there are plants which merged, but there is no information about the type of merger. These are labeled as "Unknown".

These modifications reduced the number of observations significantly:[24] from 7,801 to 165 treated, and from 1,009 to 563 untreated observations.[25] Table 4.1 presents an overview of

[23] These extreme values may be due to errors or rare events. For example, consider a plant that produces a certain machine in a year, and reports only low sales in the same year. If the plant sells the machine in the next year, it will report high sales. These extreme numbers may have a high impact on empirical results (Wagner, 2007b).

[24] Gugler and Yurtoglu (2004) use the Global Mergers and Acquisitions database for their study. They also analyze only a small fraction of the original dataset, i.e. from a large sample of 140,289 mergers there is only sufficient data for 646 mergers.

[25] As stated above, I only kept a multi-merging plant more than once in the dataset if there were at least three years between the mergers, in order to avoid overlapping effects. There are also several untreated plants that appear more often in the control group, because they were surveyed several

treated plants, i.e. plants are distinguished according to the subgroups of acquiring and target plants, and plants involved in horizontal and non-horizontal mergers. The number of observations in the subgroup of non-horizontal mergers is small, which leads to a limited robustness of estimations for this subgroup in the following empirical analysis. In addition, table 4.1 also shows that the type of merger is not known for every treated plant.

4.4 Empirical investigation

4.4.1 Empirical strategy

Employment and skill-intensity measurements: In this study, I analyze the total number of a plant's employment. The dataset allows distinguishing between employees who are liable to social security, and those who are not. Liable to social security means that employees and trainees are liable to health, pension, and unemployment insurance, or their contributions to pension insurance are partly paid by the employer. In contrast, workers who are not liable to social security are civil servants, self-employed persons, unpaid family workers, and so-called "marginal" part-time workers.[26]

In addition to employment analysis, I also analyze changes in the plants' skill-intensity. I define two groups of employees, respectively measured as proportion within total workforce. First, skilled employees, that is skilled workers, employees, and civil servants for qualified jobs, working proprietors, directors, and managers. Second, unskilled employees, that is unskilled or semi-skilled workers, employees and civil servants for menial jobs. I do not include trainees, apprentices, and candidates for civil service into any of these groups.[27]

---

times. Hence, one might think that it is not a good idea that a single plant is used more often as a control because it was surveyed several times, while another single plant will be used only once because it was surveyed less often. As a consequence, the results for controls may be biased towards the plants surveyed more often. Nevertheless, I do not worry about this for the following reasons: first, I do not assume that there is a systematic bias in relevant variables in those plants that were surveyed more often. In particular: why should a plant that was surveyed more often in the IAB Establishment Panel exhibit systematically higher or lower employment or skill-intensity changes over time? Second, if I allowed keeping untreated plants with only one observation in the control group, the control group would shrink to approximately a quarter of its size, and useful information would be lost. As a robustness check I also performed the empirical investigation with a control group with each individual plant appearing only once in the dataset. The estimations based on this smaller control group (133 plants) were similar as expected.

[26] These workers are either employed only short-term (i.e. for a maximum of two month or 50 days per year), or have an agreed working week of less than 15 hours and a monthly wage of max. EUR 400 (formerly DM 630). See the introduction to the questionnaire for the IAB Establishment Panel, available on the website of the Research Data Centre (FDZ) at the IAB (http://fdz.iab.de).

[27] At the beginning of their vocational training, trainees, apprentices, and candidates for civil service should be assigned to the group of unskilled employees, but at the end of their vocational training they

Plant- vs. firm-level analysis: Existing empirical studies differ with respect to the observation level, i.e. between plant- and firm-level. The firm-level typically includes headquarter activities like marketing, R&D, finance operations, etc., whereas plant-level includes activities like production and assembling (Barba Navaretti and Venables, 2004). Thus, firm- and plant-levels may be affected differently by mergers. For example, synergy effects in R&D or marketing should primarily be measured at firm-level, but if M&A leads to improved production processes, one would expect to observe the effects at plant-level.

The data at hand are at plant-level. That is, plant-level data cannot be used for any analysis at firm-level, and this may be seen as a drawback. However, this drawback is reduced if smaller and middle-size firms are also of interest. Firms of this size are often single-plant firms. It is a further advantage of plant-level analysis that plants can more easily be assigned to a specific industry sector compared to firms (Bellmann and Kirchhof, 2006). And, in addition, the plant is the unit which is fully involved in the transaction capturing the whole effects of M&A, whereas employment and skill-intensity effects may disperse at firm-level which measures the average employment and skill-intensity of all plants.

Observation period: I create an observation period of four years, i.e. from $t = 1$ to $t = 4$, with mergers all occurring in $t = 2$. The dataset contains information about mergers that occur between 1996 and 2005, i.e. year $t = 2$ corresponds to one of these years. Consequently, year $t = 1$ corresponds to a year between 1995 and 2004, and year $t = 4$ corresponds to a year between 1998 and 2007. This leads to ten cohorts with a respective four-year window (1995 - 1998, 1996 - 1999, ..., 2004 - 2007). For controls, the first possible year for $t = 1$ is 1993 (the starting year of the IAB Establishment Panel), and the last possible year for $t = 4$ is 2005, leading to ten cohorts with a four-year window (1993 - 1996, 1994 - 1997, ..., 2002 - 2005).[28] Figure D1 in the appendix provides a graphical illustration of this "rolling observation window", which is common in empirical studies (e.g. Schank, Schnabel, and Wagner, 2010).

The four-year observation period makes the analysis of both plants' pre-merger and post-

should be assigned to the group of skilled employees. However, the data does not include respective information. For this reason, I performed the empirical analysis for both cases, and also estimated results if these workers were excluded from the analysis. The results were all similar, and thus, I decided to exclude them, because an assignment to one of both groups would be arbitrary.

[28] The dataset includes plants which merged between 1996 and 2005. However, some of the four-year cohorts for merged plants also cover the years 1995, 2006, and 2007, but there is no information if plants also merged in these years, or only between 1996 and 2005. For example, a plant that merged in 2005 could also merge in 2007 again. If this was true, my results would be biased by overlapping effects due to more mergers. Nevertheless, I choose to keep these observations because I already excluded those multi-merges from the dataset (see section 4.3), and thus, it is not too likely that the remaining plants also merged in 1995, 2006, or 2007.

merger employment and skill-intensity performance possible. Generally, a longer observation period prior to the merger would be desirable, for example because a decreasing performance could be the trigger for a merger, or because of the existence of Ashenfelter's Dip:[29] if plants prepare themselves for the merger, employment can already be adjusted prior to the merger, and estimations about post-merger effects may then be biased. However, due to data limitation, I am not able to lengthen the observation period to several years prior to the merger. Therefore, I follow most of the existing studies which also start their observation period one year prior to the merger (e.g. Girma and Görg, 2007).

There are good reasons to assume that it takes some time for effects to arise after mergers (e.g. for the reorganization of the merging plant's production). For this reason, I calculate post-merger performance in years $t = 2$ and $t = 4$. I believe that this time window is long enough to capture most of the merger-induced changes.[30] This also allows the calculation of post-merger growth rates instead of levels, and helps to reduce selection bias that arises due to unobserved time-invariant differences between both groups, as will be discussed later in the paper. In addition to changes between years $t = 2$ and $t = 4$ ("post-merger period"), I also analyze growth rates over the whole observation period between $t = 1$ and $t = 4$ ("total period") in order to capture changes that occur within several weeks or months around the merger.[31] Finally, sales are deflated by the aggregated consumer price index over the whole observation period.

The selection problem and the methods applied: A simple comparison of performance parameters between plants that merged and plants that had not merged may show differences between both groups. However, this comparison is not able to analyze a causal effect of M&A on performance parameters like employment or skill-intensity for a simple reason: there may be a selection of plants in M&A activity, i.e. plants that merge are not randomly assigned to the group of merging plants. Instead, it is more likely that plants with certain pre-merger characteristics become acquirers or targets. Thus, if plants are not selected randomly in the merger activity, a simple comparison between merged and non-merged plants suffers from a selection bias.

The empirical analysis in this paper aims at solving this selection problem to identify causal

[29]The Ashelfelter Dip describes that the unemployeds' efforts in job-seeking decrease shortly before they participate in a labor-market program.

[30]Similar, Maksimovic, Phillips, and Prabhala (2011) found that most of the restructuring occurs within a three-year period. Changing the time window to five years does not change their results.

[31]In particular, most data in the IAB Establishment Panel refers to June 30th of the respective year. Mergers occur in year $t = 2$, i.e. somewhere between July 1st and June 30th of the following year.

effects. As usual in empirical studies, I start my investigation with an analysis of descriptive statistics. I analyze different pre-merger characteristics between both groups, compare employment (skill-intensity) over time as well as growth rates for treated and control groups.

However, employment (skill-intensity) differences between merging and non-merging plants may be the result of differences in other variables than M&A that determine employment and skill-intensity. For this, I apply regression analysis to control for variables that influence employment (skill-intensity). The construction of the framework is similar to Schank, Schnabel, and Wagner (2010): the regression allows looking at differences in the average plant employment (skill-intensity) between both groups over time, and it controls for plant characteristics that are thought to be related to a plant's average employment (skill-intensity). Nevertheless, regression analysis is able to detect correlations, but it is not a reliable method for identifying causalities (Backhaus, Erichson, Plinke, and Weiber, 2010).[32]

Econometric evaluation methods like propensity score matching introduced by Rosenbaum and Rubin (1983, 1985) can be regarded as a solution to the selection problem and identify causalities. The problem is that a faster employment (skill-intensity) growth of plants that merged does not necessarily reflect a causal effect of M&A on the plants' employment (skill-intensity). Instead, plants with higher employment (skill-intensity) could also self-select in the merged group, but would have experienced higher growth even without merging. Since both states are never observable at the same time, the problem of the missing counterfactual situation arises (Schank, Schnabel, and Wagner, 2010). With propensity score matching, this missing counterfactual can be replaced by the construction of an appropriate control group, i.e. the method pairs merged plants with controls that are similar in their pre-merger characteristics. Therefore, they exhibit a similar probability of merging. If there are performance differences, they can then be attributed to the merger. As suggested by Blundell and Costa Dias (2000) and Smith and Todd (2005a), I combine matching and a difference-in-differences estimator to compare changes instead of levels in order to conduct causal analysis without neglecting time-invariant heterogeneity between treated and control plants.

Throughout the empirical investigation, I first perform the analysis for employment effects, and then for skill-intensity effects. I discuss the results for the treatment group including all treated, and the control group. I will also perform the analysis for all four subgroups: acquirers, targets, plants in horizontal mergers, and plants in non-horizontal mergers.

---

[32] However, Angrist and Pischke (2009) argue that regression analysis is also able to solve the selection problem if it controls for the correct covariates. See Caliendo (2006) or Gelman and Hill (2007) for a further discussion of regression analysis compared to matching methods.

4.4.2 Descriptive statistics

Table 4.2 presents summary statistics for different variables which are all measured prior to the merger in $t = 1$. The table compares the different treatment groups to the control group. Plants are assigned to different size categories, and the table shows that mergers are concentrated in larger plants (percentage numbers in parentheses): for example, while approximately 67% of all merged plants have 100 and more employees, approximately 60% of all control plants have less than 100 employees. The majority of treated plants belongs to the sectors of production goods, investment goods, and trade. Most of the plants in the different groups have the legal form "Limited", but the share of plants that have the legal form "Limited by shares" is clearly higher in treated groups compared to the control group. Most of the plants are located in West Germany, and this is similar across all groups. A comparison of plants with respect to labor productivity (sales per employee), wages (per month), and export rates (as percentage of sales) shows that merging plants are on average more productive, pay higher wages, and export more. In addition, treated plants are more often in foreign ownership compared to control plants, and they offer further training more often. Finally, treated plants are less often single-plant firms (e.g. independent plants or head offices) compared to control plants. All in all, summary statistics from table 4.2 point to a substantial pre-merger heterogeneity between merging and non-merging plants, and the numbers do not substantially vary between subgroups.

Descriptive statistics for employment: In table 4.3 I present the number of employees over the years $t = 1$ to $t = 4$ for all groups, and also describe growth rates. In year $t = 1$, the average number of employees is higher in all treatment groups compared to the control group. With an average number of employees ranging from 378 (non-horizontal) to 516 (horizontal) in treated plants, the size differential of merging plants is between approximately 215% and 330%, and the size gap persists over the whole observation period. In order to test the statistical significance of differences between each treatment group and the control group in all years, I apply a t-test which does not assume equal variances in both groups.[33] In particular, I test the null hypothesis H_0: mean of employees of treated = mean of employees of controls, against the alternative hypothesis H_1: mean of employees of treated $\neq$ mean of employees of controls. P-values of at most 0.01, 0.05, or 0.1 indicate that the null hypothesis H_0 can be rejected at the 1%-, 5%-, or 10%-error level in favor of the alternative hypothesis H_1. For all treatment groups, the average size is statistically significantly higher at the 1%- or 5%-significance-level

[33] I apply Levene's statistic for a test of the equality of variances.

Table 4.2: **Summary statistics: different variables**

Variable	All	Acquirers	Targets	Horizontal	Non-horiz.	Control
Size category						
1-19 employees	2 (12.73)	9 (10.34)	11 (14.29)	9 (10.23)	12 15.58	127 (22.56)
20-49 employees	22 (13.33)	14 (16.09)	8 (10.39)	14 (15.91)	8 10.39	110 (19.54)
50-99 employees	12 (7.27)	5 (5.75)	7 (9.09)	5 (5.68)	7 9.09	96 (17.05)
100-299 employees	40 (24.24)	26 (29.89)	14 (18.18)	26 (29.55)	14 18.18	174 (30.91)
300-499 employees	24 (14.55)	10 (11.49)	14 (18.18)	11 (12.50)	13 16.88	39 (6.93)
>=500 employees	46 (27.88)	23 (26.44)	23 (29.87)	23 (26.14)	23 29.87	17 (3.02)
Total	165 (100.00)	87 (100.00)	77 (100.00)	88 (100.00)	77 (100.00)	563 (100.00)
Sector						
Mining/Quarrying/Electricity	13 (7.88)	10 (11.36)	3 (3.90)	11 (11.58)	2 (7.14)	5 (0.89)
Food	3 (1.82)	3 (3.41)	0 (0.00)	2 (2.11)	0 (0.00)	26 (4.62)
Consumer goods	12 (7.27)	6 (6.82)	6 (7.79)	8 (8.42)	0 (0.00)	34 (6.04)
Production goods	36 (21.82)	13 (14.77)	23 (29.87)	20 (21.05)	6 (21.43)	82 (14.56)
Investment goods	40 (24.24)	21 (23.86)	19 (24.68)	20 (21.05)	9 (32.14)	155 (27.53)
Construction	8 (4.85)	4 (4.55)	4 (5.19)	0 (0.00)	1 (3.57)	36 (6.39)
Trade	23 (13.94)	13 (14.77)	10 (12.99)	13 (13.68)	4 (14.29)	115 (20.43)
Transport	10 (6.06)	9 (10.23)	1 (1.30)	7 (7.37)	1 (3.57)	24 (4.26)
Hotels/Restaurants	2 (1.21)	0 (0.00)	2 (2.60)	1 (1.05)	1 (3.57)	0 (0.00)
Education	1 (0.61)	0 (0.00)	1 (1.30)	1 (1.05)	0 (0.00)	7 (1.24)
Human health	1 (0.61)	0 (0.00)	1 (1.30)	1 (1.05)	0 (0.00)	0 (0.00)
Research/Computer/Ser	9 (5.45)	5 (5.68)	4 (5.19)	6 (6.32)	3 (10.71)	74 (11.37)
Other services	7 (4.24)	4 (4.55)	3 (3.90)	5 (5.26)	1 (3.57)	15 (2.66)
Total	165 (100.00)	88 (100.00)	77 (100.00)	95 (100.00)	28 (100.00	563 (100.00)
Legal form						
Partnership, individually-owned, etc.	16 (9.82)	8 (9.20)	8 (10.53)	7 (7.45)	4 (14.82)	119 (21.99)
Limited	118 (72.39)	59 (67.82)	59 (77.63)	64 (68.09)	21 (77.78)	402 (74.31)
Limited by shares	29 (17.79)	20 (22.99)	9 (11.84)	23 (24.47)	2 (7.40)	20 (3.70)
Total	163 (100.00)	87 (100.00)	76 (100.00)	94 (100.00)	27 (100.00)	541 (100.00)
Location in East Germany	52 (31.71)	29 (32.95)	23 (30.26)	33 (34.74)	7 (25.00)	153 (27.22)
Labor productivity (mean)	204344.5	230619.3	173046.6	221995.9	167243.7	100059.1
Wage per month (mean)	2170.6	2210.6	2121.2	2215.5	2152.27	1742.3
Export rate (% of sales, mean)	24.09	23.04	25.33	24.41	26.70	12.46
Plant is in foreign ownership	19 (14.29)	8 (12.12)	11 (16.42)	14 (15.05)	2 (8.00)	5 (1.13)
Further training	150 (90.91)	79 (89.77)	71 (92.21)	86 (90.53)	26 (92.86)	373 (67.57)
Single-plant firm	51 (31.29)	20 (22.99)	31 (40.79)	27 (29.03)	15 (53.57)	539 (80.52)

Note: numbers refer to $t = 1$. Numbers in parentheses are percentage of total number of plants. Reduced number of observations is due to missing data for several variables.

compared to the control group in all years. In addition, I also test the statistical significance of size differences between subgroups, but these are not statistically significant at a usual level (not reported).

Table 4.3: **Summary statistics: employment**

All	Mean	(Std. Dev.)	N	P-value
No. of employees in t=1	461.0	(736.3)	165	0.000
No. of employees in t=2	453.2	(724.0)	165	0.000
No. of employees in t=3	450.5	(733.5)	165	0.000
No. of employees in t=4	450.0	(738.6)	165	0.000
Growth from t=2 to t=4	-0.028	(0.254)	165	0.934
Growth from t=1 to t=4	-0.040	(0.294)	165	0.664
Acquirers	**Mean**	**(Std. Dev.)**	**N**	**P-value**
No. of employees in t=1	419.7	(554.2)	88	0.000
No. of employees in t=2	413.3	(544.9)	88	0.000
No. of employees in t=3	408.8	(550.7)	88	0.000
No. of employees in t=4	406.7	(544.9)	88	0.000
Growth from t=2 to t=4	-0.033	(0.299)	88	0.820
Growth from t=1 to t=4	-0.050	(0.334)	88	0.560
Targets	**Mean**	**(Std. Dev.)**	**N**	**P-value**
No. of employees in t=1	508.3	(902.2)	77	0.000
No. of employees in t=2	498.7	(887.2)	77	0.000
No. of employees in t=3	498.1	(899.8)	77	0.000
No. of employees in t=4	499.4	(912.5)	77	0.000
Growth from t=2 to t=4	-0.021	(0.193)	77	0.850
Growth from t=1 to t=4	-0.027	(0.244)	77	0.987
Horizontal	**Mean**	**(Std. Dev.)**	**N**	**P-value**
No. of employees in t=1	515.9	(877.6)	95	0.000
No. of employees in t=2	511.3	(867.3)	95	0.000
No. of employees in t=3	503.9	(875.9)	95	0.000
No. of employees in t=4	506.0	(884.8)	95	0.000
Growth from t=2 to t=4	-0.030	(0.297)	95	0.906
Growth from t=1 to t=4	-0.052	(0.333)	95	0.517
Non-horizontal	**Mean**	**(Std. Dev.)**	**N**	**P-value**
No. of employees in t=1	377.8	(538.3)	28	0.018
No. of employees in t=2	362.9	(511.0)	28	0.018
No. of employees in t=3	379.2	(550.0)	28	0.019
No. of employees in t=4	366.7	(527.4)	28	0.019
Growth from t=2 to t=4	-0.033	(0.15)	28	0.815
Growth from t=1 to t=4	-0.026	(0.145)	28	0.948

Control	Mean	(Std. Dev.)	N
No. of employees in t=1	120.4	(133.0)	563
No. of employees in t=2	119.5	(129.9)	563
No. of employees in t=3	118.5	(128.3)	563
No. of employees in t=4	117.7	(127.8)	563
Growth from t=2 to t=4	-0.026	(0.282)	563
Growth from t=1 to t=4	-0.028	(0.33)	563

Note: p-values refer to the t-test of statistical significance of difference of means between treated and control groups.

Table 4.3 also reports growth rates for the number of employees. I calculate changes between $t = 2$ and $t = 4$, and between $t = 1$ and $t = 4$. The percentage growth rates are approximated by logarithms:

$$lnEMPLOY_{(2-4)i} = lnEMPLOY_{4i} - lnEMPLOY_{2i}, \tag{1}$$

and

$$lnEMPLOY_{(1-4)i} = lnEMPLOY_{4i} - lnEMPLOY_{1i}. \tag{2}$$

$lnEMPLOY_{(2-4)i}$ and $lnEMPLOY_{(1-4)i}$ are the variables for percentage changes of a plant i's number of employees between $t = 2$ and $t = 4$, and between $t = 1$ and $t = 4$, while $EMPLOY_{ti}$ with $t = \{1, ..., 4\}$ describes the number of employees in year t. All treatment groups and the control group exhibit negative mean employment changes for both observation periods. Again, I apply t-tests which do not assume equal variances, and I test the null hypothesis H_0: mean of employment growth of treated = mean of employment growth of controls, against the alternative hypothesis H_1: mean of employment growth of treated $\neq$ mean of employment growth of controls. The negative growth rates of the respective treatment groups are statistically not significantly different from the negative growth rates of the control group at any usual level, implying that M&A has no statistically significant effect on employment.

Descriptive statistics for skill-intensity: In table 4.4 I analyze the plants' skill-intensity over time. The merging plants exhibit a higher skill-intensity than plants that do not merge. In $t = 1$ the merging plants' average skill-intensity is between 74% (acquirer) and 81% (non-horizontal) compared to 65% in non-merging plants. Similar to above, I apply a t-test to test whether the mean differences between treated and non-treated plants are statistically significant. The table shows that a statistically significant difference in skill-intensity exists for all groups compared to controls, and the difference is statistically significant at usual levels over the whole observation, except for year $t = 4$ in the group of non-horizontally merging plants.

In order to take a closer look at the development path of skill-intensity over the four years, I estimate growth rates. Similar to employment changes, I calculate skill-intensity changes between $t = 2$ and $t = 4$, and between $t = 1$ and $t = 4$. Again, the percentage growth rates are approximated by logarithms:

$$lnSKILLED_{(2-4)i} = lnSKILLED_{4i} - lnSKILLED_{2i}, \tag{3}$$

and

$$lnSKILLED_{(1-4)i} = lnSKILLED_{4i} - lnSKILLED_{1i}. \tag{4}$$

Table 4.4: **Summary statistics: skill-intensity**

All	Mean	(Std. Dev.)	N	P-value
Proportion of skilled in t=1	0.751	(0.240)	165	0.000
Proportion of skilled in t=2	0.738	(0.228)	165	0.000
Proportion of skilled in t=3	0.753	(0.231)	165	0.000
Proportion of skilled in t=4	0.781	(0.233)	165	0.000
Growth from t=2 to t=4	0.055	(0.353)	165	0.983
Growth from t=1 to t=4	0.046	(0.396)	165	0.237
Growth from t=1 to t=2	-0.009	(0.340)	165	0.155
Acquirers	Mean	(Std. Dev.)	N	P-value
Proportion of skilled in t=1	0.737	(0.248)	88	0.004
Proportion of skilled in t=2	0.751	(0.225)	88	0.002
Proportion of skilled in t=3	0.751	(0.234)	88	0.008
Proportion of skilled in t=4	0.776	(0.238)	88	0.002
Growth from t=2 to t=4	0.025	(0.319)	88	0.449
Growth from t=1 to t=4	0.062	(0.397)	88	0.563
Growth from t=1 to t=2	0.037	(0.313)	88	0.978
Targets	Mean	(Std. Dev.)	N	P-value
Proportion of skilled in t=1	0.766	(0.231)	77	0.000
Proportion of skilled in t=2	0.724	(0.233)	77	0.057
Proportion of skilled in t=3	0.754	(0.230)	77	0.008
Proportion of skilled in t=4	0.787	(0.229)	77	0.000
Growth from t=2 to t=4	0.089	(0.388)	77	0.466
Growth from t=1 to t=4	0.027	(0.396)	77	0.213
Growth from t=1 to t=2	-0.062	(0.363)	77	0.031
Horizontal	Mean	(Std. Dev.)	N	P-value
Proportion of skilled in t=1	0.762	(0.227)	95	0.000
Proportion of skilled in t=2	0.744	(0.229)	95	0.004
Proportion of skilled in t=3	0.746	(0.236)	95	0.001
Proportion of skilled in t=4	0.815	(0.221)	95	0.000
Growth from t=2 to t=4	0.101	(0.370)	95	0.273
Growth from t=1 to t=4	0.071	(0.395)	95	0.686
Growth from t=1 to t=2	-0.030	(0.324)	95	0.080
Non-horizontal	Mean	(Std. Dev.)	N	P-value
Proportion of skilled in t=1	0.811	(0.202)	28	0.000
Proportion of skilled in t=2	0.771	(0.212)	28	0.018
Proportion of skilled in t=3	0.787	(0.224)	28	0.018
Proportion of skilled in t=4	0.757	(0.247)	28	0.162
Growth from t=2 to t=4	-0.047	(0.310)	28	0.110
Growth from t=1 to t=4	-0.101	(0.335)	28	0.074
Growth from t=1 to t=2	-0.053	(0.154)	28	0.011

Control	Mean	(Std. Dev.)	N
Proportion of skilled in t=1	0.653	(0.292)	562
Proportion of skilled in t=2	0.668	(0.283)	562
Proportion of skilled in t=3	0.677	(0.279)	562
Proportion of skilled in t=4	0.688	(0.274)	562
Growth from t=2 to t=4	0.054	(0.439)	562
Growth from t=1 to t=4	0.089	(0.463)	562
Growth from t=1 to t=2	0.036	(0.411)	559

Note: p-values refer to the t-test of statistical significance of difference of means between treated and control groups.

Table 4.4 shows that for some subgroups mean growth rates between $t = 2$ and $t = 4$ are higher than mean growth rates between $t = 1$ to $t = 4$, pointing to a U-shaped growth path.

For this, I additionally report percentage changes in plants' skill-intensity between $t = 1$ and $t = 2$:

$$lnSKILLED_{(1-2)i} = lnSKILLED_{2i} - lnSKILLED_{1i}. \quad (5)$$

I apply a t-test once more to test the statistical significance of differences in means of skill-intensity growth rates between treated and controls. The mean decrease between $t = 1$ and $t = 2$ is statistically significantly different from controls for the subgroups of targets, horizontal, and non-horizontal M&A at usual significance levels. Since changes between $t = 1$ and $t = 4$ are not statistically significant in subgroups of targets and horizontal M&A, the average growth rates imply that the skill-intensity in these two subgroups decreases after M&A, but turns back to a level that is not statistically significantly different from the pre-merger level, and hence, creates a U-shaped development path.

Summarizing the results from the descriptive statistics, tables 4.2, 4.3, and 4.4 exhibit a substantial pre-merger heterogeneity between merged and non-merged plants, i.e. merging plants are obviously different from non-merging plants with respect to different characteristics even before they merge. In particular, descriptive statistics show that merging plants are on average larger and have a higher skill-intensity compared to plants that do not merge. These findings give strong support to a self-selection hypothesis. I do not find support for employment changes over time, but with respect to plants' skill-intensity, I identify a U-shaped development path for subgroups of targets and horizontal M&A.

4.4.3 Regression analysis

Since descriptive analysis has shown considerable differences in size and skill-intensity between treated and control plants, it is of interest to analyze whether these differences are due to self-selection or if they are causally determined. In the next step, I apply an OLS-regression analysis with logarithm of employment as a proxy for plants' size in order to investigate the effect of some interaction variables of interest. I will repeat the analysis with respect to plants' skill-intensities. Then, in section 4.4.4 I will conduct a difference-in-differences propensity score matching analysis to consider issues of causality and firm heterogeneity.

Employment regressions: The regression equation with logarithm of employment as dependent variable is

$$lnEMPLOY_{it} = \beta_0 + \beta_1 MA_i + \sum_{t=2}^{4} \beta_t (MA_i * PERIOD_t) + \beta_5 CONTROL_{it} + \epsilon_i. \quad (6)$$

The variable $lnEMPLOY_{it}$ describes the logarithm of employees in plant i at year t. The dummy MA_i is one if a plant i merged and zero if the plant is a control. The regression coefficient β_1 measures the average size differential of merged plants prior to the merger in $t = 1$. To estimate employment changes over time I construct three interaction dummies $MA_i * PERIOD_t$ for the years $t = 2$ to $t = 4$, with $t = 1$ being the reference period. These dummies are a product of the M&A dummy (MA_i) and a time dummy ($PERIOD_t$, with $t \in [2, 4]$). The coefficients β_t measure whether the size differential between treated and controls becomes smaller or larger over time. The variable $CONTROL_{it}$ is a vector for several control variables, and ϵ_i is an error term. I perform this regression for all five groups of treated, i.e. the first including all treated, and the others for the respective subgroups of treated.

The results are displayed in table 4.5. The coefficient of the M&A dummy in the first regression specification is 0.485, and corresponds to a size differential of approximately 62%,[34] i.e. ceteris paribus, the number of employees in treated plants is 62% higher compared to control plants in $t = 1$. The coefficient is statistically significant at the 1%-level. Coefficients of the interaction dummies are either positive or negative, but they are small from an economic point of view, and not statistically significant at any usual significance level. In other words, the size gap between treated and control plants does not change in any statistically significant way over time. The coefficients of the control variables have all the expected signs and are in line with findings from descriptive statistics: the higher a plant's investment per employee, the larger the plant. The coefficient of the dummy variable for further training is also positive and highly significant. Moreover, plants that have the legal form "Limited", or "Limited by shares" are statistically significantly larger compared to plants that have other legal forms. Plants located in East Germany are smaller, while being foreign owned has no statistically significant correlation with a plant's size. If the plant is a single plant, its size is statistically significantly smaller.

The estimations for the regressions for subgroups of treated are similar. Regressions for acquirers, targets, and plants in horizontal mergers exhibit a highly statistically significant pre-merger size differential for treated between 50% (target) and 85% (acquirer). Only for the subgroup for non-horizontally merging plants I do not estimate a statistically significant pre-merger size difference. Similar to the first regression, the coefficients of the interaction dummies are not statistically significant, i.e. the size difference does not change over the years in any subgroup, and coefficients of control variables are also similar across these regressions.

[34] In loglinear models coefficients can be transformed into a percentage change. Here, $exp\{0.485\} \approx 1.62$, which corresponds to a differential of approximately 62% (Verbeek, 2005).

Table 4.5: **OLS-regression (dependent variable: log. employment)**

Variables	All (1)	Acquirers (2)	Targets (3)	Horizontal (4)	Non-horiz. (5)
M&A (D)	0.485*** (0.171)				
M&A*Period=2	0.013 (0.053)				
M&A*Period=3	-0.063 (0.070)				
M&A*Period=4	-0.085 (0.096)				
Acquirer (D)		0.616*** (0.221)			
Acquirer*Period=2		-0.040 (0.081)			
Acquirer*Period=3		-0.132 (0.099)			
Acquirer*Period=4		-0.162 (0.128)			
Target (D)			0.404* (0.221)		
Target*Period=2			0.058 (0.070)		
Target*Period=3			-0.000 (0.090)		
Target*Period=4			-0.018 (0.137)		
Horizontal M&A (D)				0.606*** (0.212)	
Horizontal*Period=2				-0.050 (0.071)	
Horizontal*Period=3				-0.087 (0.102)	
Horizontal*Period=4				-0.050 (0.151)	
Non-horizontal M&A (D)					0.060 (0.387)
Non-horizontal*Period=2					0.052 (0.129)
Non-horizontal*Period=3					-0.076 (0.153)
Non-horizontal*Period=4					-0.337 (0.240)
log. Investment p. employee	0.045*** (0.011)	0.038*** (0.012)	0.030*** (0.011)	0.040*** (0.011)	0.024** (0.011)
Further training (D)	1.162*** (0.101)	1.178*** (0.101)	1.130*** (0.103)	1.178*** (0.101)	1.154*** (0.103)
Legal form "Limited" (D)	0.483*** (0.138)	0.492*** (0.136)	0.625*** (0.143)	0.511*** (0.136)	0.634*** (0.144)
Legal form "Limited by shares" (D)	0.633** (0.245)	0.656*** (0.248)	0.218 (0.286)	0.731*** (0.251)	-0.139 (0.284)
Location in East Germany (D)	-0.474*** (0.132)	-0.448*** (0.141)	-0.368*** (0.133)	-0.406*** (0.134)	-0.372** (0.145)
Foreign owned plant (D)	0.155 (0.239)	0.334 (0.342)	0.325 (0.275)	0.278 (0.278)	0.715* (0.407)
Single-plant firm (D)	-0.334*** (0.122)	-0.318** (0.126)	-0.438*** (0.118)	-0.294** (0.125)	-0.513*** (0.132)
Constant	2.191*** (0.684)	2.955*** (0.357)	2.617*** (0.943)	1.688 (1.123)	3.807*** (0.463)
Observations	2566	2302	2246	2316	2077
R^2	0.413	0.399	0.428	0.417	0.404

Notes: robust standard errors in parentheses (adjusted for intragroup correlation). *** p<0.01, ** p<0.05, * p<0.1. (D) means variable is a dummy. Reference categories for legal form is "Partnership, individually-owned, public, and others". Regressions include dummies for sectors and years. The reduced number of observations is due to missing data for several variables. Data source: IAB Establishment Panel, M&A DATABASE St. Gallen.

The robustness of estimations for all five regressions is satisfying. The value for R^2 is between 0.399 and 0.428 which is sufficiently high, and the results of other tests (not reported) like F-test, variance-inflation-factor (VIF), and Durbin-Watson-test, controlling for the significance of the overall model, multicollinearity, and autocorrelation also yield good results. Moreover, the coefficients of variables have the expected signs from an economic point of view.

In addition, I also eliminated intragroup correlation and corrected biased standard errors.[35]

The results for employment so far are in line with findings from descriptive statistics. Estimations point to a self-selection of larger plants into M&A activity, and there is no evidence that plants change employment after M&A in a statistically significant way. This is true for all groups of treated. However, the estimated size differential in treated plants prior to the merger is substantially smaller compared to the results from descriptive statistics, i.e the size difference between merging and non-merging plants is obviously also explained by differences in other variables than M&A that determine the number of employees.

Skill-intensity regressions: Similar to the regressions above, I also perform regression analysis with respect to skill-intensities. The corresponding regression equation is

$$lnSKILLED_{it} = \beta_0 + \beta_1 MA_i + \sum_{t=2}^{4} \beta_t (MA_i * PERIOD_t) + \beta_5 CONTROL_{it} + \epsilon_i. \quad (7)$$

The explanation of regression equation 7 is similar to equation 6. The dependent variable $lnSKILLED_{it}$ describes the logarithm of the skill-intensity of the plant's workforce.

Table 4.6 presents the estimations. The coefficient of the M&A dummy in the first regression specification is 0.074 and corresponds to a pre-merger skill differential in treated plants of approximately 8%. In contrast to descriptive statistics, it is not statistically significant. The coefficients of the interaction terms are also statistically insignificant and do not point to a change in the plant's skill-intensity. Several control variables are also included in the regression: the larger the plant, the lower the skill-intensity. This may be because each plant - also smaller plants - has a management. The smaller the plant, the higher the proportion of the management, which is by my definition also part of the skilled workforce. The coefficient of the variable for squared employment is positive and statistically significant, i.e. as plants become very large, their skill-intensity increases. Moreover, the higher a plant's sales volume, the higher a plant's skill-intensity. Furthermore, as the proportion of female employees increases, the skill-intensity of a plant decreases, which may be caused by low-skilled part-time female workers. There is no statistically significant correlation between a plant's investment and its skill-intensity, but further training obviously increases the skill-intensity of a plant's workforce. Plants with the legal form "Limited" have a lower skill-intensity, while plants which are located

[35] In order to eliminate intragroup correlation - i.e. in a panel context, observations on the same individual may be correlated in different time periods, but observations on different individuals are not correlated (Baum, Schaffer, and Stillman, 2003) - I cluster observations of the same plants. This corrects biased standard errors.

Table 4.6: **OLS-regression (dependent variable: log. skill-intensity)**

Variables	All (1)	Acquirers (2)	Targets (3)	Horizontal (4)	Non-horiz. (5)
M&A (D)	0.074 (0.059)				
M&A*Period=2 (D)	-0.029 (0.030)				
M&A*Period=3 (D)	-0.033 (0.033)				
M&A*Period=4 (D)	-0.010 (0.043)				
Acquirer (D)		-0.024 (0.082)			
Acquirer*Period=2 (D)		0.061 (0.041)			
Acquirer*Period=3 (D)		-0.011 (0.050)			
Acquirer*Period=4 (D)		0.067 (0.061)			
Target (D)			0.106 (0.077)		
Target*Period=2 (D)			-0.138*** (0.046)		
Target*Period=3 (D)			-0.052 (0.051)		
Target*Period=4 (D)			-0.096 (0.061)		
Horizontal M&A				0.005 (0.078)	
Horizontal*Period=4 (D)				-0.003 (0.041)	
Horizontal*Period=4 (D)				-0.075 (0.053)	
Horizontal*Period=4 (D)				0.009 (0.065)	
Non-horizontal M&A					0.144 (0.110)
Non-horizontal*Period=2 (D)					-0.128*** (0.044)
Non-horizontal*Period=3 (D)					-0.023 (0.090)
Non-horizontal*Period=4 (D)					-0.158 (0.104)
log. Employment	-0.536*** (0.062)	-0.595*** (0.074)	-0.625*** (0.070)	-0.605*** (0.068)	-0.648*** (0.088)
Squared log. Employment	0.018*** (0.006)	0.022*** (0.008)	0.021*** (0.007)	0.024*** (0.006)	0.021** (0.010)
lnSales	0.248*** (0.037)	0.267*** (0.040)	0.302*** (0.041)	0.270*** (0.040)	0.317*** (0.045)
Proportion of female employees	-0.782*** (0.096)	-0.823*** (0.106)	-0.896*** (0.102)	-0.837*** (0.102)	-0.978*** (0.103)
log. Investment p. employee	0.006 (0.005)	0.010* (0.005)	0.007 (0.005)	0.009* (0.005)	0.012** (0.006)
Further training (D)	0.157*** (0.044)	0.162*** (0.046)	0.135*** (0.044)	0.149*** (0.046)	0.132*** (0.046)
Legal form "Limited" (D)	-0.197*** (0.041)	-0.189*** (0.042)	-0.197*** (0.043)	-0.195*** (0.044)	-0.183*** (0.043)
Legal form "Limited by shares" (D)	0.069 (0.062)	0.137* (0.072)	0.006 (0.073)	0.094 (0.072)	-0.025 (0.083)
Location in East Germany (D)	0.327*** (0.036)	0.323*** (0.039)	0.309*** (0.041)	0.321*** (0.040)	0.285*** (0.043)
Foreign owned plant (D)	0.102 (0.080)	0.130 (0.138)	0.216** (0.086)	0.132 (0.096)	0.334** (0.139)
Single-plant firm (D)	0.086** (0.044)	0.087* (0.049)	0.067 (0.051)	0.111** (0.048)	0.071 (0.053)
Constant	-1.944*** (0.363)	-2.259*** (0.322)	-2.519*** (0.448)	-2.590*** (0.427)	-2.428*** (0.503)
Observations	2461	2226	2169	2227	2024
R^2	0.373	0.391	0.412	0.395	0.438

Notes: robust standard errors in parentheses (adjusted for intragroup correlation). *** p<0.01, ** p<0.05, * p<0.1. (D) means variable is a dummy. Reference categories for legal form is "Partnership, individually-owned, public, and others". Regressions include dummies for sectors and years. The reduced number of observations is due to missing data for several variables. Data source: IAB Establishment Panel, M&A DATABASE St. Gallen.

in East Germany have a higher skill-intensity. Foreign ownership does not have any statistically significant impact, but if a plant is a single plant, the skill-intensity is higher.

I also found no statistically significant pre-merger difference in skill-intensity for all subgroups. However, I found a statistically significant decrease in skill-intensity in $t = 2$ of approximately 15% in target plants. Regression also finds evidence for a skill-intensity decrease of approximately 14% in non-horizontally merging plants in $t = 2$. Again, I also apply robustness tests as before, and they point to a proper model specification.

Summarizing results from regression analysis for skill-intensity, the estimations suggest that the included control variables explain the pre-merger differences in skill-intensity between treated and controls that were found in descriptive statistics. However, and in line with descriptive statistics, I find evidence for a statistically significant decrease of skill-intensity in $t = 2$ for targets which confirms the hypothesis of a U-shaped development path. I also find this U-form for the subgroup of non-horizontal M&A, but this result is not fully in line with descriptive statistics which also found a statistically significant decrease in $t = 2$, but no U-form. For the subgroup of horizontal M&A, the U-form that was found in descriptive statistics is not supported by the regression analysis.

4.4.4 Difference-in-differences propensity score matching

The regression analysis performed above is able to show correlations between variables, but it is not a proper method for detecting causality.[36] For this, I address the issue of causality and firm heterogeneity with a matching method in this section. Matching allows the construction of a comparison group that is identical or at least as similar as possible to the treatment group prior to the merger, eliminating the problem of a self-selection of plants with certain characteristics into M&A activity. Differences in size or skill-intensity are then attributed to the merger.

The fundamental evaluation problem: The fundamental evaluation problem that arises in the evaluation of effects of M&A on employment (skill-intensity) is that it is not possible to step back in time and observe how the number of employees (skill-intensity) of the same plant would have developed if it had not merged. Let Y be the variable for employment (skill-intensity), and let $D \in \{0, 1\}$ be an indicator of whether the plant merged or not. Then, $Y_i(1)$ describes the post-merger size (skill-intensity) if plant i merged, and $Y_i(0)$ if the same plant had not merged. The causal effect is then:

$$\Delta = Y_i(1) - Y_i(0). \tag{8}$$

[36] See Angrist and Pischke (2009) for a comprehensive discussion.

However, the outcome $Y_i(0)$ cannot be observed because it is the missing counterfactual, and thus, the observation of the individual treatment effect is not possible. The microeconometric evaluation literature (e.g. Caliendo, 2006; Dehejia and Wahba, 2002; Heckman, Ichimura, and Todd, 1997) defines the (population) average treatment effect (ATT) on the merging plants as

$$ATT = E[Y(1)|D = 1] - E[Y(0)|D = 1]. \tag{9}$$

Again, the second term $E[Y(0)|D = 1]$ can't be observed, because it describes the expected size (skill-intensity) of merging plants had they not merged. However, the causal inference depends on the construction of the counterfactual of this second term. One possibility is $E[Y(0)|D = 0]$, i.e. the expected size (skill-intensity) of control plants, but this is only a good idea if plants randomly assign to the treatment and control groups. In non-experimental data, it seems more realistic that plants self-select in the groups, i.e. that pre-merger characteristics that determine the decision for M&A also influence the plant's post-merger size (skill-intensity) performance. For this reason, it could be misleading if employment (skill-intensity) changes following a merger are interpreted as being caused by this merger.

Since $E[Y(0)|D = 1]$ is not a useful comparison group, a different and valid control group has to be identified. Matching provides a way to construct such a valid control group. This approach, which has been developed by Roy (1951), Rubin (1974), and Heckman, LaLonde, and Smith (1999), pairs treated and control plants which are "statistical twins", i.e. they are similar (ideally identical) to each other in relevant pre-merger characteristics.[37] Because treated and untreated only differ with respect to their treatment status, post-merger differences in size (skill-intensity) can then only be caused by the merger.

The creation of a valid control group implies that both groups have to be similar across a number of different pre-merger characteristics X. But this leads to a dimensionality problem, and therefore, it would be desirable to match plants according to only one single index that includes all information from those variables. Rosenbaum and Rubin (1983) suggest using a propensity score $P(X)$ as a measure of the plant's probability to merge, conditional on observed characteristics X. The ATT can now be estimated as

$$ATT = E[Y(1)|D = 1, P(X)] - E[Y(0)|D = 0, P(X)]. \tag{10}$$

To apply matching, several assumptions have to hold: first, the "conditional independence

[37] "Relevant" means that these plant characteristics influence the decision to merge.

assumption" (CIA) states that - conditioning on the values of a set of observable characteristics X which are not affected by the merger treatment - the number of employees (skill-intensity) in both groups would be the same in the absence of a merger (Lechner, 1999). A second assumption is the "common support condition": it ensures that propensity scores of both groups overlap and all merging plants have a counterpart in the control group. With this assumption, only plants which are sufficiently similar to each other will be matched (Caliendo, 2006). Third, the "Stable Unit-Treatment Value Assumption" (SUTVA) states that a plant's behavior has no impact on that of another plant.

Estimation of the Propensity Score The propensity score $p(X)$ estimates the probability for a plant to merge, based on characteristics X prior to M&A, and it is estimated with a probit model.[38] Economic theory and empirical literature about M&A from Girma and Görg (2007), Margolis (2006a), Harris and Robinson (2002), Conyon, Girma, Thompson, and Wright (2002a), and others provide a guideline for the variable choice. According to these studies, size is identified as a determinant for M&A, as well as legal form, location (West or East Germany), and sector. In addition, I include dummies for different years.[39] Hence, the probit model is specified as

$$P(MA_{it=2}=1) = F(size\,dummies_{it=1}, legal\,form_{it=1},$$
$$location_{it=1},\;\; industry\,dummies_{it=1},\;\; year\,dummies_{it=1}). \quad (11)$$

Table 4.7 presents the results from the probit regression. Focusing on the first regression, size is a determinant for M&A activity. In particular, the coefficients for the dummies for employees between 300 and 499 and for employees of 500 and more are both positive, of relevant magnitude, and statistically highly significant. This means that plants of this size have a higher probability for merging compared to the reference group of plants with less than 20 employees. These findings are in line with existing empirical evidence (e.g. Girma and Görg, 2007), and confirm the hypothesis of a self-selection of larger plants into M&A activity. In addition, plants with legal form "Limited by shares" are more likely to be merging plants compared to plants of

[38] Using a logit model is also possible and yields similar results (Caliendo, 2006).

[39] There are different views about the proper number of explanatory variables in the model: Bryson, Dorsett, and Purdon (2002) recommend a smaller set of variables whereas Rubin and Thomas (1996) argue for a broader and more generous model. I performed the whole matching process with different model specifications including additional explaining variables (e.g. productivity, wages, etc.). However, I obtained the most robust matching results with a reduced number of variables. Nevertheless, it should be noted that the objective of matching is not an exact estimation of the propensity score, but the balancing of relevant variables (Caliendo, 2006).

the reference group "Partnership, individually-owned, public, and others", and plants located in East Germany are also more likely to be involved in M&A. Dummies for sectors and years are also included, but not reported in the table. The results from the probit regression are consistent with the results from descriptive statistics and regression analysis above. In addition, the McFadden's pseudo-R^2-value is acceptable.[40]

Table 4.7: **Probit regression (dependent variable: M&A dummy)**

Variables	All (1)	Acquirers (2)	Targets (3)	Horizontal (4)	Non-horiz. (5)
Employees 20-49 (D)	0.05 (0.27)	0.26 (1.06)	-0.21 (-0.81)	0.22 (0.89)	-0.42 (-1.22)
Employees 50-99 (D)	-0.37 (-1.59)	-0.42 (-1.35)	-0.40 (-1.38)	-0.94*** (-2.58)	-0.65 (-1.63)
Employees 100-299 (D)	0.27 (1.46)	0.49** (2.13)	-0.11 (-0.46)	0.32 (1.39)	-0.58* (-1.76)
Employees 300-499 (D)	0.86*** (3.54)	0.90*** (2.95)	0.75** (2.51)	0.86*** (2.81)	0.85** (2.30)
Employees >=500 (D)	1.71*** (6.92)	1.73*** (5.51)	1.52*** (5.21)	1.80*** (5.87)	0.84** (2.13)
Legal form "Limited" (D)	0.20 (1.18)	0.28 (1.30)	0.01 (0.04)	0.28 (1.20)	-0.05 (-0.16)
Legal form "Limited by shares" (D)	1.22*** (4.35)	1.47*** (4.51)	1.01** (2.55)	1.68*** (4.88)	0.41 (0.58)
Location in East Germany (D)	0.26** (1.98)	0.33** (2.08)	0.12 (0.69)	0.39** (2.32)	0.06 (0.24)
Constant	-1.09*** (-2.88)	-1.60*** (-3.45)	-1.11** (-2.26)	-1.35*** (-2.90)	-0.96 (-1.51)
Observations	699	620	588	597	502
Pseudo-R^2	0.20	0.23	0.20	0.28	0.16

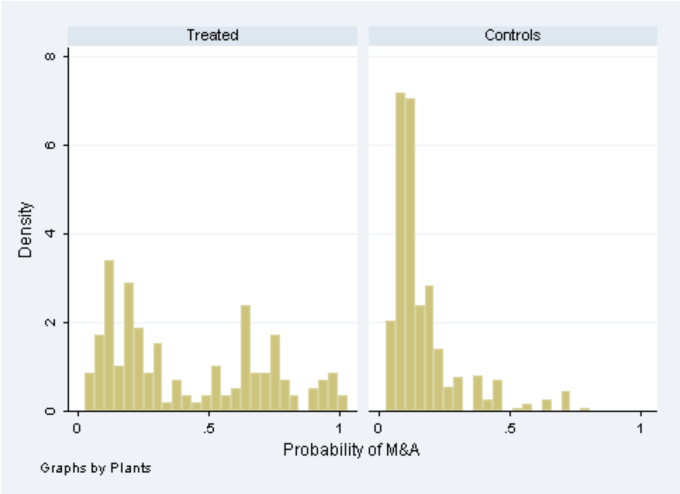
Notes: t-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1. (D) means variable is a dummy. The reference category for employees is "Employees 1-19", and for legal form it is "Partnership, individually-owned, public, and others". Regressions also include dummies for sectors and years. The reduced number of observations is due to missing data for several variables. Data source: IAB Establishment Panel, M&A DATABASE St. Gallen.

In addition to the first regression specification, I also perform probit regressions for all subgroups of treated plants. That is, the respective dependent variables are dummies with value one if the plant is an acquirer - analogous for targets, horizontal, and non-horizontal M&A - and zero if the plant is a control. The results are similar to estimations from the first regression, with few exceptions (e.g. in regression 4, plants with employees between 50 and 99 are significantly less likely to merge horizontally, and regression 5 yields a poor value for pseudo-R^2 which may be due to the low number of treated observations in this group).

Based on this probit regression, for every treated and control plant, a propensity score $P(X)$ is now estimated. Figure 4.1 presents the distribution of propensity scores estimated in the first regression specification which includes all treated plants: the distribution is different between both groups, and for certain ranges of propensity scores, there are no or only few matching partners. If matching partners are not sufficiently similar in terms of propensity scores, the matching quality suffers. This has to be kept in mind when choosing how to pair treated and controls.

[40]Values of 0.2 and above as seen as sufficient (e.g. Backhaus, Erichson, Plinke, and Weiber, 2010).

Figure 4.1: **Distribution of propensity scores for treated and control group**



Matching algorithm: The propensity score is a continuous variable, and thus, it is difficult to find matching pairs with exactly the same score. For this, the matching literature has developed different algorithms about how to assign treated and control plants to each other, and how to weight each of the matching partners.[41] In this paper, I choose a kernel matching algorithm.[42] The kernel algorithm assigns all controls j to each single treated i. That is, for each treated i a neighborhood $C(P_i)$ is defined which contains the whole control group I_0:

$$C(P_i) = \{I_0\}. \tag{12}$$

Controls are assigned with different weights depending on the distance to the treated cases in terms of propensity scores:

$$W_{N_0}^{KM}(i,j) = \frac{G_{ij}}{\sum_{k \in I_0} G_{ik}}, \tag{13}$$

with N_0 denoting the number of controls in the control group I_0, and $G_{ik} = G[(P_i - P_j)/a_{N_0}]$ as a kernel function that downweighs controls j which have a larger distance to the treated i. I use a kernel based on a Gaussian normal function with a bandwidth parameter a_{N_0} of 0.06.[43] In addition, I only match observations which lie in the region of common support, and I also apply

[41] I use the STATA-module PSMATCH2 of Leuven and Sianesi (2003).

[42] Because the smaller the dataset, the more important the choice of the algorithm (Caliendo, 2006). For this reason, I tested several algorithms (nearest neighbor, caliper, radius caliper, and kernel algorithms in different modifications) and found out that kernel yields the most robust results.

[43] A bandwidth parameter impacts the form of the kernel function. The choice of the bandwidth parameter affects the results more strongly (Silverman, 1986; Pagan and Ullah, 1999), whereas the choice of the kernel function (i.e. Gaussian (normal), biweight, epanechnikov, uniform, or tricube kernel) is of minor relevance (DiNardo and Tobias, 2001).

a trimming procedure as suggested by Smith and Todd (2005a): 10% of treated for which the density of controls is the lowest are dropped. This is useful if there are no observations for several regions of the controls' propensity scores, as figure 4.1 shows.

Due to the availability of longitudinal data, I am able to analyze changes instead of levels, i.e. I combine the standard matching approach with a difference-in-differences (DiD) estimator. This estimator measures the difference between the arithmetic mean values of changes in employment and skill-intensity. The combination of the standard approach with a DiD estimator is suggested by Smith and Todd (2005a): even though the combined estimator is still based on the assumption of "selection on observables", it relaxes this strong assumption, because it eliminates all unobserved time-invariant plant characteristics between treated and control plants that the standard matching estimator fails to eliminate. For this, a combined estimator has the potential to improve the quality of the results significantly (Blundell and Costa Dias, 2000).

Under the consideration of the common support condition, the estimator can be implemented as

$$\Delta_{ATT}^{DiD} = \frac{1}{N_1} \sum_{i \in I_1 \cap S_P} \left[(Y_{it}^1 - Y_{it'}^0) - \sum_{j \in I_0 \cap S_P} W(i,j)(Y_{jt}^0 - Y_{jt'}^0) \right], \tag{14}$$

with N_1 denoting the number of treated in the treatment group I_1, and S_P the region of common support. t and t' represent the years of the respective observation periods.

Matching results for employment changes: The results of the matching procedure with respect to changes in employment are shown in table 4.8. The upper part of the table presents the results for the observation period between $t = 2$ to $t = 4$, and the lower part of the table for the period between $t = 1$ and $t = 4$.

Taking the group including all treated ("All"), percentage changes in both groups are negative (first and second column), but the difference in means for the matched sample (ATT) is 0.035, indicating that merging plants exhibit a lower negative employment growth than non-merged plants. However, this effect is statistically not significantly different from zero at any level, which implies that matching does not find any effect of M&A on employment changes for the post-merger period.[44] With respect to subgroups of treated, the results are similar. The ATTs are positive except for the subgroup of horizontal M&A, but they are not statistically significant.

[44]For the test of the statistical significance of the ATT, I apply bootstrapping with 150 replications in order to yield robust standard errors. However, Abadie and Imbens (2008) criticize that no formal justification for the use of bootstrapping methods in the context of matching has been provided. Nevertheless, many empirical studies applying matching used bootstrapping (e.g. Girma, Görg, and Wagner, 2009).

Turning to the observation period between $t = 1$ and $t = 4$, the results are similar, too. For all groups of treated, the estimated ATTs are positive except for the subgroup of horizontal M&A, but they are not statistically significant at any level.

Table 4.8: **ATT for employment changes**

Employment changes between $t = 2$ and $t = 4$:						
Group of treated	Treated	Controls	ATT	S.E.	T-statistic	p-value
All	-0.028	-0.063	0.035	0.031	1.12	0.305
Acquirers	-0.025	-0.047	0.022	0.043	0.51	0.582
Targets	-0.020	-0.071	0.050	0.033	1.51	0.250
Horizontal	-0.021	-0.015	-0.006	0.045	-0.14	0.883
Non-horizontal	-0.034	-0.061	0.026	0.037	0.71	0.587

Employment changes between $t = 1$ and $t = 4$:						
Group of treated	Treated	Controls	ATT	S.E.	T-statistic	p-value
All	-0.038	-0.078	0.040	0.036	1.10	0.290
Acquirers	-0.041	-0.055	0.013	0.049	0.27	0.792
Targets	-0.020	-0.071	0.050	0.033	1.51	0.166
Horizontal	-0.042	-0.021	-0.021	0.051	-0.42	0.692
Non-horizontal	-0.021	-0.074	0.052	0.037	1.41	0.256

Notes: p-values are estimated for bootstrapped standard errors with 150 replications.

Matching results for skill-intensity changes: Table 4.9 presents the results for the estimated ATT with respect to potential merger induced skill-intensity changes. The upper part of the table displays percentage changes for the matched sample for treated and controls between $t = 2$ and $t = 4$ and the respective ATT. The treatment group including all treated estimates both positive changes for treated and controls, and the ATT is 0.010, but it is not statistically significant at any usual level. That is, merging does not affect the plants' skill-intensity. For the subgroups, the ATTs are also not statistically significant, except the negative ATT for the subgroup of non-horizontal M&A.

With respect to the observation period between $t = 1$ and $t = 4$, changes in the treatment groups are positive except for the subgroup of non-horizontal M&A. The estimated ATTs are all negative, and this implies that the skill-intensity of merging plants increases more slowly over time compared to plants that do not merge. But again, this effect is not statistically significant, except for the subgroup of non-horizontal M&A.

Because of the results from descriptive statistics and regression analysis which point to a U-shaped development path of the merging plants' skill-intensity over time for some subgroups, I also analyze changes between $t = 1$ and $t = 2$. The results are displayed in the lower part of table 4.9. Changes are negative for all treatment groups except for acquirers, and the same is true for the respective estimated ATTs. For the subgroup of targets and non-horizontal M&A, the ATTs are statistically significant at usual significance levels. Targets exhibit a decrease in their workforce's skill-intensity of approximately 7.5% immediately around the merger, and this change is around 10.5% lower compared to plants that do not merge. For non-horizontally merging plants, the difference is around 11.8%. For targets, the results are fully in line with findings from descriptive statistics and regression analysis. For the subgroup of non-horizontal M&A the results are in line with regression analysis.

Table 4.9: **ATT for skill-intensity changes**

Skill-intensity changes between $t = 2$ and $t = 4$:						
Group of treated	Treated	Controls	ATT	S.E.	T-statistic	p-value
All	0.044	0.034	0.010	0.045	0.23	0.795
Acquirers	-0.002	0.037	-0.038	0.047	-0.82	0.431
Targets	0.098	0.045	0.053	0.065	0.80	0.393
Horizontal	0.097	0.052	0.045	0.056	0.80	0.363
Non-horizontal	-0.068	0.071	-0.139	0.072	-1.93	0.092
Skill-intensity changes between $t = 1$ and $t = 4$:						
Group of treated	Treated	Controls	ATT	S.E.	T-statistic	p-value
All	0.041	0.066	-0.025	0.050	-0.51	0.580
Acquirers	0.039	0.071	-0.033	0.059	-0.55	0.625
Targets	0.023	0.072	-0.048	0.065	-0.75	0.462
Horizontal	0.065	0.086	-0.021	0.061	-0.34	0.738
Non-horizontal	-0.133	0.123	-0.256	0.078	-3.26	0.001
Skill-intensity changes between $t = 1$ and $t = 2$:						
Group of treated	Treated	Controls	ATT	S.E.	T-statistic	p-value
All	-0.004	0.033	-0.037	0.043	-0.85	0.378
Acquirers	0.041	0.035	0.005	0.049	0.11	0.919
Targets	-0.074	0.030	-0.105	0.057	-1.84	0.077
Horizontal	-0.032	0.035	-0.067	0.050	-1.34	0.167
Non-horizontal	-0.064	0.053	-0.118	0.044	-2.69	0.017

Notes: p-values are estimated for bootstrapped standard errors with 150 replications.

Robustness tests: In order to assess the reliability of the results, I check the quality of the matching procedure. First, the balancing property should be satisfied. That is, the distribution of variables should be balanced and no statistically significant differences in the variables between both groups should remain. Rosenbaum and Rubin (1985) suggest the calculation of a standardized bias. This is an indicator to assess the distance in marginal distributions of the variables: for each variable the standardized bias calculates the difference of sample means in the subsamples of treated and matched controls as a percentage of the square root of the average of sample variances in both groups (Caliendo, 2006):

$$SB = 100\frac{(\bar{X}_1 - \bar{X}_0)}{\sqrt{0.5[V_1(X) + V_0(X)]}}. \tag{15}$$

After matching, the standardized bias is given by

$$SB^M = 100\frac{(\bar{X}_1^M - \bar{X}_0^M)}{\sqrt{0.5[V_1^M(X) + V_0^M(X)]}}. \tag{16}$$

$\bar{X}_1$ and $\bar{X}_0$ are the mean values for a specific variable of the treatment and control group, and V_1 and V_0 are the respective variances. $\bar{X}_1^M$, $\bar{X}_0^M$, V_1^M, and V_0^M are the corresponding values after matching.

The results for the mean standardized bias (MSB), i.e. the average standardized bias for all covariates, are reported in table 4.10. For the first matching procedure including all treated the MSB is 8. This is acceptable according to Rosenbaum and Rubin (1985) who argued that a bias reduction below 20 is sufficient. However, Caliendo (2006) pointed out that in most empirical studies a bias reduction below 3 or 5% is seen as sufficient, and this is not achieved here. The subgroups yield similar balancing results with values for MSB ranging between 6 and 9. Even if values of MSB are higher than recommended by Caliendo (2006), I consider the balancing as acceptable.

The same table also reports about the region of common support, which can be seen as an indicator for the representativeness of the matching results. Treated observations with propensity scores higher than the maximum or lower than the minimum propensity score of the control plants are excluded, and they are not considered for the estimations any more. If the fraction of observations lying off support is large (e.g. up to one fourth in the subgroup of horizontal M&A), the generality of the results is reduced, and the interpretation of estimations must be restricted only to those plants that are matched.[45]

There are further possibilities to test the matching quality. Sianesi (2004) recommends performing the probit regression and calculation of propensity scores again, but now based on the matched sample. If the matching was successful, there should be no differences in the covariates, and the pseudo-R^2 should be low. Moreover, a reestimation of the likelihood ratio test should generate low values and a p-value close to one, i.e. the independent variables in

[45]See Lechner (2008) for a discussion about common support.

Table 4.10: **Robustness tests**

Mean standardized bias (MSB)			
Group of treated	Mean/Std. Err.	Before matching	After matching
All	Mean	19.592	8.202
	Std. Dev.	17.375	6.613
Acquirers	Mean	20.614	7.271
	Std. Dev.	19.600	6.355
Targets	Mean	21.317	6.770
	Std. Dev.	18.541	4.608
Horizontal	Mean	23.375	9.540
	Std. Dev.	20.823	6.283
Non-horizontal	Mean	20.912	7.177
	Std. Dev.	18.014	5.480

Common Support				
Group of treated	Sample	Off support	On support	Total
All	Untreated	0	540	540
	Treated	21	138	159
Acquirers	Untreated	0	533	533
	Treated	17	70	87
Targets	Untreated	0	516	516
	Treated	15	57	72
Horizontal	Untreated	0	505	505
	Treated	24	68	92
Non-horizontal	Untreated	0	476	476
	Treated	2	24	26

Pseudo-R^2 and log likelihood test				
Group of treated	Sample	Pseudo-R2	LR chi2	p>chi2
All	Unmatched	0.205	153.34	0.000
	Matched	0.028	10.61	0.910
Acquirers	Unmatched	0.230	115.71	0.000
	Matched	0.028	5.41	0.996
Targets	Unmatched	0.196	85.70	0.000
	Matched	0.024	3.79	1.000
Horizontal	Unmatched	0.279	143.09	0.000
	Matched	0.041	7.79	0.971
Non-horizontal	Unmatched	0.158	32.26	0.006
	Matched	0.038	2.53	1.000

the model have no explanatory power. These robustness tests yield mostly satisfying results, as the lower part of table 4.10 shows.

Summarizing the matching process with respect to employment changes, I confirm the results from descriptive statistics and regression analysis, i.e. mergers do not statistically significantly change a plant's employment in a positive or negative way. Differentiating between acquirers and targets, or between plants that merge horizontally and non-horizontally does not change the results. The results from matching with respect to skill-intensity changes also corroborate earlier findings in this paper: estimates point to a U-shaped development path of the workforce's skill-intensity in targets. For the subgroup of non-horizontal M&A, the matching results are also in line with findings from regression analysis, providing evidence for a U-shaped skill-intensity growth. However, the robustness of this subgroup is limited due to the small number of only 28 observations. Matching does not identify effects for the other treatment groups.

4.5 Conclusion - what do we learn?

In this paper, I use a new dataset which is a combination of the IAB Establishment Panel and the M&A DATABASE. It includes plants that merged between 1995 and 2005 and control plants, and I analyze self-selection of plants into M&A activity and merger induced effects on employment and skill-intensity. I take an observation period of four years. Theory does not give clear predictions, and the results from existing empirical literature are ambiguous. The data allow for differentiation between subsamples of acquirers and targets, horizontal and non-horizontal mergers. I choose a three-step estimation strategy with descriptive statistics, regression analysis, and a difference-in-differences propensity score matching.

I find evidence for a self-selection of larger plants into merger activity for all groups of treated. With respect to acquirers this supports the argument that larger firms have better opportunities to merge due to lower credit constraints and a better equity basis. Moreover, an explanation why targets are large may be that acquirers seek for market power increase, assuming that a plant's market power increases with its size. However, the findings of a higher pre-merger size contradict Mueller (2003a) who stated that in the presence of scale economies mostly smaller firms merge horizontally. With respect to plants' pre-merger skill-intensity, I found no statistically significant difference between treated and untreated when I controlled for other pre-merger differences.

According to the results from this study, I do not estimate statistically significant employment effects due to M&A for any treatment group, independent of the underlying estimation method. In other words: employment neither increases nor decreases due to M&A in any statistically significant way. This means that estimations do not support the widely held fear of employment losses after mergers. Moreover, these results are in line with estimations from existing studies, e.g. from Arndt and Mattes (2010) and Mattes (2010), but nevertheless, contradict the majority of studies for Europe which mostly find negative employment effects (e.g.

Gugler and Yurtoglu, 2004). As one strand of theory suggests, employment losses are more likely if mergers occur for profit maximizing reasons (Conyon, Girma, Thompson, and Wright, 2002a). In this case, one can conclude from my findings that a certain number of plants also merges for other, non-profit maximizing reasons, e.g. for empire building. However, unchanged employment is also possible in a neoclassical framework, because the results may be evidence that M&A are a capital investment, or improve the plants' financial possibilities. In addition, another difference in comparison to other studies lies in the fact that the dataset only includes domestic mergers. Since foreign acquirers may be less committed to fulfill implicit contracts, foreign M&A may lead to greater job losses than domestic M&A (Lehto, 2006).

I also do not find any effects of M&A on a plant's skill-intensity with two exceptions. First, I found robust evidence for a U-shaped skill-intensity development path over time for targets. In other words: considering the whole observation period, findings from this paper imply that the skill-intensity of targets statistically significantly decreases immediately around the merger, but increases again, reaching a level that does not differ in any statistically significant way from the targets' pre-merger level. The interpretation of the estimated results is not easy, since skill-intensity changes cannot be due to changes in employment (which is held constant), but must be due to shifts between skilled and unskilled workers. One explanation may be that mostly targets are affected by merger induced organizational changes including job rotations or restructuring of departments. Nevertheless, these result are in accordance with Lipsey and Sjöholm (2003) who found an increase in the number of blue-collar workers and a reduction of the number of white-collar workers after M&A. Second, I also found statistically significant effects on the skill-intensity in non-horizontally merging plants. However, I no not consider the results as robust enough because the estimates from the different methods are not fully in line with each other. Moreover, the number of observations in this subgroup is small, reducing the reliability of results.

C Remarks on the dataset

For the empirical analysis of this paper I use a combined dataset from the IAB Establishment Panel and the M&A DATABASE from St. Gallen, performed by TNS Infratest Sozialforschung GmbH München. Two methodology reports exist about the creation of the treatment group and the control group.[46] Prior to the creation of this dataset, a pilot study from Bellmann and Kirchhof (2006) showed that the IAB Establishment Panel is capable of an analysis of effects of mergers. The focus of the pilot study was on the effects on employment, and data about M&A was generated from the dataset Thomson ONE Banker. However, the dataset M&A DATABASE from St. Gallen is more comprehensive in comparison to Thomson ONE Banker because it also includes small and medium-size firms, whereas Thomson ONE Banker only includes firms of a larger size. Moreover, for each deal the M&A DATABASE also lists the seller[47] (Thomson ONE Banker only reports acquirers and targets) as well as additional information like location, number of employees, sales, etc.

The treatment group: The observation period for plants covers the time from January 1996 to December 2005. The creation of the treatment group was carried out in several steps by TNS Infratest Sozialforschung. First, in order to combine both datasets, companies in the M&A DATABASE were compared to plants in the IAB Establishment Panel and classified according to the degree of similarity with respect to name, location, and sector. Next, merged plants were only kept if they were surveyed at least once before and once after the merger. This led to some complications due to the set up of the survey: information about employees refers to June 30th of the respective year, whereas information about sales, investments, etc. refer to the previous year. To take these circumstances into account, the following definition was chosen:

- If M&A was between January 1st and June 30th of year T, the survey in year T was considered to be conducted after M&A, even if some information refers to a point of time before M&A.
- If M&A was between July 1st and December 31th of year T, the survey in year T was considered to be conducted before M&A, even if some information refers to a point of

[46] TNS Infratest: Beschäftigungseffekte von Fusionen und Übernahmen - Methodenbericht Untersuchungsgruppe (März 2007); Beschäftigungseffekte von Fusionen und Übernahmen - Methodenbericht Untersuchungsgruppe (December 2007).

[47] Nevertheless, this study does not analyze the effects on sellers because the number of observations is too small.

time after M&A.

This restriction reduced the number of plants which were found in both datasets to 7,801. According to the degree of similarity of plants in both datasets, observations are distributed across four different categories:

- Quality class 1: name, location, and sector match exactly (1,426);
- Quality class 2: name, location, and superior sector match exactly (146);
- Quality class 3: name and sector match exactly; multi-plant firm (5,961);
- Quality class 4: name and location match exactly (268).

These 7,801 merger cases consist of 958 different plants in the IAB Establishment Panel. This is because one plant may be involved in several M&A within the observation period.

The control group: Next, a group of control plants that had not merged between 1980 and 2005 has to be found. These controls must be as similar as possible to plants in the treatment group. Each of the 7,801 treatment observations exhibits an individual combination of sector, size, legal form, and location in West or East Germany. Therefore, TNS Infratest Sozialforschung defined 2,143 categories which differ with respect to these characteristics, and each of the 7,801 treatment observations was assigned to one of these 2,143 categories. Now, the objective was to find controls for each category. This is, within a category, treated and controls are homogeneous with respect to the characteristics. An example: there are three treated plants which all belong to the agriculture sector, have less than 10 employees, have "GmbH" as legal form, and are located in West Germany. The combination of these specific characteristics constitutes one of the 2,143 categories. After that, three control plants should also be identified which exhibit the same characteristics of this specific category.

30,110 plants from the IAB Establishment Panel were identified as statistical twins to treated plants (with respect to sector, size, legal form, and location in West or East Germany), and they may potentially act as a control. Within these 30,110 observations, several plants appear more often if they were surveyed for the IAB Establishment Panel for several years. The challenge is to identify those "true" controls within the 30,110 potential controls, that is, plants for which we can be sure that they have not been involved in any M&A activity. Hence, plants which already appeared in the treatment group were excluded, and plants which were similar to plants from the M&A DATABASE, but which were not in the treatment group, because they merged outside the reference period between 1996 and 2005. This step excludes 1,204 from the 30,110 observations. Then, plants which were not surveyed at least twice between 1993, the starting year of the IAB Establishment Panel and 2006,[48] the end of the observation period, were dropped.

The rest of the remaining 27,676 potential controls had to be checked manually: as a first check, plants were eliminated if their name appeared in the M&A DATABASE. In addition, the dataset "Markus" from Bureau van Dijk also provided information about M&A activities of plants. Finally, for most plants websites were used as a source of information about merger activity.

As stated above, the number of treated should equal the number of controls within each of the 2,143 categories. Hence, for each category, potential controls were checked for whether they were "true" controls. This was repeated until the number of true controls equaled the number of treated, and the remaining potential controls for the respective category were no longer considered. However, for several categories no controls could have been found, because potential controls have all merged.[49] Figure C1 presents a graphical illustration of this process.

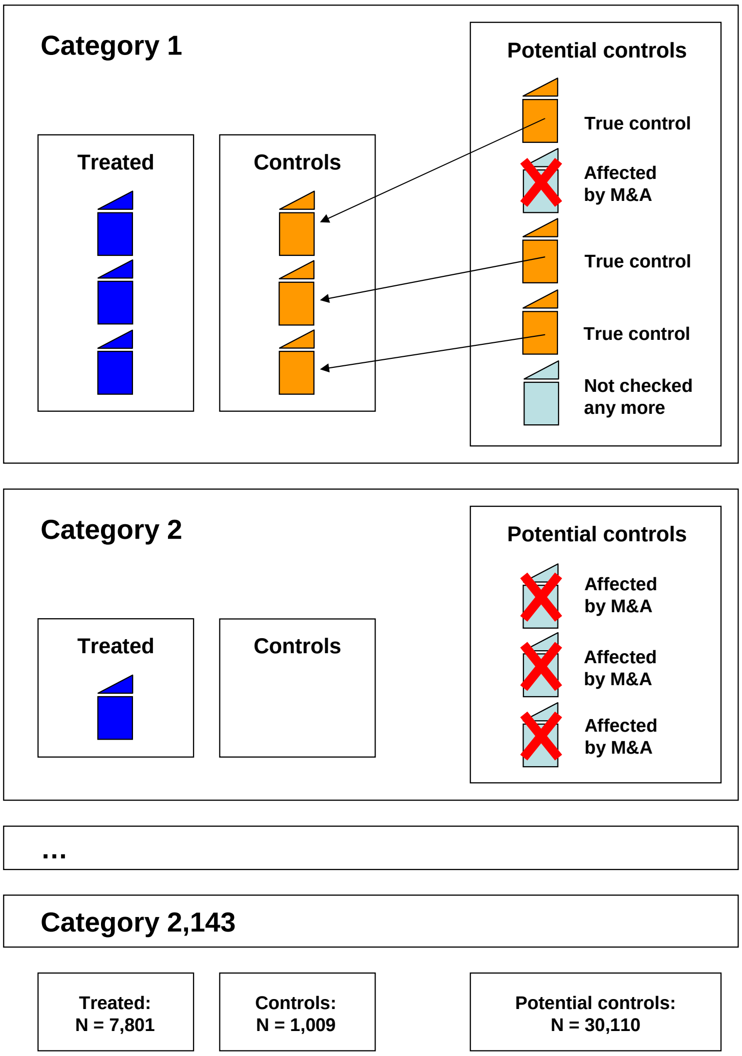
In total, 12,755 plants were checked in 400 hours of research by TNS Infratest. As a result, 1,009 controls from 291 different plants appear in the control group. This is because a plant can act as a control over several years and for different categories as well. The structure of the control group is similar to the structure of the treated group with respect to sector, size, legal form, and location, but controls were not involved in any merger activity during the reference period.

A note on sector classification: The 2-digit sector classification of the IAB Establishment Panel follows the NACE code (Nomenclature Générale des Activités Économiques). The NACE code changed in 1999 and in 2003, leading to different classifications of plants over time. Whereas the classification change in 2003 is not a problem, the break in 1999 is more severe. Due to this, I transfered the sector classification of the year 2000 to the years before in order to achieve a consistent sector classification of plants. However, this leads to a drawback if plants changed sectors due to a merger: they may not be classified correctly. I checked this aspect manually and could not identify incorrect classifications.

[48] The observation period for treated plants ends in December 2005. To gain information about controls for the year 2005, the 2006 survey is relevant, because several questions refer to the year before, e.g. sales.

[49] For example, almost no controls were found in the financial sector.

Figure C1: **Graphical illustration of the creation of the control group by TNS Infratest**

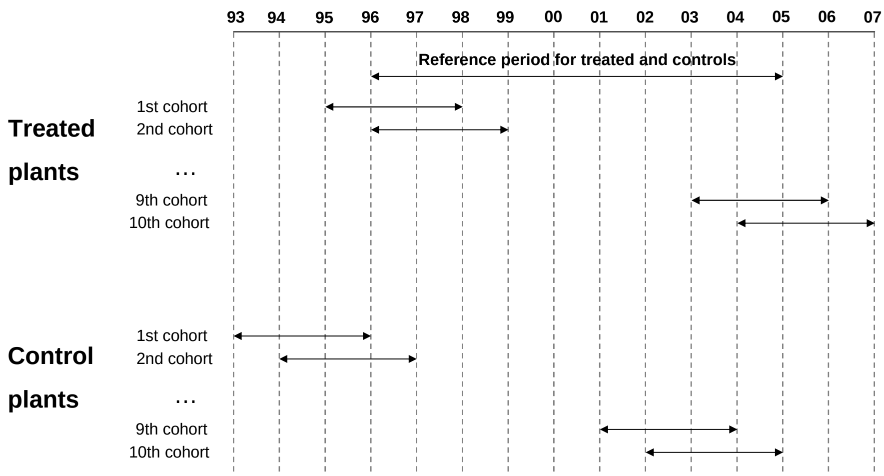


Note: the large boxes represent different categories. In each of the 2,134 categories, plants are homogeneous with respect to sector, size classification, location in West or East Germany and legal form. The number of controls should equal the number of treated in each category. Therefore, 30,110 potential controls from the IAB Establishment Panel, which represent statistical twins to the treated, were assigned to these different categories. Then, each potential control was checked for whether it had merged since 1980. If this was true, the respective plant was identified as a "true" control. If, however, a potential control had merged since 1980, it was discarded. If the required number of controls was found, remaining controls were not considered anymore (category 1). It could also be that no controls were found for a certain category (category 2). Finally, 1,009 controls were identified.

D The observation period

For the empirical investigation in this paper, I create a "rolling observation window" which is illustrated by figure D1. For treated plants, there are ten four-year windows. The first covers the years between 1995 and 1998, and the last between 2004 and 2007. Similarly, for controls the first four-year window is from 1993 to 1996, and the last is from 2002 to 2005.

Figure D1: **Graphical illustration of cohorts**



Note: for the first cohort in the treatment group, year $t = 1$ corresponds to 1995, $t = 2$ to 1996, $t = 3$ to 1997, and $t = 4$ to 1998. This applies to all cohorts in both groups analogously.

Concluding remarks

This thesis is about the economics of mergers and acquisitions. It starts with a survey about the literature of M&A and performance effects. Then, it presents propensity score matching as a newer econometric method and its implementation in the computer software STATA. The heart of the thesis is about the effects of M&A on plants' labor productivity, employment, and skill-intensity, and about self-selection of plants into merger activity. It uses a new dataset about German plants, and places the focus of the econometric analysis on a propensity score matching approach. Each of the four chapters closes with concluding remarks, but it is worth to draw some final conclusions from the whole thesis.

First, as discussed in the first chapter, the majority of newer studies found positive productivity effects and negative employment effects from M&A, or they did not find any changes. I mostly confirm these trends, because I also find a positive productivity effect for acquirers, even if the evidence is weak. With respect to employment effects, I do not identify causal effects. Nevertheless, the results of my thesis also suggest that plants merge not only for profit maximizing reasons: if all mergers were motivated by profit maximizing reasons, the positive productivity effect for acquirers in my analysis should have been stronger, and other subgroups should also have been affected by these changes. Additionally, I should also have estimated negative employment effects.

Second, the phenomenon "M&A" is complex and requires a differentiated analysis. Research has to differ between the types of firms involved in the mergers (acquirers and targets), or between the types of mergers (e.g. horizontal and non-horizontal). This is also true for the distinction between domestic and cross-border, or friendly and hostile. Any researcher that ignores these crucial differences runs the risk of biased or misleading conclusions. For example, in chapter 3 I identify a causal effect of M&A on plants' labor productivity, but only a differentiation between different types of plants shows that the causality only holds for the subgroup of acquirers.

Third, we learned that plant heterogeneity is an important issue when assessing the effects of mergers. The arguments from Melitz (2003), who showed that only more productive firms can bear the extra costs of exporting, and thus, self-select in export activities, also apply to questions about merger activity. In my thesis, I confirm these arguments and show that plants that merge are more productive, larger, have a higher skilled workforce, etc. Hence, I assume that the results from earlier studies about effects of M&A that did not control for self-selection are seriously biased, and therefore, I suspect that ignoring firm heterogeneity is a major reason for the different results between earlier and newer studies.

And fourth, some lessons can be learned from the estimation strategy I used in this thesis. As usual, I started with descriptive statistics. Then, I performed a classical regression analysis, and finally, I applied a matching approach. Most findings are in line with each other, but some estimation results differ between the methods. This makes the presentation of a consistent picture of empirical findings more difficult, but I consider this strategy as more reliable. Advanced econometric methods like propensity score matching improved empirical work, but they are not per se a guarantee to get unassailable results. This is, for example, because details of the methods like propensity score matching are not yet standardized, yielding different estimations from the same data. Hence, starting with a classical regression analysis in addition to other newer methods, as also suggested by Angrist and Pischke (2009), makes the empirical strategy more reliable.

Bibliography

Abadie, A./ Drukker, D./ Herr, J./ Imbens, G. W. (2004): Implementing Matching Estimators for Average Treatment Effects in STATA, The Stata Journal, 4(3), 290-311.

Abadie, A./ Imbens, G. W. (2008): On the Failure of the Bootstrap for Matching Estimators, Econometrica, 76(6), 1537-1557.

Abadie, A./ Imbens, G. W. (2011): Bias-Corrected Matching Estimators for Average Treatment Effects, Journal of Business and Economic Statistics, 29(1), 1-11.

Abraham, K. G./ Houseman, S. N. (1993): Job Security and Work Force Adjustment: How Different are US and Japanese Practices? in Buechtemann, C. F. (Ed.): Employment Security and Labor Market Behavior: Interdisciplinary Approaches and International Evidence, New York, 180-199.

Abraham, K. G./ Houseman, S. N. (1995): Labour Adjustment Under Different Institutional Structures: A Case Study of Germany and the United States, in Buttler, F./ Franz, W./ Schettkat, R./ Soskice, D. (Eds.): Institutional Frameworks and Labor Market Performance: Comparative Views on the U.S. and German Economies, London, 285-315.

Aldrich, J./ Nelson, F. (1984): Linear Probability, Logit, and Probit Models, Beverly Hills & London.

Almeida, R. (2003): The Effects of Foreign Owned Firms on the Labour Market, IZA Discussion Paper 785.

Amess, K./ Wright, M. (2007): The Wage and Employment Effects of Leveraged Buyouts in the UK, International Journal of the Economics of Business, 14(2), 179-195.

Andrade, G./ Mitchell, M./ Stafford, E. (2001): New Evidence and Perspectives on Mergers, Journal of Economic Perspectives, 15(2), 103-120.

Andrade, G./ Stafford, E. (2004): Investigating the Economic Role of Mergers, Journal of Corporate Finance, 10(1), 1-36.

Andrews, M./ Bellmann, L./ Schank, T., Upward, R. (2009): The Takeover and Selection Effects of Foreign Ownership in Germany: An Analysis Using Linked Worker-Firm Data, Review of World Economics, 145(2), 293-317.

Angrist, J. D./ Pischke, J.-S. (2009): Mostly Harmless Econometrics, Princeton, New Jersey.

Angrist, J. D./ Pischke, J.-S. (2010): The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con Out of Econometrics, Journal of Economic Perspectives, 24(2), 3-30.

Arndt, C./ Mattes, A. (2010): Cross-Border Mergers and Acquisitions of Multinational Firms. New Firm-Level Evidence, IAW Discussion Paper 62.

Arnold, J. M./ Javorcik, B. S. (2009): Gifted Kids or Pushy Parents? Foreign Direct Investment and Plant Productivity in Indonesia, Journal of International Economics, 79(1), 42-53.

Ashenfelter, O. (1978): Estimating the Effects of Training Programs on Earnings, Review of Economics and Statistics, 60(1), 47-57.

Audretsch, D. B./ Elston, J. A. (2002): Does Firm Size Matter? Evidence on the Impact of Liquidity Constraints on Firm Investment Behavior in Germany, International Journal of Industrial Organization, 20(1), 1-17.

Augurzky, B./ Schmidt, C. (2001): The Propensity Score: A Means to an End, IZA Discussion Paper 271.

Austin, P. C. (2007): A Critical Appraisal of Propensity-Score Matching in the Medical Literature Between 1996 and 2003, Statistics in Medicine, 27(12), 2037-2049.

Backhaus, K./ Erichson, B./ Plinke, W./ Weiber, R. (2010): Multivariate Analysemethoden, 13. Aufl., Heidelberg.

Baldwin, J. R. (1998): The Dynamics of Industrial Competition, Cambridge.

Baldwin, J. R./ Gorecki, P. (1990): Mergers Placed in the Context of Firm Turnover, in Bureau of the Census: 1990 Annual Research Conference Porceedings, Washington, DC: US Department of Commerce, 53-73.

Balsvik, R./ Haller, S. A. (2011): Picking "Lemons" or Picking "Cherries"? Domestic and Foreign Acquisitions in Norwegian Manufacturing, Scandinavian Journal of Economics, 112(2), 361-387.

Bandick, R. (2011): Foreign Acquisition, Wages and Productivity, The World Economy, 34(6), 931-951.

Barba Navaretti, G./ Castellani, D./ Disdier, A.-C. (2010): How Does Investing in Cheap Labour Countries Affect Performance at Home? Firm-Level Evidence from France and Italy, Oxford Economic Papers, 62(2), 234-260.

Barba Navaretti, G./ Venables, A. J. (2004): Multinational Firms in the World Economy, Princeton.

Bauer, T. K. (2003): Flexible Workplace Practices and Labor Productivity, IZA Discussion Paper 700.

Baum, Ch. F./ Schaffer, M./ Stillman, S. (2003): Instrumental Variables and GMM: Estimation and Testing, The Stata Journal, 3(1), 1-31.

Baumol, W. J. (1959): Business Behavior, Value, and Growth, New York.

Baye, M. R./ Crocker, K. J./ Ju, J. (1996): Divisionalization, Franchising, and Divestiture Incentives in Oligopoly, American Economic Review, 86(1), 223-236.

Beck, Th./ Demirguc-Kunt, A. (2006): Small and Medium-Size Enterprises: Access to Finance as a Growth Constraint, Journal of Banking & Finance, 30(11), 2931-2943.

Becker, S. O./ Ichino, A. (2002): Estimation of Average Treatment Effects Based on Propensity Scores, The Stata Journal, 2(4), 358-377.

Beckmann, M. (2000): Unternehmenspolitik, Managerkontrolle und Personalabbau in Deutschland, Mitteilungen aus der Arbeitsmarkt- und Berufsforschung, 33(4), 594-608.

Bellak, C./ Pfaffermayr, M. (2002): Why Foreign-Owned Firms are Different: A Conceptual Framework and Empirical Evidence for Austria, in Jungnickel, R. (Ed.): Foreign-Owned Firms - Are They Different? Palgrave, Houndsmill, Basingstoke, 13-57.

Bellak, C./ Pfaffermayr, M./ Wild, M. (2006): Firm Performance After Ownership Change: A Matching Estimator Approach, Applied Economics Quarterly, 52(1), 29-54.

Bellmann, L./ Kirchhof, K. (2006): Akquisitionen und Unternehmenszusammenschlüsse im IAB-Betriebspanel, in Bellmann, L./ Wagner, J. (Hrsg.): Betriebsdemographie. Beiträge zur Arbeitsmarkt- und Berufsforschung, 305, Nürnberg, 191-203.

Bellmann, L./ Kohaut, S. (1999): Betriebliche Beschäftigungsentwicklung und Innovationsaktivitäten - Ergebnisse aus dem IAB-Betriebspanel 1998, in Mitteilungen aus der Arbeitsmarkt- und Berufsforschung, 32(4), 416-422.

Bellmann, L./ Pahnke, A. (2006): Auswirkungen organisatorischen Wandels auf die betriebliche Arbeitsnachfrage, Zeitschrift für Arbeitsmarktforschung, 39(2), 201-233.

Bernard, A. B./ Jensen, J. B./ Redding, S. J./ Schott, P. K. (2007): Firms in International Trade, Journal of Economic Perspectives, 21(3), 105-130.

Bertrand, O./ Zitouna, H. (2008): Domestic Versus Cross-Border Acquisitions: Which Impact on the Target Firms' Performance?, Applied Economics, Taylor and Francis Journals, 40(17), 2221-2238.

Bhagat, S./ Shleifer, A./ Vishny, R. (1990): Hostile Takeovers in the 1980s: The Returns to Corporate Specialisation, Brookings Papers on Economic Activity, Microeconomics, 1-72.

Bhuyan, S. (2002): Impact of Vertical Mergers on Industry Profitability: An Empirical Evaluation, Review of Industrial Organization, 20(1), 61-79.

Black, B. S. (2000): The First International Merger Wave (and the Fifth and Last U.S. Wave), University of Miami Law Review, 54, 799-818.

Blackwell, M./ Iacus, S./ King, G./ Porro, G. (2009): cem: Coarsened Exact Matching in STATA, Stata Journal, 9(4), 524-546.

Blundell, R./ Costa Dias, M. (2000): Evaluation Methods for Non-Experimental Data, Fiscal Studies, 21(4), 427-468.

Brand, J. E./ Halaby, Ch. N. (2006): Regression and Matching Estimates of the Effects of Elite College Attendance on Educational and Career Achievement, Social Science Research, 35, 749-770.

Brealey, R. A./ Myers, S. C./ Allen, F. (2008): Principles of Corporate Finance, 9th ed., Boston.

Breinlich, H. (2008): Trade Liberalization and Industrial Restructuring Through Mergers and Acquisitions, Journal of International Economics, 76(2), 254-266.

Brown, C./ Medoff, J. L. (1988): The Impact of Foreign Acquisition on Labor, in Auerbach, A. J. (Ed.): Corporate Takeovers: Causes and Consequences, Chicago, 9-25.

Brownstone, D./ Valletta, R. (2001): The Bootstrap and Multiple Imputations: Harnessing Increased Computing Power for Improved Statistical Tests, Journal of Economic Perspectives, 15(4), 129-141.

Bryson, A./ Dorsett, R./ Purdon, S. (2002): The Use of Propensity Score Matching in the Evaluation of Labour Market Policies, Working Paper 4, Department of Work and Pension.

Bühner, R. (2002): Fusionen aus betriebswirtschaftlicher Sicht, in Franz, W./ Ramser, H. J./ Stadler, M. (Hrsg): Fusionen, Wirtschaftswissenschaftliches Seminar Ottobeuren, 53-65.

Cable, J. R./ Palfrey, J. P. R./ Runge, J. W. (1980): Federal Republic of Germany, 1964-1974, in Mueller, D. C. (Ed.): The Determinants and Effects of Mergers: An International Comparison, Cambridge, 99-132.

Caliendo, M. (2006): Microeconometric Evaluation of Labour Market Policies, Berlin, Heidelberg.

Caliendo, M./ Hujer, R. (2006): The Microeconometric Estimation of Treatment Effects - An Overview, Journal of the German Statistical Society, 90(1), 197-212.

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

Capron, L./ Mitchell, W. (1998): The Role of Acquisitions in Reshaping Business Capabilities in the International Telecommunications Industry, Industrial and Corporate Change, 7(4), 715-730.

Carlton, D. W. / Perloff, J. M. (2005): Modern Industrial Organization, 4th ed., Boston.

Cartwright, S./ Cooper, C. L. (1992): Mergers and Acquisitions: The Human Factor, Oxford.

Castellani, D./ Zanfei, A. (2004): "Cherry-Picking" and Self-Selection, Applied Economics Quarterly, 50(1), 5-20.

Caves, R. E. (1989): Mergers, Takeovers and Economic Efficiency: Foresight vs. Hindsight, International Journal of Industrial Organization, 7(1), 151-174.

Caves, R. E. (1996): Multinational Enterprise and Economic Analysis, 2nd ed., Cambridge.

Caves, R. E./ Barton, D. R. (1990): Efficiency in U.S. Manufacturing Industries, Cambridge.

Church, J. (2004): The Impact of Vertical and Conglomerate Mergers on Competition, Final Report to the European Commission.
http://ec.europa.eu/competition/mergers/studies_reports/merger_impact.pdf
March 31th 2011, 10.22h.

Church, J. (2008a): Conglomerate Mergers, Issues in Competition Law and Policy, 2, American Bar Association, 1503-1552.

Church, J. (2008b): Vertical Mergers, Issues in Competition Law and Policy, 2, American Bar Association, 1455-1502.

Cochran, W. G./ Rubin, D. B. (1973): Controlling Bias in Observational Studies: A Review, Sankhya, Series A 35(4), 417-446.

Comanor, W. S. (1967): Vertical Mergers, Market Powers, and the Antitrust Laws, American Economic Review, 57, 254-265.

Conn, C./ Cosh, A./ Guest, P./ Hughes, A. (2001): Long-Run Share Performance of UK Firms Engaging in Cross-Border Acquisitions, ESRC Centre for Business Research - Working Papers 214, ESRC Centre for Business Research.

Conyon, M. J./ Girma, S./ Thompson, S./ Wright, P. (2001): Do Hostile Mergers Destroy Jobs? Journal of Economic Behavior and Organization, 45(4), 427-440.

Conyon, M. J./ Girma, S./ Thompson, S./ Wright, P. (2002a): The Impact of Mergers and Acquisitions on Company Employment in the United Kingdom, European Economic Review, 46(1), 31-49.

Conyon, M. J./ Girma, S./ Thompson, S./ Wright, P. (2002b): The Productivity and Wage Effects of Foreign Acquisition in the United Kingdom, The Journal of Industrial Economics, 50(1), 85-102.

Conyon, M. J./ Girma, S./ Thompson, S./ Wright, P. (2004): Do Wages Rise or Fall Following Merger?, Oxford Bulletin of Economics and Statistics, 66(5), 847-862.

Cosh, A./ Hughes, A./ Singh, A. (1980): The Causes and Effects of Takeovers in the United Kingdom, in Mueller, D. C. (Ed.): The Determinants and Effects of Mergers, Cambridge, 227-270.

Davies, R./ Kim, S. (2004): Matching and the Estimated Impact of Interlisting, Working Paper, University of Reading.

Dehejia, R. H. (2005): Program Evaluation as a Decision Problem, Journal of Econometrics, 125(1-2), 141-173.

Dehejia, R. H./ Wahba, S. (1999): Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs, Journal of the American Statistical Association, 94(448), 1053-1062.

Dehejia, R. H./ Wahba, S. (2002): Propensity Score Matching Methods for Nonexperimental Causal Studies, Review of Economics and Statistics, 84(1), 151-161.

Deneckere, R./ Davidson, C. (1985): Incentives to Form Coalitions with Bertrand Competition, RAND Journal of Economics, 16(4), 473-486.

Dewey, D. (1961): Mergers and Cartels: Some Reservations about Policy, American Economic Review, 51(2), 255-262.

DiNardo, J./ Tobias, J. L. (2001): Nonparametric Density and Regression Estimation, Journal of Economic Perspectives, 15(4), 11-28.

DiPrete, T./ Gangl, M. (2004): Assessing Bias in the Estimation of Causal Effects: Rosenbaum Bounds on Matching Estimators and Instrumental Variables Estimation with Imperfect Instruments, Sociological Methodology, 34(1), 271-310.

Dunning, J. H. (1998): Location and the Multinational Enterprise: A Neglected Factor?, Journal of International Business Studies, 29(1), 45-66.

Dutz, M. A. (1989): Horizontal Mergers in Declining Industries, International Journal of Industrial Organisation, 7(1), 11-37.

Efron, B. (1979): Bootstrap Methods: Another Look at the Jackknife, The Annals of Statistics, 7(1), 1-26.

Efron, B. (1990): More Efficient Bootstrap Computations, Journal of the American Statistical Association, 85, 79-89.

Essama-Nssah, B. (2006): Propensity Score Matching and Policy Impact Analysis: A Demonstration in EViews, World Bank Policy Research, Working Paper 3877.

Fahrmeir, L./ Künstler, R./ Pigeot, I./ Tutz, G. (2009): Statistik - Der Weg zur Datenanalyse, 5. verb. Aufl., Berlin, Heidelberg, New York.

Fama, E. F./ Fisher, L./ Jensen, M. C./ Roll, R. W. (1969): The Adjustment of Stock Prices to New Information, International Economic Review, 10(1), 1-21.

Farrell, J./ Shapiro, C. (1990): Horizontal Mergers: An Equilibrium Analysis, The American Economic Review, 80(1), 107-126.

Farrell, J./ Shapiro, C. (2001): Scale Economies and Synergies in Horizontal Merger Analysis, Antitrust Law Journal, 68(3), 685-711.

Fischer, G./ Janik, F./ Müller, D./ Schmucker, A. (2009): The IAB Establishment Panel * Things Users Should Know., Schmollers Jahrbuch. Zeitschrift für Wirtschafts- und Sozialwissenschaften, 129(1), 133-148.

Fox, J. (1997): Applied Regression Analysis, Linear Models, and Related Methods, Thousand Oaks (CA), London.

Franks, J. R./ Mayer, C. (1996): Hostile Takeovers and the Correction of Management Failure, Journal of Financial Economics, 40(1), 163-181.

Gelman, A./ Hill, J. (2007): Data Analysis Using Regression and Multilevel/Hierarchical Models, New York.

Gensler, S./ Skiera, B./ Böhm, M. (2005): Einsatzmöglichkeiten der Matching Methode zur Berücksichtigung von Selbstselektion, Journal für Betriebswirtschaft, 55(1), 37-62.

Gioia, C./ Thomsen, S. (2004): International Acquisitions in Denmark 1990-1997: Selection and Performance, Applied Economics Quarterly, 1(1), 61-88.

Girma, S. (2005): Safeguarding Jobs? Acquisition FDI and Employment Dynamics in U.K. Manufacturing, Review of World Economics, 141(1), 165-178.

Girma, S./ Görg, H. (2004): Blessing or Curse? Domestic Plants' Employment and Survival Prospects after Foreign Acquisition, Applied Economics Quarterly, 50, 89-110.

Girma, S./ Görg, H. (2006): Multinationals' Productivity Advantage: Scale or Technology, CEPR Discussion Paper 5841.

Girma, S./ Görg, H. (2007): Evaluating the Foreign Ownership Wage Premium Using a Difference-in-Differences Matching Approach, Journal of International Economics, 72(1), 97-112.

Girma, S./ Görg, H./ Wagner, J. (2009): Subsidies and Exports in Germany: First Evidence from Enterprise Panel Data, Applied Economics Quarterly, 55(3), 179-195.

Girma, S./ Thompson, S./ Wright, P. (2006): International Acquisitions, Domestic Competition and Firm Performance, International Journal of the Economics of Business, Taylor and Francis Journals, 13(3), 335-349.

Glaum, M./ Hutzschenreuter, T. (2010): Mergers & Acquisitions - Management des externen Unternehmenswachstums, Stuttgart.

Goldberg, L. G. (1973): The Effect of Conglomerate Mergers on Competition, Journal of Law and Economics, 16(1), 137-158.

Görg, H. (2000): Analysing Foreign Market Entry: The Choice Between Greenfield Investment and Acquisitions, Journal of Economic Studies, 27(3), 165-181.

Görg, H./ Henry, M./ Strobl, E. (2007): Grant Support and Exporting Activity, CEPR Discussion Paper 6054.

Gort, M. (1969): An Economic Disturbance Theory of Mergers, Quarterly Journal of Economics, 83(4), 624-642.

Graßhoff, U./ Schwalbach, J. (1997): Managervergütung und Unternehmenserfolg, Zeitschrift für Betriebswirtschaft, 67(2), 203-217.

Greenaway, D./ Kneller, R. (2007): Firm Heterogeneity, Exporting and Foreign Direct Investment: A Survey, Economic Journal, Royal Economic Society, 117(517), F134-F161.

Greene, W. H. (2011): Econometric Analysis, 7th ed., New Jersey.

Griffith, R./ Simpson, H. (2004): Characteristics of Foreign-Owned Firms in British Manufacturing, NBER Chapters, in Seeking a Premier Economy: The Economic Effects of British Economic Reforms, 1980-2000, 147-180.

Grimpe, C. (2007): Der ZEW-ZEPHYR M&A-Index: Konzeption und Berechnung eines Barometers für weltweite Fusions- und Akquisitionstätigkeit, ZEW Dokumentation Nr. 07-01.

Grossman, S./ Hart, O. (1986): The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration, Journal of Political Economy, 94(4), 691-719.

Gugler, K./ Mueller, D. C./ Yurtoglu, B. B./ Zulehner, C. (2003): The Effects of Mergers: An International Comparison, International Journal of Industrial Organization, 21(5), 625-653.

Gugler, K./ Yurtoglu, B. B. (2004): The Effects of Mergers on Company Employment in the USA and Europe, International Journal of Industrial Organization, 22, 481-502.

Hagen, T./ Steiner, V. (2000): Von der Finanzierung der Arbeitslosigkeit zur Förderung von Arbeit - Analysen und Handlungsempfehlungen zur Arbeitsmarktpolitik, in Schriftenreihe des ZEW, 51, Baden-Baden.

Harris, R. S./ Stewart, J. F./ Carleton, W. T. (1982): Financial Characteristics of Acquired Firms, in Keenan, M./ White, L. J. (Eds.): Mergers and Acquisitions: Current Problems in Perspective, Lexington, 223-241.

Harris, R./ Robinson, C. (2002): The Effect of Foreign Acquisitions on Total Factor Productivity: Plant-Level Evidence From U.K. Manufacturing, 1987-1992, Review of Economics and Statistics, 84(3), 562-568.

Hartford, J. (2005): What Drives Merger Waves?, Journal of Financial Economics, 77(3), 529-560.

Healy, P. M./ Palepu, K. G./ Ruback, R. S. (1992): Does Corporate Performance Improve after Mergers?, Journal of Financial Economics, 31(2), 135-175.

Healy, P. M./ Palepu, K. G./ Ruback, R. S. (1997): Which Takeovers are Profitable? Strategic of Financial?, MIT Sloan Management Review, 38(4), 45-57.

Heckman, J. J./ LaLonde, R. J./ Smith, J. A. (1999): The Economics and Econometrics of Active Labour Market Programs, in Ashenfelter, O. C./ Card, D. (Eds.): Handbook of Labour Economics, 3A, Amsterdam, 1865-2097.

Heckman, J. J./ Ichimura, H./ Smith, J. A./ Todd P. (1998): Characterizing Selection Bias Using Experimental Data, Econometrica, 66(5), 1017-1098.

Heckman, J. J./ Ichimura, H./ Todd, P. (1997): Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program, Review of Economic Studies, 64(4), 605-654.

Heckman, J. J./ Robb, R. (1985): Alternative Methods for Evaluating the Impact of Interventions, in Heckman, J. J./ Singer, B. (Eds.): Longitudinal Analysis of Labour Market Data, New York, 156-245.

Heckman, J. J./ Smith, J. A./ Clements, N. (1997): Making the Most Out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts, Review of Economic Studies, 64, 487-535.

Helpman, E./ Melitz, M./ Yeaple, S. (2004): Exports versus FDI with Heterogeneous Firms, American Economic Review, 94(1), 300-316.

Hijzen, A./ Görg, H. / Manchin, M. (2008): Cross-Border Mergers and Acquisitions and the Role of Trade Costs, European Economic Review, 52(5), 849-866.

Hijzen, A./ Inui, T./ Todo, Y. (2007): The Effects of Multinational Production on Domestic Performance: Evidence from Japanese Firms, RIETI Discussion Paper Series 07-E-006.

Hirano, K./ Imbens, G. W. (2004): The Propensity Score with Continuous Treatments, in Gelman, A./ Meng, X. (Eds.): Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives, Chichester.

Hirshleifer, D./ Thakor, A. (1994): Managerial Performance, Boards of Directors and Takeover Bidding, Journal of Corporate Finance, 1(1), 63-90.

Holland, P. (1986): Statistics and Causal Inference, Journal of the American Statistical Associations, 81(396), 945-960.

Horn, H./ Persson, L. (2001): The Equilibrium Ownership of an International Oligopoly, Journal of International Economics, 53(2), 307-333.

Hosmer, D. W./ Lemeshow (2000): Applied Logistic Regression, 2. ed., New York.

Hughes, A. (1989): The Impact of Merger: A Survey of Empirical Evidence for the UK, in Fairburn, J. A./ Kay, J. A. (Eds.): Mergers and Merger Police, Oxford.

Hujer, R./ Caliendo, M./ Radic, D. (2001): Nobody Knows... How Do Different Evaluation Estimators Perform in a Simulated Labour Market Experiment?, Diskussionspapier der Universität Frankfurt/Main.

Huttunen, K. (2007): The Effect of Foreign Acquisition on Employment and Wages: Evidence from Finnish Manufacturing, The Review of Economics and Statistics, 89(3), 497-509.

Ikeda, K./ Doi, N. (1983): The Performance of Merging Firms in Japanese Manufacturing Industry 1964-75, Journal of Industrial Economics, 31(3), 257-266.

Imai K./ van Dyk, D. A. (2004): Causal Inference with General Treatment Regimes: Generalizing the Propensity Score, Journal of the American Statistical Association, 99(467), 854-866.

Imbens, G. W. (2000): The Role of the Propensity Score in Estimating Dose-Response Functions, Biometrika, 87(3), 706-710.

Imbens, G. W. (2004): Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review, The Review of Economics and Statistics, 86(1), 4-29.

Jäckle, R./ Wamser, G. (2010): Going Multinational: What are the Effects on Home-Market Performance?, German Economic Review, 11(2), 188-207.

Jansen, S. (2008): Mergers & Acquisitions: Unternehmensakquisitionen und -kooperationen. Eine strategische, organisatorische und kapitalmarkttheoretische Einführung, 5. überarb. und erw. Aufl., Wiesbaden.

Jensen, M. C. (1986): Agency Costs of Free Cash Flow, Corprate Finance, and Takeovers, American Economic Review, 76(2), 323-329.

Jensen, M. C. (1988): Takeovers: Their Causes and Consequences, Journal of Economic Perspectives, 2(1), 21-48.

Jensen, M. C./ Murphy, K. J. (1990): CEO Incentives - It's Not How Much You Pay, But How, Harvard Business Review, 68(3), 138-149.

Jensen, M. C./ Ruback, R. S. (1983): The Market for Corporate Control: The Scientific Evidence, Journal of Financial Economics, 11(1-4), 5-50.

Jovanovic, B. (1979): Job Matching and the Theory of Turnover, The Journal of Political Economy, 87(5), 972-990.

Kamien, M. I./ Zang, I. (1990): The Limits of Monopolization Through Acquisition, Quarterly Journal of Economics, 105(2), 465-499.

Kini, O./ Kracawb, W./ Mian, S. (1995): Corporate Takeovers, Firm Performance, and Board Composition, Journal of Corporate Finance, 1(3-4), 383-412.

Kirchner, M. (1991): Strategisches Akquisitionsmanagement im Konzern, Wiesbaden.

Kleinert, J./ Klodt, H. (2002): Fusionswellen und ihre Ursachen, in Franz, W./ Ramser, H. J./ Stadler, M. (Hrsg): Fusionen, Wirtschaftswissenschaftliches Seminar Ottobeuren, 27-49.

Kohler, U./ Kreuter, F. (2008): Datenanalyse mit Stata, 3. Aufl., München.

Kölling, A./ Schank, T. (2002): Skill-Biased Technological Change, International Trade and the Wage Structure. New Evidence on the Determinants of the Employment Structure from Linked Employer-Employee Panel Data for Germany, 10th Conference on Panel Data in Aarhus, Juli 6/7th 2002.

Krafft, M. (1997): Der Ansatz der logistischen Regression und seine Interpretation, Zeitschrift für Betriebswirtschaft, 67(5/6), 625-542.

Kumar, M. S. (1985): Growth, Acquisition Activity and Firm Size: Evidence from the United Kingdom, Journal of Industrial Economics, 33(3), 327-338.

Kunisch, S./ Wahler, C. (2010): Deutscher M&A-Markt im "Tal der Tränen" - Rückblick auf das M&A-Geschehen im Jahr 2009, M&A-REVIEW, 21(2), 53-62.

LaLonde, R. (1986): Evaluating the Econometric Evaluations of Training Programs with Experimental Data, American Economic Review, 76(4), 604-620.

Lechner, M. (1999): Earnings and Employment Effects of Continuous Off-the-Job Training in East Germany after Unification, Journal of Business and Economic Statistics, 17(1), 74-90.

Lechner, M. (2008): A Note on the Common Support Problem in Applied Evaluation Studies, Annales d'Économie et de Statistique, 91/92, 217-235.

Lehto, E. (2006): Motives to Restructure Industries Finnish Evidence from Cross-Border and Domestic Mergers and Acquisitions, Papers in Regional Science, 85(1), 1-22.

Lehto, E./ Böckerman, P. (2008): Analysing the Employment Effects of Mergers and Acquisitions, Journal of Economic Behavior and Organization, 68(1), 112-124.

Leuven, E./ Sianesi, B. (2003): PSMATCH2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing, Version 3.0.0.
http://ideas.repec.org/c/boc/bocode/s432001.html
May 10th 2010, 13.44h.

Lichtenberg, F. R. (1992): Industrial De-Diversification and its Consequences for Productivity, Journal of Economic Behavior and Organization, 18(3), 427-438.

Lichtenberg, F. R./ Siegel, D. (1990): The Effect of Ownership Changes on the Employment and Wages of Central Office and Other Personnel, Journal of Law and Economics, 33, 383-408.

Lichtenberg, F. R. / Siegel, D. (1992a): Productivity and Changes in Ownership of Manufacturing Plants, in Lichtenberg, F. R. (Ed.): Corporate Takeovers and Productivity, Cambridge, MA, 25-43.

Lichtenberg, F. R./ Siegel, D. (1992b): Takeovers and Corporate Overhead, in Lichtenberg, F. R. (Ed.): Corporate Takeovers and Productivity, Cambridge, Cambridge, MA, 45-67.

Lindbeck, A./ Snower, D. J. (2000): Multitask Learning and the Reorganization of Work: From Tayloristic to Holistic Organization, Journal of Labor Economics, 18(3), 353-376.

Lipsey, R. E./ Sjöholm, F. (2003): Foreign Firms and Indonesian Manufacturing Wages: An Analysis with Panel Data, NBER Working Paper 9417.

Loughran, T./ Vijh, A. M. (1997): Do Long-Term Shareholders Benefit From Corporate Acquisitions?, Journal of Finance, 52(5), 1765-1790.

Lucks, K./ Meckl, R. (2002): Internationale Mergers & Acquisitions. Der prozessorientierte Ansatz, Berlin.

Lyons, B. R. (2001): What Do We Conclude from the Success and Failure of Mergers?, Journal of Industry, Competition and Trade, 1(4), 411-422.

Maksimovic, V./ Phillips, G. (2001): The Market for Corporate Assets: Who Engages in Mergers and Asset Sales and Are There Efficiency Gains?, Journal of Finance, 56(6), 2019-2065.

Maksimovic, V./ Phillips, G./ Prabhala, N. R. (2011): Post-Merger Restructuring and the Boundaries of the Firm, Journal of Financial Economics, 102(2), 317-343.

Manne, H. G. (1965): Mergers and the Market for Corporate Control, Journal of Political Economy, 73, 110-120.

Margolis, D. N. (2006a): Compensation Policy, Human Resource Management Practices and Takeovers, in Bryson, A./ Forth, J./ Barber, C. (Eds.): Making Linked Employer-Employee Data Relevant to Policy, DTI Occasional Paper 4.

Margolis, D. N. (2006b): Should Employment Authorities Worry About Mergers and Acquisitions?, Portuguese Economic Journal, 5(2), 167-194.

Markowitz, H. (1952): Portfolio Selection, The Journal of Finance, 7(1), 77-91.

Markusen, J. (2004): Multinational Firms and the Theory of International Trade, Cambridge.

Marris, R. (1963): A Model of the Managerial Enterprise, Quarterly Journal of Economics, 77(2), 185-209.

Marris, R. (1964): The Economic Theory of Managerial Capitalism, Glencoe.

Martin, K. J./ McConnell, J. J. (1991): Corporate Performance, Corporate Takeovers and Management Turnover, Journal of Finance, 46(2), 671-688.

Mattes, A. (2010): International M&A: Evidence on Effects of Foreign Takeovers, IAW Discussion Papers 60.

Mayer, T./ Ottaviano, G. I. P. (2007): The Happy Few: The Internationalisation of European Firms. New Facts Based on Firm-Level Evidence, Bruegel Blueprint Series 3.

McDougall, F. M./ Round, D. K. (1986): The Determinants and Effects of Corporate Takeovers in Australia, 1970-1981, Victoria, Australian Institute of Management.

McGuckin, R./ Nguyen, S. V. (1995): On Productivity and Plant Ownership Change: New Evidence from the Longitudinal Research Database, RAND Journal of Economics, 26(2), 257-276.

McGuckin, R./ Nguyen, S. V. (2001): The Impact of Ownership Changes: A View from Labour Markets, International Journal of Industrial Organization, 19(5), 739-762.

McGuckin, R./ Nguyen, S. V./ Reznek, A. P. (1995): The Impact of Ownership Change on Employment, Wages and Labor Productivity in US Manufacturing 1977-1987, Center for Economic Studies, US Bureau of the Census, 95-98.

McGuckin, R./ Nguyen, S. V./ Reznek, A. P. (1998): On the Impact of Ownership Change on Labor: Evidence from Food Manufacturing Plant Level Data, in Haltiwanger, J./ Manser, M./ Toppel, R. (Eds.): Labor Statistics Measurement, National Bureau of Economic Research, Studies in Income and Wealth, 60, Chicago and London, 207-246.

Meeks, G. (1977): Disappointing Mariagge: A Study of the Gains from Merger, Cambridge.

Melicher, R. W./ Rush, D. F. (1973): The Performance of Conglomerate Firms: Recent Risk and Return Experience, Journal of Finance, 28(2), 381-388.

Melicher, R. W./ Rush, D. F. (1974): Evidence on the Acquisition-Releated Performance of Conglomerate Firms, Journal of Finance, 29(1), 141-149.

Melitz, M. (2003): The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity, Econometrica, 71(6), 1695-1725.

Menard, S. (2001): Applied Logistic Regression Analysis, 2nd ed., Thousand Oaks.

Morck, R./ Shleifer, A./ Summers, L. (1988): The Characteristics of Hostile Takeovers, in Auerbach, A. J. (Ed.): Corporate Takeovers: Causes and Consequences, Chicago and London.

Morck, R./ Shleifer, A./ Summers, L. (1990): Do Managerial Objectives Drive Bad Acquisitions?, Journal of Finance, 45(1), 31-48.

Mueller, D. C. (1980a): The Determinants and Effects of Mergers: An International Comparison, Cambridge.

Mueller, D. C. (1980b): The United States, 1962-1972, in Mueller, D.C. (Ed.): The Determinanats and Effects of Mergers: An International Comparison, Cambridge, 271-298.

Mueller, D. C. (1985): Mergers and Market Shares, Review of Economics and Statistics, 67(2), 259-267.

Mueller, D. C. (1986): Profits in the Long Run, Cambridge.

Mueller, D. C. (2003a): The Corporation: Investment, Mergers, and Growth, London.

Mueller, D. C. (2003b): The Finance Literature on Mergers: A Critical Survey, in Waterson, M. (Eds.): Competition, Monopoly and Corporate Governance, Essays in Honour of Keith Cowling, Cheltenham: Edward Elgar, 161-205.

Mueller, D. C./ Gugler, K./ Weichselbaumer, M. (2012): The Determinants of Merger Waves: An International Perspective, International Journal of Industrial Organization, 30(1), 1-15.

Mueller, D. C./ Sirower, M. L. (2003): The Causes of Mergers: Tests Based on the Gains to the Acquiring Firm's Shareholders and the Size of Premia, Managerial and Decision Economics, 24(5), 373-391.

Müller-Stewens, G./ Kunisch, S./ Binder, A. (2010): Mergers & Acquisitions. Analysen, Trends und Best Practices, Stuttgart.

Nalebuff, B. (2003): Bundling, Tying, and Portfolio Effects, DTI Economics Paper 1.

Neary, P. J. (2007): Cross-Border Mergers as Instruments of Comparative Advantage, Review of Economic Studies, 74(4), 1229-1257.

Nelson, R. L. (1959): Merger Movements in American Industry, 1895-1956, Princton.

Nickell, S. J. (1996): Competition and Corporate Performance, Journal of Political Economy, 104(4), 724-746.

Nocke, V./ White, L. (2007): Do Vertical Mergers Facilitate Upstream Collusion?, American Economic Review, 97(4), 1321-1339.

Nocke, V./ White, L. (2010): Vertical Merger, Collusion, and Disruptive Buyers, International Journal of Industrial Organization, 28, 350-354.

Nocke, V./ Yeaple, S. (2007): Cross-Border Mergers and Acquisitions vs. Greenfield Foreign Direct Investment: The Role of Firm Heterogeneity, Journal of International Economics, 72(2), 336-365.

Ollinger, M./ Nguyen, S. V./ Blayney, D./ Chambers, B./ Nelson, K. (2006): Food Industry Mergers and Acquisitions Lead to Higher Labour Productivity, United States Department of Agriculture, Economic Research Report 27.

Ollinger, M./ Nguyen, S. V./ Blayney, D./ Nelson, K./ Chambers, B.(2005): Effect of Food Industry Mergers and Acquisitions on Employment and Wages, United States Department of Agriculture, Economic Research Report 13.

Pagan, A./ Ullah, A. (1999): Nonparametric Econometrics, New York.

Pausenberger, E. (1989): Zur Systematik von Unternehmenszusammenschlüssen, Das Wirtschaftsstudium, 18(11), 621-626.

Peer, H. (1980): The Netherlands, 1962-1973, in Mueller, D.C. (Ed.): The Determinants and Effects of Mergers: An International Comparison, Cambridge, 163-191.

Perry, M./ Porter, R. H. (1985): Oligopoly and the Incentive for Horizontal Merger, American Economic Review, 75(1), 219-227.

Pesendorfer, M. (2003): Horizontal Mergers in the Paper Industry, RAND Journal of Economics, 34(3), 495-515.

Petkova, N. (2009): Essays on Firm Ownership, Performance and Value, Dissertation, University of Michigan.

Piscitello, L./ Rabbiosi, L. (2005): The Impact of Inward FDI on the Local Companies' Labour Productivity: Evidence from the Italian Case, International Journal of the Economics of Business, 12(1), 35-51.

Porter, M. E. (1992): Wettbewerbsvorteile - Spitzenleistungen erreichen und behaupten, 4. Aufl., Frankfurt am Main.

Ravenscraft D. J./ Scherer, F. M. (1987): Mergers, Sell-Offs, and Economic Efficiency, Washington.

Reid, S. R. (1971): A Reply to the Weston/Mansinghka Criticisms Dealing with Conglomerate Mergers, Journal of Finance, 26(4), 937-946.

Reinowski, E. (2008): Matching kleiner Stichproben. Ein Vergleich verschiedener Verfahren, Dissertation, Martin-Luther-Universität Halle-Wittenberg, Saarbrücken.

Rhoades, S. A. (1987): The Operating Performances of Acquired Firms in Banking, in Wills, R. L./ Caswell, J. A./ Culvertson, J. D. (Eds.): Issues After a Century of Federal Competition Policy, Lexington, 277-292.

Rhodes-Kropf, M./ Robinson, D. T./ Viswanathan, S. (2005): Valuation Waves and Merger Activity: The Empirical Evidence, Journal of Financial Economics, 77(9), 561-603.

Riordan, M. H./ Salop, S. C. (1995): Evaluating Vertical Mergers: A Post-Chicago Approach, Antitrust Law Journal, 63(2), 513-568.

Rohrlack, C. (2009): Logistische und Ordinale Regression, in Albers, S./ Klapper, D./ Konradt, U./ Walter, A./ Wolf, J. (Eds.): Methodik der empirischen Forschung, 3. Aufl., Wiesbaden, 267-282.

Roll, R. (1986): The Hubris Hypothesis of Corporate Takeovers, Journal of Business, 59(2), 197-216.

Rosenbaum, P. R. (2002): Observational Studies, 2nd ed., New York.

Rosenbaum, P. R./ Rubin, D. B. (1983): The Central Role of the Propensity Score in Observational Studies for Causal Effects, Biometrika, 70(1), 41-50.

Rosenbaum, P. R./ Rubin, D. B. (1985): Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score, The American Statistican, 39(1), 33-38.

Roy, A. (1951): Some Thoughts on the Distribution of Earnings, Oxford Economic Papers, 3(2), 135-145.

Rubin, D. B. (1974): Estimating Causal Effects to Treatments in Randomised and Nonrandomised Studies, Journal of Educational Psychology, 66(5), 688-701.

Rubin, D. B. (1990): Comment on: Neyman (1923) and Causal Inference in Experiments and Observational Studies, Statistical Science, 5(4), 472-480.

Rubin, D. B./ Thomas, N. (1996): Matching Using Estimated Propensity Scores: Relating Theory to Practice, Biometrics, 52(1), 249-264.

Ryden, B./ Edberg, J. O. (1980): Large Mergers in Sweden, 1962-1976, in Mueller, D. C. (Ed.): The Determinants and Effects of Mergers: An International Comparison, Cambridge, 193-226.

Salant, S. W. / Switzer, S./ Reynolds, R. J. (1983): Losses from Horizontal Merger: The Effects of an Exogenous Change in Industry Structure on Cournot-Nash Equilibrium, Quarterly Journal of Economics, 98(2), 185-199.

Salis, S. (2008): Foreign Acquisition and Firm Productivity: Evidence from Slovenia, World Economy, 31(8), 1030-1048.

Schank, T./ Schnabel, C./ Wagner, J. (2010): Higher Wages in Exporting Firms: Self-Selection, Export Effect, or Both? First Evidence from Linked Employer-Employee Data, Review of World Economics, 146(2), 303-322.

Scharfstein, D. (1988): The Disciplining Role of Takeovers, Review of Economic Studies, 55(2), 185-199.

Scherer, F. M. (2002): The Merger Puzzle, in Franz, W./ Ramser, H. J./ Stadler, M. (Eds.): Fusionen, Wirtschaftswissenschaftliches Seminar Ottobeuren, 1-22.

Schoar, A. (2002): Effects of Corporate Diversification on Productivity, The Journal of Finance, 67(6), 2379-2403.

Schwert, G. W. (2000): Hostility in Takeovers: In the Eyes of the Beholder?, Journal of Finance, 55(6), 2599-2640.

Seth, A. (1990): Value Creation in Acquisitions: A Reexamination of Performance Issues, Strategic Management Journal, 11(2), 99-105.

Shleifer, A./ Summers, L. (1988): Breach of Trust in Hostile Takeovers, in Auerbach, A. J. (Ed.): Corporate Takeovers: Causes and Consequences, Chicago and London.

Shleifer, A./ Vishny, R. W. (1989): Management Entrenchment: The Case of Manager-Specific Investments, Journal of Financial Economics, 25(1), 123-140.

Sianesi, B. (2004): An Evaluation of the Swedish System of Active Labor Market Programs in the 1990s, Review of Economics and Statistics, 86(1), 133-155.

Siegel, D. S./ Simons, K. L. (2008): Evaluating the Effects of Mergers and Acquisitions on Employees: Evidence from Matched Employer-Employee Data, Working Papers 2, Jerusalem Institute for Market Studies.

Siegel, D. S./ Simons, K. L. (2010): Assessing the Effects of Mergers and Acquisitions on Firm Performance, Plant Productivity, and Workers: New Evidence from Matched Employer-Employee Data, Strategic Management Journal, 31(8), 903-916.

Silverman, B. W. (1986): Density Estimation, London.

Smith, J. A./ Todd, P. (2005a): Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators?, Journal of Econometrics, 125(1-2), 305-353.

Smith, J. A./ Todd, P. (2005b): Rejoinder, Journal of Econometrics, 125(1-2), 365-375.

Spanninger, J. (2011a): Hinkt Deutschland der Welt hinterher? - Jahresrückblick auf das deutsche M&A-Geschehen 2010, M&A-REVIEW, 22(2), 49-57.

Spanninger, J. (2011b): M&A-Geschehen (bisher) nicht durch Krise(n) gefährdet - Deutsche M&A-Aktivitäten im 1. Halbjahr 2011, M&A-REVIEW, 22(9), 356-365.

Spanninger, J. (2012): Gut angefangen und stark nachgelassen - Jahresrückblick auf das deutsche M&A-Geschehen 2011, M&A-REVIEW, 23(2), 42-51.

Spearot, A. C. (2007a): Firm Heterogeneity and Acqusition Incentives, mimeo University of Wisconsin.

Spearot, A. C. (2007b): Acqusition, Productivity, and Trade Liberalization, mimeo University of California - Santa Cruz.

Spengler, J. (1950): Vertical Integration and Antitrust Policy, Journal of Political Economy, 53, 347-352.

Steiner, P. O. (1975): Mergers: Motives, Effects, Policies, Ann Arbor.

Stephan, G. (2008): The Effects of Active Labour Market Programs in Germany - An Investigation Using Different Definitions of Non-Treatment, Jahrbücher für Nationalökonomie und Statistik, 228, 586-611.

Tichy, G. (2001): What Do We Know About Success and Failure of Mergers?, Journal of Industry, Competition and Trade, 1(4), 347-394

Town, R. J. (1992): Merger Waves and the Structure of Merger and Acquisition Time-Series, Journal of Applied Econometrics, 7(S), 83-100.

Uhlenbruck, K. (2004): Developing Acquired Foreign Subsidiaries: The Experience of MNEs in Transition Economies, Journal of International Business Studies, 35(2), 109-123.

UNCTAD (2012): World Investment Report (WIR) 2011, Towards a New Generation of Investment Policies, United Nations Conference on Trade and Development, Geneva.

Verbeek, M. (2005): A Guide to Modern Econometrics, 2nd ed., Chichester.

Vogel, D. H. (2002): M&A. Ideal und Wirklichkeit, Wiesbaden.

Wagner, J. (2002): The Causal Effect of Exports on Firm Size and labour Productivity: First Evidence from a Matching Approach, Economic Letters, 77(2), 287-292.

Wagner, J. (2007a): Exports and Productivity: A Survey of the Evidence from Firm-level Data, The World Economy, 30(1), 60-82.

Wagner, J. (2007b): Export Entry, Export Exit, and Productivity in German Manufacturing Industries, Working Paper Series in Economics 54, University of Lüneburg, Institute of Economics.

Weston, J. F. (1970): The Nature and Significance of Conglomerate Firms, St. John's Law Review, 44, 66-80.

Weston, J. F./ Mansinghka, S. K. (1971): Tests of the Efficiency Performance of Conglomerate Firms, Journal of Finance, 26(4), 919-936.

Williamson, O. E. (1970): Corporate Control and Business Behavior: An Inquiry into the Effects of Organization Form on Enterprise Behavior, Englewood Cliffs, New Jersey.

Williamson, O. E. (1975): Markets and Hierarchies: Analysis and Antitrust Implications, New York.

Williamson, O. E. (1988): Mergers, Acquisitions, and Leveraged Buyouts: An Efficiency Assessment, in Libecap, G. (Ed.): Corporate Reorganization Through Mergers, Acquisition, and Leveraged Buyouts, Greenwich, 55-80.

Wirtz, B. W. (2003): Mergers & Acquisitions Management. Strategie und Organisation von Unternehmenszusammenschlüssen, Wiesbaden.